

also im Minimum der Gradient verschwinden. Die partielle Ableitung nach  $x_k$  ist

$$\sum_{i=1}^N 2a_{ik} \left( \sum_{j=1}^m a_{ij} x_j - b_i \right) = 2 \sum_{j=1}^m \sum_{i=1}^N a_{ik} a_{ij} x_j - 2 \sum_{i=1}^N a_{ik} b_i.$$

Falls es ein Minimum gibt, muß dieses also eine Lösung des linearen Gleichungssystems

$$\sum_{j=1}^m \left( \sum_{i=1}^N a_{ik} a_{ij} \right) x_j = \sum_{i=1}^N a_{ik} b_i \quad \text{für alle } k = 1, \dots, m$$

aus  $m$  Gleichungen in  $m$  Unbekannten sein.

Schreiben wir das Ausgangssystem in Matrixform als  $Ax = b$  mit

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{Nm} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix},$$

so läßt sich das neue Gleichungssystem schreiben als

$$(A^T A)x = A^T b,$$

wir müssen also das ursprüngliche System einfach mit der transponierten Matrix

$$A^T = \begin{pmatrix} a_{11} & \cdots & a_{N1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \cdots & a_{Nm} \end{pmatrix}$$

multiplizieren.

Von der Problemstellung her ist klar, daß ein lokales Extremum einer Summe von Quadraten nur ein Minimum sein kann. Trotzdem wollen wir zur Vorsicht noch das Kriterium aus dem vorigen Abschnitt überprüfen, also die HESSE-Matrix auf Definitheit untersuchen.

Leiten wir die partielle Ableitung der Quadratsumme nach  $x_k$  weiter ab nach  $x_\ell$ , erhalten wir

$$2 \sum_{i=1}^N a_{ik} a_{i\ell},$$

die HESSE-Matrix ist also gerade das Doppelte der Matrix  $A^T A$ . Für Definitheitseigenschaften ist der Faktor zwei natürlich ohne Bedeutung; wir können also die quadratische Form zu  $A^T A$  betrachten. Für  $x \in \mathbb{R}^m$  ist

$$x^T (A^T A)x = (x^T A^T)(Ax) = (Ax)^T (Ax)$$

das Skalarprodukt des Vektors  $Ax \in \mathbb{R}^N$  mit sich selbst und somit nie negativ. Es verschwindet genau dann, wenn  $Ax$  der Nullvektor in  $\mathbb{R}^N$  ist, also insbesondere natürlich, wenn  $x$  der Nullvektor in  $\mathbb{R}^m$  ist. Ob es Vektoren  $x \neq 0$  gibt mit  $Ax = 0$  hängt ab von der Matrix  $A$ : Ist  $M < m$ , muß es solche Vektoren geben; die Matrix  $A^T A$  ist also nur positiv semidefinit.

Wenn es um lineare Regression geht, ist dagegen  $N$  im allgemeinen deutlich größer als  $m$ ; hier wird es nur sehr selten vorkommen, daß es Vektoren  $x \neq 0$  aus  $\mathbb{R}^m$  gibt mit  $Ax = 0$ . Mit den aus der Linearen Algebra bekannten Begriffen können wir das auch exakt formulieren: Genau dann, wenn die Matrix  $A$  kleineren Rang als  $m$  hat, gibt es solche Vektoren. Andernfalls ist  $Ax = 0$  nur für  $x = 0$ , die Matrix  $A^T A$  ist also positiv definit. Da dann insbesondere auch ihre Determinante nicht verschwindet, folgt:

**Satz:**  $Ax = b$  sei ein lineares Gleichungssystem aus  $N$  Gleichungen in  $m$  Unbekannten. Falls die Matrix  $A$  den Rang  $m$  hat, gibt es genau einen Vektor  $x \in \mathbb{R}^m$ , für den die (EUKLIDISCHE) Länge von  $Ax - b$  minimal wird. Er ist die eindeutig bestimmte Lösung des linearen Gleichungssystems  $A^T Ax = A^T b$ . ■

Als erstes Beispiel wollen wir den bekanntesten Spezialfall der linearen Regression betrachten, die *Ausgleichsgerade*. Hier geht es um  $N$  Datenpaare  $(x_i, y_i)$ , zwischen denen wir einen Zusammenhang der Form  $y_i \approx ax_i + b$  vermuten; gesucht sind diejenigen Werte  $a$  und  $b$  für die der Differenzvektor zwischen linker und rechter Seite minimale Länge hat. Im Gegensatz zu unserer sonstigen Konvention bezeichnen hier also  $x_i$  und  $y_i$  bekannte Größen, während  $a$  und  $b$  gesucht sind.

Das lineare Gleichungssystem, das wir näherungsweise lösen wollen, ist daher ein Gleichungssystem mit den Unbekannten  $a$  und  $b$ ; es besteht

aus den  $N$  Gleichungen

$$x_i a + b = y_i \quad \text{für } i = 1, \dots, N.$$

Matrix und rechte Seite sind daher

$$A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}.$$

Nach dem gerade bewiesenen Satz müssen wir das System  $A \begin{pmatrix} a \\ b \end{pmatrix} = y$  von links mit der transponierten Matrix  $A^T$  multiplizieren; da

$$A^T A = \begin{pmatrix} x_1 & x_2 & \dots & x_N \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & N \end{pmatrix}$$

und

$$A^T y = \begin{pmatrix} x_1 & x_2 & \dots & x_N \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N x_i y_i \\ \sum_{i=1}^N y_i \end{pmatrix}$$

ist, müssen wir also das lineare Gleichungssystem

$$\begin{pmatrix} \sum_{i=1}^N x_i^2 \\ \sum_{i=1}^N x_i \end{pmatrix} a + \begin{pmatrix} \sum_{i=1}^N x_i \\ N \end{pmatrix} b = \begin{pmatrix} \sum_{i=1}^N x_i y_i \\ \sum_{i=1}^N y_i \end{pmatrix}$$

lösen. Nach obigem Satz hat es genau dann eine eindeutig bestimmte Lösung, wenn der Rang der Matrix  $A$  gleich zwei ist, wenn also der Vektor mit Komponenten  $x_i$  linear unabhängig ist vom Vektor, dessen sämtliche Komponenten Einsen sind. Das ist offenbar genau dann der Fall, wenn die  $x_i$  nicht allesamt denselben Wert haben

Falls alle  $x_i$  gleich sind, ist auch unabhängig von jeder Linearer Algebra klar, daß wir keine Aussage darüber machen können, wie sich  $y$  in

Abhängigkeit von  $x$  verändert – wir haben schließlich nur einen einzigen  $x$ -Wert.

