oder, anders ausgedrückt,

$$\operatorname{grad} f(\boldsymbol{a}) - \frac{f_{x_n}(\boldsymbol{a})}{g_{x_n}(\boldsymbol{a})} \operatorname{grad} g(\boldsymbol{a}) = \vec{0}\,.$$

Also sind die beiden Gradienten in der Tat linear abhängig. ■

Falls der Gradient von g im Punkt $\boldsymbol{a}$ nicht verschwindet, gibt es somit eine Zahl $\lambda \in \mathbb{R}$, so daß

$$\operatorname{grad} f(\boldsymbol{a}) + \lambda \operatorname{grad} g(\boldsymbol{a}) = \vec{0}$$

ist, nämlich

$$\lambda = -\frac{f_{x_n}(\boldsymbol{a})}{g_{x_n}(\boldsymbol{a})}\,.$$

Diese Zahl bezeichnet man als LAGRANGEschen Multiplikator; mit seiner inhaltlichen Interpretation werden wir uns in Kürze beschäftigen.

JOSEPH-LOUIS LAGRANGE (1736–1813) wurde als GIUSEPPE LODOVICO LAGRANGIA in Turin geboren und studierte dort zunächst Latein. Erst eine alte Arbeit von HALLEY über algebraische Methoden in der Optik weckte sein Interesse an der Mathematik, woraus ein ausgedehnter Briefwechsel mit EULER entstand. In einem Brief vom 12. August 1755 berichtete er diesem unter anderem über seine Methode zur Berechnung von Maxima und Minima; 1756 wurde er, auf EULERs Vorschlag, Mitglied der Berliner Akademie; zehn Jahre später zog er nach Berlin und wurde dort EULERs Nachfolger als mathematischer Direktor der Akademie. 1787 wechselte er an die Pariser Académie des Sciences, wo er bis zu seinem Tod blieb und unter anderem an der Einführung des metrischen Systems beteiligt war. Seine Arbeiten umspannen weite Teile der Analysis, Algebra und Geometrie.

Zur praktischen Bestimmung von Extremwerten unter Nebenbedingungen geht man wie folgt vor: Über die Punkte, in denen der Gradient von g verschwindet, macht obiger Satz keine verwertbare Aussage; diese Punkte müssen also vorab berechnet und untersucht werden.

Danach müssen die Punkte gefunden werden, in denen es ein $\lambda \in \mathbb{R}$ gibt, so daß

$$\begin{aligned} f_{x_1}(\boldsymbol{x}) + \lambda g_{x_1}(\boldsymbol{x}) &= 0 \\ &\vdots \\ f_{x_n}(\boldsymbol{x}) + \lambda g_{x_n}(\boldsymbol{x}) &= 0 \\ g(\boldsymbol{x}) &= 0 \end{aligned}$$

ist. Dies ist ein System von $n+1$ Gleichungen für die $n+1$ Unbekannten, allerdings ist dieses Gleichungssystem nur selten linear und damit oft nicht mit bekannten Methoden lösbar. Manchmal kann man das Gleichungssystem durch geeignete Umformungen und Fallunterscheidungen vollständig lösen, in anderen Fällen helfen nur die aus der Numerik bekannten Näherungsverfahren wie etwa die Methode von NEWTON-RALSTON.

Falls alle Gleichungen Polynomgleichungen sind (oder durch Einführung geeigneter zusätzlicher Variablen auf Polynomgleichungen zurückgeführt werden können), kann man im Falle einer endlichen Lösungsmenge diese auch exakt bestimmen: Genau wie der GAUSS-Algorithmus zur Lösung eines linearen Gleichungssystems dieses auf eine Treppengestalt bringt, aus der man die Lösungen einfach ermitteln kann, gibt es in der Computeralgebra einen Algorithmus, der dasselbe für beliebige Systeme von Polynomgleichungen versucht; die Gleichungen, die dieser Algorithmus liefert, bezeichnet man als GRÖBNER-Basis oder Standardbasis. Zum Verständnis dieses Algorithmus, den man als eine Art Synthese aus EUKLIDischen Algorithmus und GAUSS-Algorithmus ansehen kann, sind Kenntnisse der kommutativen Algebra erforderlich, für die die Zeit in dieser Vorlesung nicht ausreicht; bei einigen Implementierungen werden zusätzlich auch noch Algorithmen aus der Informatik eingesetzt, die typischerweise nicht in Grundvorlesungen behandelt werden. Deshalb sei hier nur darauf hingewiesen, daß die gängigen universellen Computeralgebrasysteme wie Maple, Mathematica, MuePad allesamt entsprechende Routinen enthalten, mit denen man auch dann experimentieren kann, wenn man die dahinterstehende Theorie nicht versteht.