Andernfalls sind  $a$  und  $b$  durch das obige Gleichungssystem eindeutig bestimmt; durch GAUSS-Elimination oder nach der CRAMERSchen Regel erhalten wir die Werte

$$a = \frac{N \sum_{i=1}^N x_i y_i - \left( \sum_{i=1}^N x_i \right) \left( \sum_{i=1}^N y_i \right)}{N \sum_{i=1}^N x_i^2 - \left( \sum_{i=1}^N x_i \right)^2}$$

und

$$b = \frac{\left( \sum_{i=1}^N x_i^2 \right) \left( \sum_{i=1}^N y_i \right) - \left( \sum_{i=1}^N x_i \right) \left( \sum_{i=1}^N x_i y_i \right)}{N \sum_{i=1}^N x_i^2 - \left( \sum_{i=1}^N x_i \right)^2}.$$

In konkreten Anwendungen dürfte es freilich einfacher sein, diese Formeln zu vergessen und stattdessen das lineare Gleichungssystem direkt zu lösen.

Diejenigen, die Ausgleichsgeraden aus der Schule kennen, hatten dort wohl vor allem mit Paaren  $(x_i, y_i)$  von Meßwerten zu tun, bei denen klar war, daß sie auf Grund eines Naturgesetzes in einem linearen Zusammenhang stehen sollten; die Abweichungen zwischen  $y_i$  und  $a x_i + b$  waren ausschließlich auf Meßfehler zurückzuführen.

In den Wirtschafts- und Sozialwissenschaften sind exakte Gesetze selten; trotzdem spielt Regression auch hier eine große Rolle, wenn es darum geht, ungefähre Zusammenhänge aufzustellen.

Als Beispiel betrachten wir das Problem der Korruption. Die vom ehemaligen Weltbankdirektor für Ostafrika PETER EIGEN 1993 gegründete Organisation *Transparency International* stellt jedes Jahr einen *Corruption Perceptions Index (CPI)* auf. Er beruht auf Befragungen vor allem von Geschäftsleuten, die in den untersuchten Ländern tätig sind und Auskunft geben sollen, für wie korrupt sie die dortige Regierung und Verwaltung halten. Für jedes Land wird auf Grund entsprechender Studien der letzten drei Jahre ein Indexwert zwischen 0 und 10 berechnet.

Eine Zehn bedeutet, daß den Befragten trotz ihrer Tätigkeit dort nichts über Korruption in diesem Land bekannt ist, eine Null entsprechend daß ohne Bimbes nichts geht.

Die neueste derzeit unter [www.transparency.org](http://www.transparency.org) verfügbare Liste ist die von 2009; sie enthält 180 Staaten mit Neuseeland (9,4) an der Spitze und Somalia (1,1) an letzter Stelle. Blättert man durch die Liste, drängt sich schnell der Eindruck auf, als seien am Ende vor allem arme Staaten zu finden, an der Spitze eher die reichen.

Um diese Hypothese zumindest für die 27 EU-Staaten quantitativ zu untersuchen, können wir deren CPI vergleichen mit dem Bruttoinlandsprodukt pro Einwohner in Kaufkraftstandards (KKS), wobei ein KKS so definiert ist, daß ein vorgegebener Warenkorb in jedem Land denselben Preis in KKS hat und im EU-Durchschnitt ein KKS genau einem Euro entspricht.

Die Bruttoinlandsprodukte pro Einwohner sind bei eurostat zu finden unter [ec.europa.eu/eurostat](http://ec.europa.eu/eurostat) zu finden. Die derzeitig aktuellsten mehr oder weniger vollständigen Daten stammen von 2008; lediglich für Rumänien liegt nur ein 2007er-Wert vor. (Der Wert für Griechenland ist, wie alle seit 2004, nur vorläufig.)

Die Tabelle auf der nächsten Seite listet für alle 27 Länder sowohl den CPI als auch das Bruttoinlandsprodukt per Einwohner auf; weiter hinten ist der Zusammenhang auch graphisch dargestellt.

Ganz offensichtlich liegen die Punkte nicht auf einer Geraden, und in der Tat gibt es keinen Grund für einen festen, deterministischen Zusammenhang zwischen den beiden Größen. Trotzdem sind sie auch nicht völlig unabhängig voneinander; zumindest tendenziell wird in den reicheren Staaten weniger Korruption wahrgenommen.

Versuchen wir, eine Ausgleichsgerade durch die obige Punktwolke zu legen! Unsere  $x_i$  sind die Bruttoinlandsprodukte per Einwohner, ihre Summe ist 655 800, die Summe ihrer Quadrate 22 520 220 000. Die  $y_i$  sind die CPIs; ihre Summe ist 171, 6 und die Summe der  $x_i y_i$  schließlich 4 717 610. Wir bekommen also das lineare Gleichungssystem

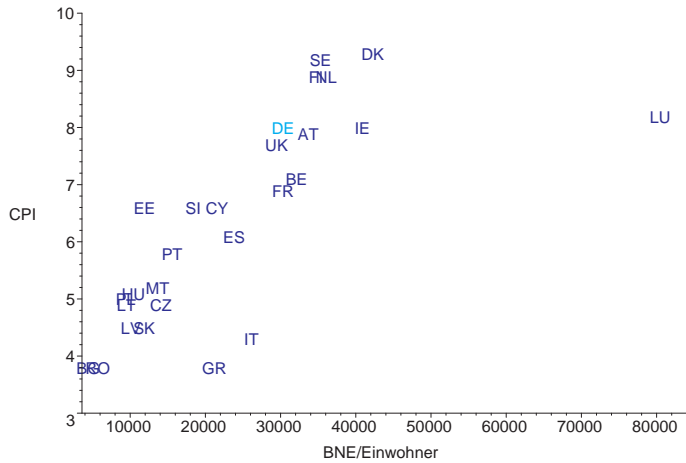
$$\begin{aligned} 22\,520\,220\,000 a + 655\,800 b &= 4\,717\,610 \\ 655\,800 a + 27 b &= 171,6 \end{aligned}$$

Land	BIP/Einwohner	CPI
Belgien	32200	7,1
Bulgarien	4500	3,8
Dänemark	42400	9,3
Deutschland	30400	8,0
Estland	12000	6,6
Finnland	34700	8,9
Frankreich	30400	6,9
Griechenland	21300	3,8
Irland	40900	8,0
Italien	26200	4,3
Lettland	10200	4,5
Litauen	9600	4,9
Luxemburg	80500	8,2
Malta	13800	5,2
Niederlande	36200	8,9
Österreich	33800	7,9
Polen	9500	5,0
Portugal	15700	5,8
Rumänien	5800	3,8
Schweden	35400	9,2
Slowakei	12000	4,5
Slowenien	18400	6,6
Spanien	23900	6,1
Tschechische Republik	14200	4,9
Ungarn	10500	5,1
Vereinigtes Königreich	29600	7,7
Zypern	21700	6,6

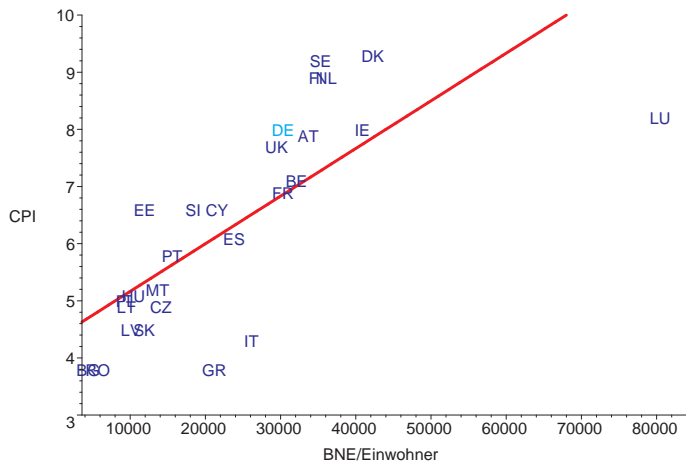
mit der Lösung

$$a = \frac{164891}{1977470000} \approx 8,338 \cdot 10^{-5} \quad \text{und} \quad b = \frac{128443519}{29662050} \approx 4,330.$$

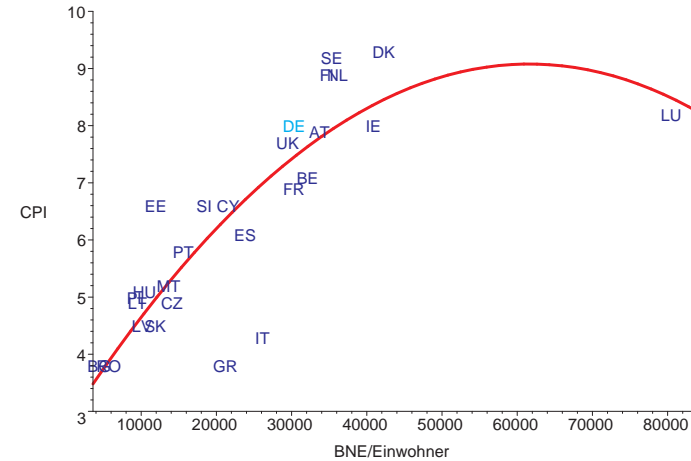
Die nächste Abbildung zeigt die Daten zusammen mit der Geraden  $\text{CPI} = a \cdot \text{BIP} + b$ . Sie wird zwar ihrem Namen *Ausgleichsgerade* durchaus gerecht, ist aber keine optimale Beschreibung des Zusammenhangs



zwischen den beiden Größen. Insbesondere fällt ins Auge, daß die Werte für Schweden und Dänemark ziemlich weit von der Geraden entfernt sind; noch schlimmer ist es bei Luxemburg, wo uns die Geradengleichung einen gar nicht existierenden (und auch nicht dargestellten) Wert von knapp über elf vorhersagen würde oder bei Griechenland und Italien, die ebenfalls sehr weit unterhalb der Gerade liegen.



Nun gibt es wie gesagt keinen vernünftigen Grund, daß es hier einen linearen Zusammenhang geben sollte: Da der CPI nur Werte bis zu zehn annehmen kann, es aber keine absolute Obergrenze für das Bruttonationaleinkommen pro Einwohner gibt, kann es in der Tat schon aus theoretischen Gründen keinen solchen Zusammenhang geben. Wir sollten daher versuchen, an Stelle einer Geraden eine andere Kurve zu finden, zum Beispiel eine Parabel:



Auch die bestmögliche „Ausgleichsparabel“ läßt sich mittels linearer Regression bestimmen: Wir suchen nach einem Zusammenhang der Art

$$y_i = ax_i^2 + bx_i + c \quad \text{für alle } i = 1, \dots, N,$$

und diese  $N$  Gleichungen liefern uns ein (im allgemeinen natürlich unlösbares) lineares Gleichungssystem für die drei Koeffizienten  $a, b$  und  $c$ . Einzelheiten seien dem Leser überlassen (*siehe Übungsblatt!*); hier ist deshalb nur das Ergebnis dargestellt. Selbstverständlich ist die Übereinstimmung auch hier alles andere als perfekt, aber sie sieht doch schon deutlich besser aus als im Falle der Geraden.

Dies führt uns auf die Frage, wie wir, möglichst schon vor der Berechnung einer Ausgleichsgeraden und ohne Aufzeichnen der Datenpunkte, entscheiden können, ob es sich überhaupt lohnt, nach einem linearen Zusammenhang zu suchen.

Dazu betrachten wir als ersten Extremfall  $N$  Paare  $(x_i, y_i)$ , zwischen denen ein perfekter linearer Zusammenhang besteht:  $y_i = ax_i + b$  für alle  $i$  mit  $a \neq 0$ . Für die Mittelwerte

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{und} \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

ist dann

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i = \frac{1}{N} \sum_{i=1}^N (ax_i + b) = a \cdot \frac{1}{N} \sum_{i=1}^N x_i + \frac{1}{N} \cdot Nb = a\bar{x} + b,$$

die Abweichung eines Werts  $y_i$  vom Mittelwert  $\bar{y}$  läßt sich also mittels der Formel

$$y_i - \bar{y} = a(x_i - \bar{x})$$

leicht aus der entsprechenden Abweichung  $x_i - \bar{x}$  berechnen. Insbesondere hat das Produkt

$$(y_i - \bar{y})(x_i - \bar{x}) = a(x_i - \bar{x})^2$$

für alle  $i$  mit  $x_i \neq \bar{x}$  dasselbe Vorzeichen, und zwar das Vorzeichen von  $a$ . Die Summe

$$\sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x}) = a \sum_{i=1}^N (x_i - \bar{x})^2 = \frac{1}{a} \sum_{i=1}^N (y_i - \bar{y})^2.$$