Als Beispiel, wie gelegentlich auch ein nichtlineares Gleichungssystem elementar gelöst werden kann, betrachten wir eine Anwendung aus den Wirtschaftswissenschaften: Die Gesamtproduktion eines Unternehmens oder eines Staates in Abhängigkeit von n eingesetzten Ressourcen $x_1, \ldots, x_n$ wird oft modelliert durch eine sogenannte COBB-DOUGLAS-Funktion der Form

$$P(x_1, \ldots, x_n) = \alpha x_1^{e_1} \ldots x_n^{e_n}\,,$$

benannt nach den beiden Wissenschaftlern, die dieses Modell 1928 für die amerikanische Gesamtproduktion in Abhängigkeit von Kapital und Arbeit in den Jahren 1899 bis 1922 entwickelten. (Sie fanden $P \approx 1{,}01 A^{3/4} K^{1/4}$ mit A = Anzahl der Beschäftigten und K = Kapitaleinsatz.)

Betrachten wir stattdessen die Produktion eines Wirtschaftsguts aus zwei Resourcen x, y gemäß der Funktion

$$f(x,y) = P(x,y) = x^{1/2} y^{1/4}\,.$$

Falls wir der Einfachheit halber annehmen, daß die Kosten pro Einheit für x und y gleich sind und die Gesamtkosten höchstens gleich zwölf sein dürfen, müssen wir f maximieren unter der Nebenbedingung

$$x + y \le 12\,.$$

Nun ist aber f eine monoton wachsende Funktion sowohl von x als auch von y, d.h. die maximale Produktion wird sicherlich erreicht in einem Punkt, für den $x + y = 12$ ist, denn für jeden anderen Punkt (x, y) mit $x + y < 12$ ist $f(x, y) < f(x, 12 - x)$. Daher können wir die Nebenbedingung in der gewohnten Form

$$g(x,y) = x + y - 12 = 0$$

schreiben. Diese Nebenbedingung sowie die zu maximierende Funktion sind in Abbildung 45 dargestellt.

Ableitung beider Funktionen zeigt, daß

$$\operatorname{grad} g = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \operatorname{grad} f = \begin{pmatrix} y^{1/4}/2x^{1/2} \\ x^{1/2}/4y^{3/4} \end{pmatrix}$$

ist; das zu lösende Gleichungssystem wird also zu

$$\begin{aligned} \frac{y^{1/4}}{2x^{1/2}} + \lambda &= 0 \\ \frac{x^{1/2}}{4y^{3/4}} + \lambda &= 0 \\ x + y - 12 &= 0 \end{aligned}\,.$$

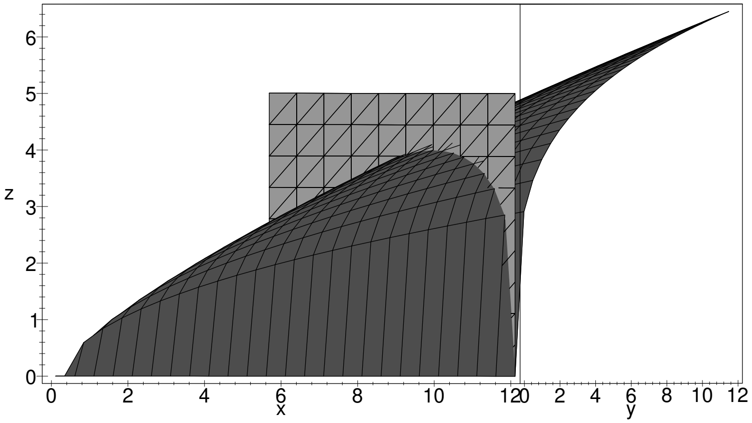


Abb. 45: Maximierung einer Produktionsfunktion bei festem Kapitaleinsatz

(Die Nenner brauchen uns nicht zu stören, denn da $f(0, y) = f(x, 0) = 0$ ist, kommen Lösungen mit $x = 0$ oder $y = 0$ für das Maximum ohnehin nicht in Frage; wir können sie also getrost ausschließen.)