Somit ist

$$\left( \sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x}) \right)^2 = \left( \sum_{i=1}^N (y_i - \bar{y})^2 \right) \left( \sum_{i=1}^N (x_i - \bar{x})^2 \right)$$

oder, anders ausgedrückt,

$$\frac{\sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} = \begin{cases} 1 & \text{falls } a > 0 \\ -1 & \text{falls } a < 0 \end{cases}.$$

Als zweiten Extremfall betrachten wir Punktepaare  $(x_i, y_i)$ , bei denen keinerlei Zusammenhang zwischen  $x_i$  und  $y_i$  besteht. Dann sollte man, zumindest bei hinreichend vielen Wertepaaren, erwarten, daß auf eine

positive Differenz  $x_i - \bar{x}$  ungefähr gleich oft eine positive wie eine negative Differenz  $y_i - \bar{y}$  trifft und daß diese Differenzen im Mittel auch etwa denselben Betrag haben. Somit sollte

$$\sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x})$$

im Vergleich zum ersten Fall einen ziemlich kleinen Betrag haben; insbesondere sollte der Betrag des Quotienten

$$\frac{\sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$

deutlich kleiner sein als eins.

**Definition:** Für  $N$  Wertepaare  $(x_i, y_i)$ , bei denen weder alle  $x_i$  noch alle  $y_i$  denselben Wert haben, bezeichnen wir

$$\kappa \stackrel{\text{def}}{=} \frac{\sum_{i=1}^N (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$

als den *Korrelationskoeffizienten*.

Dieser Korrelationskoeffizient liegt immer zwischen  $-1$  und  $1$ , denn bezeichnen wir mit  $u \in \mathbb{R}^N$  den Vektor mit Komponenten  $x_i - \bar{x}$  und mit  $v \in \mathbb{R}^N$  den mit Komponenten  $y_i - \bar{y}$ , so ist offensichtlich

$$\kappa = \frac{\langle u, v \rangle}{\sqrt{\langle u, u \rangle} \sqrt{\langle v, v \rangle}} = \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|},$$

wobei  $\|\cdot\|$  die EUKLIDISCHE Norm bezeichnet, und nach der CAUCHY-SCHWARZschen Ungleichung ist

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\|$$

mit Gleichheit genau dann, wenn  $u$  und  $v$  linear abhängig sind, wenn also zwischen den  $x_i$  und den  $y_i$  ein linearer Zusammenhang besteht.

Somit ist  $\kappa = 1$  genau dann, wenn wir eine Gleichung  $y_i = ax_i + b$  mit  $a > 0$  haben,  $\kappa = -1$  bei einer entsprechenden Gleichung mit  $a < 0$ , und  $\kappa \approx 0$ , wenn es keinerlei Zusammenhang zwischen den  $x_i$  und den  $y_i$  gibt.

Im obigen Beispiel ist  $\kappa \approx 0,7395$ , wir sind also weit von einem perfekten linearen Zusammenhang entfernt, noch weiter allerdings von zwei Größen, zwischen denen keinerlei Zusammenhang besteht.

Da das Skalarprodukt zweier Vektoren gleich dem Produkt der Längen mal dem Kosinus des eingeschlossenen Winkels ist, können wir  $\kappa$  auch geometrisch interpretieren als den Kosinus des Winkels zwischen den Vektoren  $u$  und  $v$ . Im Falle  $\kappa = \pm 1$  ist dieser Winkel gleich  $0^\circ$  oder  $180^\circ$ , die beiden Vektoren liegen also auf einer Geraden; ist  $\kappa = 0$ , bilden sie einen rechten Winkel. Im obigen Fall ist der Winkel ungefähr  $42,31^\circ$ , also weit entfernt von beiden Extremen.

**e) Höhere Ableitungen impliziter Funktionen**

Die Kriterien aus Abschnitt c) helfen uns bei der Klassifikation von Extrema ohne Nebenbedingungen; den Fall von Extrema unter Nebenbedingungen haben wir zumindest theoretisch via den Satz über implizite Funktionen darauf zurückgeführt.

In den meisten praktischen Anwendungen wird man bei der Lösung einer Optimierungsaufgabe unter Nebenbedingungen *ad hoc* entscheiden können, wo die Maxima und/oder Minima liegen; teilweise helfen auch die im nächsten Paragraphen behandelten Existenzsätze.

Im Prinzip können wir allerdings auch für eine implizit gegebene Funktion höhere Ableitungen berechnen – vorausgesetzt, natürlich, daß sie existieren.

Beim Satz über implizite Funktionen hatten wir eine differenzierbare Funktion  $F(x, y)$  zweier Veränderlicher und einen Punkt  $(x_0, y_0) \in \mathbb{R}^2$  mit  $F(x_0, y_0) = 0$ , aber  $F_y(x_0, y_0) \neq 0$ . Unter diesen Voraussetzungen konnten wir zeigen, daß es um  $x_0$  eine differenzierbare Funktion  $f(x)$  gibt, so daß  $F(x, f(x))$  identisch verschwindet. Im Beweis berechneten wir auch gleich die Ableitung  $f'(x)$ ; sobald wir allerdings wissen, daß diese existiert, können wir sie auch einfacher bestimmen: Die Funktion

$x \mapsto F(x, f(x))$  ist gleich der Nullfunktion, und damit verschwindet natürlich ihre Ableitung. Andererseits ist diese Ableitung nach der Kettenregel gleich

$$F_x(x, f(x)) + F_y(x, f(x)) \cdot f'(x),$$

also folgt

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))}.$$

Genauso kann nun auch die zweite Ableitung von  $f$  berechnet werden –falls sie existiert. Wenn  $F$  und  $f$  zweimal stetig differenzierbar sind, können wir  $F(x, f(x))$  zweimal ableiten, was natürlich immer noch Null ist. Nach der Kettenregel ist aber die zweite Ableitung von  $F(x, f(x))$  (der Übersichtlichkeit halber jeweils ohne das Argument  $(x, f(x))$  geschrieben) gleich

$$\begin{aligned} & \frac{\partial}{\partial x}(F_x + F_y \cdot f'(x)) + \frac{\partial}{\partial y}(F_x + F_y \cdot f'(x)) \cdot f'(x) \\ &= F_{xx} + F_{yx} \cdot f'(x) + F_y \cdot f''(x) + (F_{xy} + F_{yy} \cdot f'(x)) \cdot f'(x), \end{aligned}$$

d.h.

$$\begin{aligned} f''(x) &= -\frac{1}{F_y}(F_{xx} + 2F_{xy} \cdot f'(x) + F_{yy} \cdot f'(x)^2) \\ &= -\frac{1}{F_y^3}(F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2). \end{aligned}$$

Für Extremwertbetrachtungen bei implizit definierten Funktionen braucht man, die zweite Ableitung vor allem in den Punkten, in denen die erste verschwindet; dort vereinfacht sich die Formel zu

$$f''(x) = -\frac{F_{xx}(x, f(x))}{F_y(x, f(x))} \quad \text{falls } f'(x) = 0.$$

Entsprechend lassen sich auch zunehmend komplizierter werdende Formeln für höhere Ableitungen herleiten, und wenn  $F$  von mehr als zwei Variablen abhängt auch solche für partielle Ableitungen.

**§4: Die Topologie des  $\mathbb{R}^n$**

Für Funktionen einer Veränderlichen haben wir Sätze wie den Zwischenwertsatz, der uns im wesentlichen sagt, daß jede stetige Funktion

ein abgeschlossenes Intervall wieder auf ein abgeschlossenes Intervall abbildet, und auch den Satz vom Maximum, wonach eine stetige Funktion auf einem abgeschlossenen Intervall sowohl ihr Infimum als auch ihr Supremum annimmt. Funktionen mehrerer Veränderlicher haben kompliziertere Definitionsbereiche als Intervalle; wir brauchen daher etwas mehr Aufwand, um auch hier analoge Sätze zu formulieren.

### a) Kompakte Mengen

Kompakte Teilmengen des  $\mathbb{R}^n$  sollen im wesentlichen die Rolle spielen, die abgeschlossene Intervalle in  $\mathbb{R}$  spielen. Wegen der teils sehr komplizierten Gestalt der Definitionsbereiche unserer Funktionen kommen wir zu ihrer Definition allerdings nur über einen auf den ersten Blick eher seltsamen Umweg.

Im Laufe sowohl der *Analysis I* als auch der *Analysis II* war immer wieder die Rede davon, daß gewisse Aussagen in einer *kleinen* Umgebung eines Punktes  $x$  gelten; wie groß diese Umgebung ist, hat uns im Einzelnen nicht weiter interessiert; wir forderten nur, daß es irgendein  $\varepsilon > 0$  geben solle, so daß alle  $y$  mit  $\|x - y\| < \varepsilon$  dazu gehören. Wenn wir verschiedene Punkte  $x$  betrachten, werden wir dabei eventuell für jeden von diesen ein anderes  $\varepsilon$  haben.

Nehmen wir etwa an, wir haben in  $\mathbb{R}$  die Punkte  $x_n = \frac{1}{n}$  und dazu die Umgebungen  $U_n = (\frac{1}{2n}, \frac{3}{2n})$ , d.h. also alle Punkte, deren Abstand von  $x_n = \frac{1}{n}$  kleiner ist als  $\frac{1}{2n}$ . Innerhalb jeder der Mengen  $U_n$  soll irgendeine für uns interessante Aussage gelten; beispielsweise soll es möglich sein, eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  in jeder der Mengen  $U_n$  durch eine deutlich einfachere Funktion  $f_n: U_n \rightarrow \mathbb{R}$  anzunähern, so daß der Fehler eine gewisse Schranke nicht überschreitet.

Um die Funktion auf dem offenen Intervall  $(0, 1)$  näherungsweise zu berechnen, können wir uns für jeden Punkt  $x \in (0, 1)$  eine Menge  $U_n$  wählen, die  $x$  enthält: Für  $x \in (0, 1)$  ist  $1/x > 1$ , also liegt zwischen  $1/x$  und  $3/x$  mindestens eine gerade Zahl  $2n$ , und für diese ist

$$\frac{1}{x} < 2n < \frac{3}{x}, \quad \text{also} \quad \frac{1}{2n} < x < \frac{3}{2n}.$$

Wir können also immer mindestens ein  $n$  finden, für das  $x$  in  $U_n$  liegt, und  $f(x)$  dann durch  $f_n(x)$  annähern. Für praktische Zwecke, zum Bei-

spiel für ein Computerprogramm, ist das vor allem dann nützlich, wenn wir mit endlich vielen Mengen  $U_n$  und damit auch mit endlich vielen Näherungsfunktionen  $f_n$  auskommen können.

Das ist hier aber leider nicht möglich: In  $U_n$  liegen nur Zahlen, die größer sind als  $1/2n$ . Wenn wir eine endliche Auswahl  $U_{n_1}, \dots, U_{n_k}$  dieser Mengen betrachten mit  $n_1 < \dots < n_k$ , so enthält also keine dieser Mengen eine Zahl kleiner oder gleich  $1/2n_k$ .

Hätten wir allerdings zusätzlich zu den Mengen  $U_n$  noch irgendeine offene Menge  $U_0$ , die die Null enthält, so würden endlich viele Mengen ausreichen: Da die Null ein innerer Punkt von  $U_0$  sein müßte, gäbe es ein  $\delta > 0$ , so daß  $U_0$  das Intervall  $(-\delta, \delta)$  enthielte, und damit würde es ausreichen, nur das Intervall  $U_0$  sowie die Intervalle  $U_n$  mit  $n < 1/(2\delta)$  zu betrachten (oder sogar nur eine Auswahl davon).

Bei der praktischen Approximation von Funktionen dürfte dieses Beispiel zwar kaum eine Rolle spielen, aber es gibt eine Vielzahl von Situationen sowohl in der Analysis als auch der Geometrie und anderen Gebieten, in denen es von entscheidender Bedeutung ist, daß wir mit endlich vielen der vorgegebenen Mengen überdecken können. Deshalb ist die folgende Definition, so künstlich sich bei der ersten Lektüre auch erscheinen mag, in weiten Teilen der Mathematik von fundamentaler Bedeutung:

**Definition:** a) Ein System  $\mathcal{U} = \{U_i \mid i \in I\}$  von offenen Teilmengen  $U_i \in \mathbb{R}^n$ , wobei  $I$  eine beliebige Indexmenge bezeichnet, heißt *offene Überdeckung* der Teilmenge  $X \subseteq \mathbb{R}^n$ , wenn

$$X \subseteq \bigcup_{i \in I} U_i$$

in der Vereinigungsmenge aller  $U_i$  liegt.

b) Ist  $J \subseteq I$  eine Teilmenge von  $I$  und liegt  $X$  bereits in der Vereinigung aller  $U_i$  mit  $i \in J$ , bezeichnen wir  $\mathfrak{B} = \{U_i \mid i \in J\}$  als eine *Teilüberdeckung* von  $\mathcal{U}$ . Ist speziell  $J$  eine endliche Menge, so sprechen wir von einer *endlichen Teilüberdeckung*.

c) Eine Teilmenge  $X \subseteq \mathbb{R}^n$  heißt *kompakt*, wenn jede offene Überdeckung  $\mathcal{U}$  von  $X$  eine endliche Teilüberdeckung hat.