Als Ansatz zu einer möglichen Lösung können wir ausnutzen, daß λ in den beiden ersten Gleichungen isoliert steht; wenn wir danach auflösen und gleichsetzen, erhalten wir die Gleichung

$$\frac{y^{1/4}}{2x^{1/2}} = \frac{x^{1/2}}{4y^{3/4}}\,.$$

Multiplikation mit dem Hauptnenner macht daraus

$$4y^{1/4}y^{3/4} = 2x^{1/2}x^{1/2} \quad \text{oder} \quad 2y = x\,.$$

Einsetzen in die dritte Gleichung ergibt $3y = 12$, also ist

$$y = 4 \quad \text{und} \quad x = 8\,,$$

und der Maximalwert von f ist

$$f(8,4) = 8^{1/2} \cdot 4^{1/4} = 2\sqrt{2} \cdot \sqrt{2} = 4\,.$$

Auch den LAGRANGEschen Multiplikator λ können wir noch ausrechnen:

$$\lambda = \frac{y^{1/4}}{2x^{1/2}} = \frac{4^{1/4}}{2 \cdot 8^{1/2}} = \frac{\sqrt{2}}{2 \cdot 2\sqrt{2}} = \frac{1}{4}\,.$$

Die Berechnung von λ war für die Bestimmung des Optimums eigentlich überflüssig; λ ist nur eine Hilfsgröße zur Berechnung des Extremums. Wir wollen uns als nächstes überlegen, daß wir λ auch inhaltlich interpretieren können: Dazu betrachten wir eine Nebenbedingung

$$g(x_1, \ldots, x_n) = c$$

mit *variabler* rechter Seite c und ein Extremum der Funktion

$$f(x_1, \ldots, x_n)\,.$$

Dieses Extremum wird natürlich von c abhängen; wir schreiben es in der Form

$$\big(x_1(c), \ldots, x_n(c)\big)$$

und nehmen an, daß die Funktionen $x_i(c)$ stetig differenzierbar seien. (Ein interessierter Leser kann sich anhand des Satzes über implizite Funktionen überlegen, welche Bedingungen f und g erfüllen müssen, damit dies garantiert ist.) Der Optimalwert von f in Abhängigkeit von c ist dann

$$F(c) \underset{\text{def}}{=} f\big(x_1(c), \ldots, x_n(c)\big)\,.$$

Nach der Kettenregel aus §2 ist

$$F'(c) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{dx_i(c)}{dc}\,.$$

Genauso können wir

$$G(c) \underset{\text{def}}{=} g\big(x_1(c), \ldots, x_n(c)\big)$$

betrachten und erhalten

$$G'(c) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i} \frac{dx_i(c)}{dc}\,.$$

Da $\big(x_1(c), \ldots, x_n(c)\big)$ ein Optimum ist, sind dort die Gradienten von f und $g - c$ proportional mit Proportionalitätsfaktor λ. Da wir bei der Gradientenbildung nur nach den x_i ableiten, von denen die rechte Seite c nicht abhängt, ist der Gradient von $g - c$ gleich dem von g selbst, d.h.

$$\frac{\partial f}{\partial x_i} = \lambda \frac{\partial g}{\partial x_i} \qquad \text{für alle } i\,.$$

Somit ist $F'(c) = \lambda G'(c)$. Da der Punkt $\big(x_1(c), \ldots, x_n(c)\big)$ die Nebenbedingung mit rechter Seite c erfüllt, ist aber $G(c) = c$ und damit $G'(c) \equiv 1$. Also ist $\lambda = F'(c)$ die Wachstumsrate für das Optimum bei Änderung der rechten Seite der Nebenbedingung.