Beginnen wir zur Veranschaulichung mit einigen Beispielen von Überdeckungen;  $\|\cdot\|$  soll dabei stets die EUKLIDISCHE Norm bezeichnen:

$\mathcal{U}_1$  bestehe aus allen offenen Kreisscheiben

$$U_x = \{y \in \mathbb{R}^n \mid \|x - y\| < 1\}$$

vom Radius eins um Punkte  $x \in \mathbb{R}^n$ ; hier ist also die Menge  $I$  der gesamte  $\mathbb{R}^n$ . Offensichtlich ist  $\mathcal{U}_1$  eine offene Überdeckung sowohl von  $\mathbb{R}^n$  als auch von jeder Teilmenge  $X \subseteq \mathbb{R}^n$ . Zumindest als offene Überdeckung von ganz  $\mathbb{R}^n$  hat sie keine endliche Teilüberdeckung, denn sonst gäbe es ja endlich viele Punkte, so daß jeder beliebige Punkt in  $\mathbb{R}^n$  von mindestens einem dieser Punkte höchstens den Abstand eins hätte. Damit ist klar, daß  $\mathbb{R}^n$  nicht kompakt ist.

$\mathcal{U}_2$  bestehe aus denselben Kreisscheiben  $U_x$ , jetzt aber nur für Punkte  $x$  mit ganzzahligen Koordinaten, d.h.  $x \in J = \mathbb{Z}^n$ . Für  $n \leq 3$  ist dies ebenfalls eine offene Überdeckung von  $\mathbb{R}^n$ , denn für einen beliebigen Punkt  $y \in \mathbb{R}^n$  erhalten wir einen Punkt  $x \in \mathbb{Z}^n$  mit ganzzahligen Koordinaten, indem wir einfach jede Koordinate  $y_i$  von  $y$  zur nächstgelegenen ganzen Zahl runden. In jeder einzelnen Koordinate ist der Fehler höchstens gleich  $1/2$ , insgesamt also höchstens

$$\sqrt{\sum_{i=1}^n \left(\frac{1}{2}\right)^2} = \frac{\sqrt{n}}{2},$$

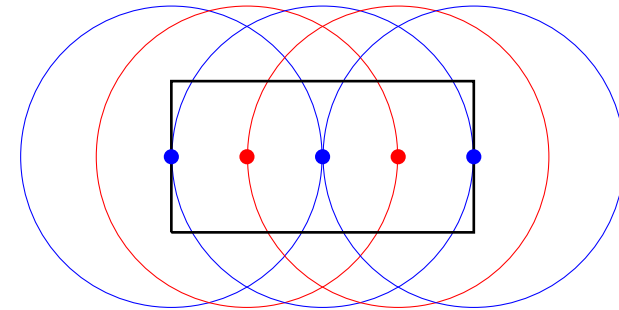
was für  $n \leq 3$  kleiner als eins ist. Für  $n = 4$  jedoch hat beispielsweise der Punkt  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$  von jedem Punkt mit ganzzahligen Koordinaten mindestens den Abstand eins, liegt also in keiner der Mengen  $U_x$ . Somit haben wir nur für  $n \leq 3$  eine Überdeckung von  $\mathbb{R}^n$ , die dann natürlich eine Teilüberdeckung von  $\mathcal{U}_1$  ist.

$\mathcal{U}_3$  bestehe aus allen Intervallen der Form  $(\frac{1}{2n}, \frac{3}{2n})$  für  $n \in \mathbb{N}$ ; hier ist also die Indexmenge  $I = \mathbb{N}$ . Wie wir oben gesehen haben, ist  $\mathcal{U}_3$  eine offene Überdeckung des offenen Intervalls  $(0, 1)$ , die keine endliche Teilüberdeckung hat; somit ist das offene Intervall  $(0, 1)$  nicht kompakt.

$\mathcal{U}_4$  schließlich bestehe aus den offenen Kreisscheiben vom Radius zwei um die Punkte  $(i, 1)$  mit  $i = 0, 1, 2, 3, 4$ . Dies ist eine offene Überdeckung des Rechtecks

$$X = [0, 4] \times [0, 2] = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 4 \text{ und } 0 \leq y \leq 2\}.$$

Teilüberdeckungen sind zum Beispiel die Überdeckung, aus den beiden (roten) Kreisscheiben um die Punkte  $(1, 1)$  und  $(3, 1)$ , aber auch die Überdeckung aus den drei (blauen) Kreisscheiben um  $(0, 1)$ ,  $(2, 1)$  und  $(4, 1)$ .



Die obigen Beispiele haben uns gezeigt, daß  $\mathbb{R}^n$  und das offene Intervall  $(0, 1)$  nicht kompakt sind; für Beispiele kompakter Mengen reicht es natürlich nicht aus, nur spezielle Überdeckungen zu betrachten; hier müssen wir zeigen, daß jede irgendwie gegebene Überdeckung eine endliche Teilüberdeckung hat.

Wie eingangs erwähnt, sollen kompakte Mengen im  $\mathbb{R}^n$  ähnliche Eigenschaften haben wie abgeschlossene Intervalle in  $\mathbb{R}$ ; wenn dies mit obiger Definition der Fall ist, sollten insbesondere alle abgeschlossenen Intervalle kompakt sind. Wir beweisen gleich etwas mehr:

**Lemma:** Jeder Quader

$$Q = [a_1, b_1] \times \dots \times [a_n, b_n]$$

$$= \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \text{ für alle } i = 1, \dots, n\}$$

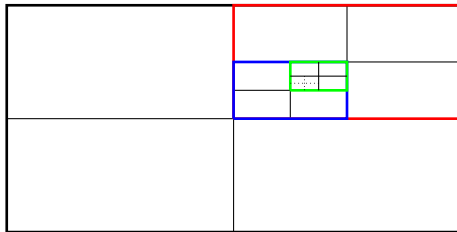
in  $\mathbb{R}^n$  ist kompakt.

*Beweis:* Wir nehmen an, es gebe eine Überdeckung  $\mathcal{U} = \{U_i \mid i \in I\}$  von  $Q$ , die keine endliche Teilüberdeckung habe, und wollen daraus einen Widerspruch herleiten, indem wir eine Art mehrdimensionale Intervallschachtelung konstruieren. Startpunkt ist der Quader  $Q$ , den wir zu diesem Zweck als  $Q^{(1)} = [a_1^{(1)}, b_1^{(1)}] \times \dots \times [a_n^{(1)}, b_n^{(1)}]$  schreiben. Nach Voraussetzung kann er nicht durch endlich viele Mengen  $U_i$  aus  $\mathcal{U}$  überdeckt werden.



Um aus einem Quader  $Q^{(k)}$  dessen Nachfolger  $Q^{(k+1)}$  zu konstruieren, teilen wir jedes der Intervalle  $[a_j^{(k)}, b_j^{(k)}]$  bei seinem Mittelpunkt Mittelpunkt  $c_j^{(k)} = \frac{1}{2}(a_j^{(k)} + b_j^{(k)})$  in die beiden Halbinservalle  $[a_j^{(k)}, c_j^{(k)}]$  und  $[c_j^{(k)}, b_j^{(k)}]$ . Damit können wir den Quader  $Q^{(k)}$  in  $2^n$  Teilquader zerlegen, die sich jeweils als Produkte von  $n$  solchen Teilintervallen darstellen lassen.

Wenn  $Q^{(k)}$  nicht durch endlich viele der Mengen aus  $\mathcal{U}$  überdeckt werden kann, muß für mindestens einen dieser  $2^n$  Teilquader dasselbe gelten: Da die Vereinigung aller Teilquader gleich  $Q^{(k)}$  ist, hätten wir sonst auch eine endliche Teilüberdeckung von  $Q^{(k)}$ . Einen solchen Quader, für den es keine endliche Teilüberdeckung gibt, bezeichnen wir als  $Q^{(k+1)}$  und zerteilen ihn weiter.



$$Q = Q^{(1)} \supset Q^{(2)} \supset Q^{(3)} \supset Q^{(4)} \supset \dots$$

Damit haben wir eine Folge von Quadern  $Q^{(1)} \supset Q^{(2)} \supset Q^{(3)} \supset \dots$ , von denen keiner durch endlich viele der  $U_i$  überdeckt werden kann. In den  $n$  Koordinaten haben wir jeweils eine entsprechende Folge von Intervallen

$$[a_j^{(1)}, b_j^{(1)}] \supset [a_j^{(2)}, b_j^{(2)}] \supset [a_j^{(3)}, b_j^{(3)}] \supset \dots,$$

von denen jedes die halbe Länge hat wie sein Vorgänger. Damit geht die Länge  $b_j^{(k)} - a_j^{(k)}$  für  $k \rightarrow \infty$  gegen Null, die obige Folge von Intervallen ist also eine Intervallschachtelung und definiert somit eine reelle Zahl  $x_j$ .

Wir betrachten den Punkt  $(x_1, \dots, x_n) \in \mathbb{R}^n$ . Nach Konstruktion liegt er in jedem der Quader  $Q^{(k)}$ , insbesondere also in  $Q$  selbst. Da  $Q$  in der Vereinigung der Mengen  $U_i$  liegt, muß es daher ein  $i \in I$  geben, so daß  $x$  in  $U_i$  liegt.

Da  $U_i$  eine offene Menge ist, gibt es ein  $\delta > 0$ , so daß mit  $x$  auch jedes  $y \in \mathbb{R}^n$  mit  $\|y - x\| < \delta$  in  $U_i$  liegt. Damit müssen aber auch ab einem gewissen  $k_0$  alle  $Q^{(k)}$  mit  $k \geq k_0$  in  $U_i$  liegen: Der Ausgangsquader  $Q^{(1)}$  hat den Durchmesser

$$d = \sqrt{(b_1 - a_1)^2 + \dots + (b_n - a_n)^2},$$

und da alle Kanten von  $Q^{(k+1)}$  genau halb so lang sind wie die entsprechenden Kanten von  $Q^{(k)}$ , ist auch der Durchmesser nur halb so lang, d.h. der Durchmesser von  $Q^{(k)}$  ist  $d/2^{k-1}$ , und das ist ab einem gewissen  $k_0$  kleiner als  $\delta$ . Somit hat für  $k \geq k_0$  jeder von  $Q^{(k)}$  höchstens den Abstand  $d/2^{k-1}$  von  $x$ , also einen kleineren Abstand als  $\delta$ . Dies wiederum bedeutet, daß  $Q^{(k)}$  in  $U_i$  liegt, im Widerspruch zur Annahme, daß  $Q^{(k)}$  nicht durch endlich viele der Mengen aus  $\mathcal{U}$  überdeckt werden kann. Damit ist das Lemma bewiesen. ■

Wenn Quader die einzigen kompakten Teilmengen von  $\mathbb{R}^n$  wären, hätte sich der Aufwand für die Definition eines so komplizierten Begriffs nicht gelohnt. Das folgende Lemma liefert uns zusammen mit dem gerade bewiesenen eine Fülle von weiteren Beispielen, die insbesondere auch krummlinig begrenzt sein können:

**Lemma:** Ist die abgeschlossene Menge  $Z \subset \mathbb{R}^n$  Teilmenge einer kompakten Menge  $K \subset \mathbb{R}^n$ , ist auch  $Z$  kompakt.

*Beweis:*  $\mathcal{U} = \{U_i \mid i \in I\}$  sei eine offene Überdeckung der Menge  $Z$ . Wegen der Abgeschlossenheit von  $Z$  ist  $V = \mathbb{R}^n \setminus Z$  offen; nehmen wir  $V$  noch mit dazu, erhalten wir eine offene Überdeckung  $\mathfrak{V} = \mathcal{U} \cup \{V\}$  von  $K$ , denn jeder Punkt aus  $K \setminus Z$  liegt erst recht in  $V = \mathbb{R}^n \setminus Z$ , und jeder Punkt aus  $Z$  liegt in mindestens einer der Mengen  $U_i$ .

Da  $K$  kompakt ist, hat  $\mathfrak{V}$  eine endliche Teilüberdeckung (von  $K$ ). Wenn diese Teilüberdeckung ohne die Menge  $V$  auskommt, ist sie gleichzeitig eine endliche Teilüberdeckung von  $\mathcal{U}$  für die Menge  $Z$ . Andernfalls betrachten wir die Teilüberdeckung *ohne* die Menge  $V$ . Das ist dann eine endliche Teilmenge von  $\mathcal{U}$ , und es ist auch eine offene Überdeckung von  $Z$ , denn jeder Punkt von  $Z \subseteq K$  muß in einer der offenen Mengen aus der Teilüberdeckung liegen, und er liegt sicher nicht in  $V = \mathbb{R}^n \setminus Z$ . Somit hat  $\mathcal{U}$  eine endliche Teilüberdeckung von  $Z$ . ■

Damit kennen wir im wesentlichen bereits alle kompakten Teilmengen von  $\mathbb{R}^n$ :

**Definition:** Eine Teilmenge  $X \subseteq \mathbb{R}^n$  heißt *beschränkt*, wenn es ein  $M \in \mathbb{R}$  gibt, so daß  $\|x\| \leq M$  ist für alle  $x \in X$ .

Da die beiden Normen äquivalent sind, ist es hierbei unwesentlich, ob wir mit der EUKLIDischen Norm oder der Maximumsnorm arbeiten.

**Satz von Heine-Borel:** Eine Teilmenge  $X \subseteq \mathbb{R}^n$  ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

*Beweis:* Sei zunächst  $X \subseteq \mathbb{R}^n$  kompakt. Wir betrachten die offene Überdeckung von  $X$  aus den Mengen

$$U_x = \{y \in \mathbb{R}^n \mid \|y - x\| < 1\}.$$

Wegen der Kompaktheit von  $X$  hat sie eine endliche Teilüberdeckung; diese bestehe aus den Mengen  $U_{x_1}, \dots, U_{x_r}$ . Nach der Dreiecksungleichung ist

$$\|y\| \leq \|x_i\| + \|y - x_i\| \leq \|x_i\| + 1 \quad \text{für alle } y \in U_{x_i};$$

bezeichnet  $R$  die größte unter den endlich vielen Normen  $\|x_i\|$ , ist also  $\|y\| \leq R + 1$  für alle  $y \in X$ . Somit ist  $X$  beschränkt.

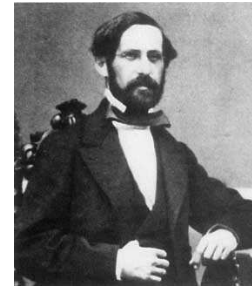
Um zu sehen, daß  $X$  auch abgeschlossen ist, zeigen wir, daß das Komplement  $\mathbb{R}^n \setminus X$  offen ist. Dazu sei  $z \in \mathbb{R}^n \setminus X$  ein beliebiger Punkt aus diesem Komplement; wir müssen zeigen, daß es ein  $\varepsilon > 0$  gibt, so daß  $\{y \in \mathbb{R}^n \mid \|y - z\| < \varepsilon\}$  ganz in  $\mathbb{R}^n \setminus X$  liegt.

Die offenen Mengen

$$U_n = \{x \in \mathbb{R}^n \mid \|z - y\| > 1/n\}$$

überdecken  $\mathbb{R}^n \setminus \{y\}$ , denn für jeden Punkt  $z \neq y$  gibt es irgendein  $n \in \mathbb{N}$ , so daß die Norm von  $y - z$  größer ist als  $1/n$ . Damit überdecken Sie insbesondere auch  $X$ , und wegen der Kompaktheit von  $X$  reicht dazu bereits eine endliche Teilüberdeckung bestehend aus gewissen Mengen  $U_{n_1}, \dots, U_{n_r}$ . Ist  $n_r$  der größte unter den  $r$  Indizes, ist die Vereinigung dieser Mengen gleich  $U_{n_r}$ , d.h.  $X \subseteq U_{n_r}$ . Damit hat jeder Punkt aus  $X$  von  $y$  einen größeren Abstand als  $1/n_r$ , wir können also  $\varepsilon = 1/n_r$  setzen.

Umgekehrt sei die Teilmenge  $X \subseteq \mathbb{R}^n$  abgeschlossen und beschränkt. Wegen der Beschränktheit gibt es einen Quader  $Q$ , der  $X$  enthält. Dieser Quader ist nach dem ersten der obigen Lemmata kompakt, und nach dem zweiten gilt dasselbe für jede darin enthaltene abgeschlossene Teilmenge. Somit ist  $X$  kompakt. ■



HEINRICH EDUARD HEINE (1821–1881) wurde in Berlin als achtes der neun Kinder eines Bankiers geboren. Ab 1838 studierte er zunächst an der Universität Berlin, wechselte aber schon nach zum zweiten Semester nach Göttingen, wo er unter anderem Vorlesungen von GAUSS über Zahlentheorie hörte. Drei Semester später kehrte er nach Berlin zurück, wo er 1842 promovierte. Nach einem kurzen Aufenthalt an der Universität Königsberg habilitierte er sich 1844 an der Universität Bonn, wo er zunächst als Privatdozent, dann als außerplanmäßiger Professor lehrte.

1856 bekam er einen Lehrstuhl an der Universität Halle, den er bis zu seinem Tod innehatte. Seine Arbeiten befassen sich unter anderem mit partiellen Differentialgleichungen, Kettenbrüchen und elliptischen Funktionen; auch der Begriff der gleichmäßigen Stetigkeit geht auf ihn zurück.



FÉLIX EDOUARD JUSTIN EMILE BOREL (1871–1956), kurz EMILE BOREL, wurde im französischen Saint Affrique nahe der Pyrenäen als Sohn eines protestantischen Pfarrers geboren. Mit elf Jahren verließ er Saint Affrique, um zunächst in Montauban, dann in Paris weiterführende Schulen zu besuchen. Er legte sowohl die Aufnahmeprüfung zur Ecole Polytechnique als auch die zur Ecole Normale als Bester seines Jahrgangs ab und entschied sich dann zum Studium an der Ecole Normale, wo er 1893 promovierte. Danach arbeitete er als Maître de Conférence zunächst an der Universität Lille, dann an der Ecole Normale. Nach einigen weiteren

Positionen unter anderem am Collège de France erhielt er 1909 einen Lehrstuhl an der Sorbonne. Trotz vielfältiger politischer Aktivitäten unter anderen als Marineminister von 1925 bis 1940 behielt er diesen Lehrstuhl bis zu seiner Verhaftung 1941 wegen seines Kampfs in der Resistance. Nach dem Krieg war er unter anderem Präsident des Wissenschaftsrats der UNESCO. Er publizierte rund zwanzig Lehrbücher und zahlreiche Arbeiten aus so unterschiedlichen Gebieten wie der reellen und komplexen Analysis, der Differentialgleichungen, der Arithmetik, der Numerik, Maßtheorie, Wahrscheinlichkeitstheorie und Spieltheorie.

Damit haben wir einen vollständigen Überblick über die kompakten Teil-