Im obigen Beispiel steigt also die Maximalmenge $f(x, y)$, die mit Kapitaleinsatz 12 produziert werden kann, für kleines h ungefähr um $h/4$, wenn wir den Kapitaleinsatz auf $12 + h$ erhöhen. Die Erhöhung des Kapitaleinsatzes lohnt sich, wenn für das fertige Produkt ein Preis pro Einheit erzielt werden kann, der größer ist als vier.

Als letztes wollen wir uns noch überlegen, was passiert, wenn wir nicht nur eine, sondern mehrere Nebenbedingungen erfüllen müssen. Es geht also wieder darum, eine Funktion $f\big(x_1, \ldots, x_n)$ zu optimieren, jetzt aber unter den Nebenbedingungen

$$g_1(x_1, \ldots, x_n) \ge 0, \quad \ldots \quad g_r(x_1, \ldots, x_n) \ge 0\,.$$

(Es genügt, Bedingungen mit $\ge$ zu betrachten, denn durch Multiplikation mit minus Eins kann man jede Ungleichung mit $\le$ in eine mit $\ge$ überführen. Auch Gleichungen $g_i = 0$ kann man zumindest formal durch die beiden Ungleichungen $g_i \ge 0$ und $-g_i \ge 0$ ausdrücken.)

Die wichtigsten Beispiele solcher Optimierungsaufgaben sind die Fälle mit linearen Funktionen f und g_i; hier redet man von *linearen Programmen.* (Das Wort *Programme* in diesem Zusammenhang hat natürlich nichts mit Computerprogrammen zu tun.) Das wichtigste Verfahren zur Lösung solcher Aufgaben, der Simplex-Algorithmus, wird in der Vorlesung *Numerik I* behandelt, so daß wir uns hier auf die *nichtlineare Programmierung* beschränken können.

Man überlegt sich leicht, daß im linearen Fall die Nebenbedingungen ein (endliches oder unendliches) Polyeder im $\mathbb{R}^n$ definieren und eine lineare Funktion, so sie ein endliches Maximum oder Minumum hat, dieses auf dem Rand dieses Polyeders annimmt, und dort sogar in einer Ecke, Man muß daher „nur“ die Ecken dieses Polyeders untersuchen – deren Anzahl allerdings exponentiell mit der Anzahl der Variablen wächst. Trotzdem führt der Simplex-Algorithmus selbst im Fall von

Zehntausenden von Variablen im allgemeinen recht schnell ans Ziel, so daß das theoretische Problem der exponentiellen Komplexität im schlimmsten Fall für praktische Anwendungen keine Bedeutung hat.

Bei nichtlinearen Funktionen ist die Situation komplizierter, denn nun kann es auch im Innern Extrema geben: Die Funktion

$$f(x,y) = e^{-x^2-y^2} \quad \text{mit der Nebenbedingung} \quad x^2+y^2 \le 1$$

etwa nimmt ihr Maximum im Punkt $(0,0)$ an; auf dem Rand des Einheitskreises liegen nur die Minima. Im allgemeinen Fall eines nichtlinearen Programms kann ein Optimum also entweder ganz im Innern liegen oder aber eine beliebige Teilmenge der Nebenbedingungen exakt erfüllen.

Falls wir es mit inneren Punkte zu tun haben, sind diese lokale Maxima oder Minima ohne Nebenbedingungen, und wir haben uns bereits am Ende von §1 überlegt, wie man diese bestimmt: In jedem solchen Punkt verschwindet der Gradient der zu optimierenden Funktion.

Im Falle einer einzigen *Gleichung* als Nebenbedingung ist der Gradient von f linear abhängig vom Gradienten der Nebenbedingung; da der Nullvektor von jedem anderen Vektor linear abhängig ist, schließt dies auch den Fall der Optima bei inneren Punkten mit ein. Die naheliegende Verallgemeinerung auf den Fall mehrerer Nebenbedingungen ist daher der

Satz: Die Funktion $f\colon D \to \mathbb{R}$ auf $D \subseteq \mathbb{R}^n$ habe im Punkt $\boldsymbol{a} \in D$ ein Extremum unter den Nebenbedingungen

$$g_1(\boldsymbol{a}) \ge 0, \quad g_2(\boldsymbol{a}) \ge 0, \quad \dots, \quad g_r(\boldsymbol{a})\,.$$

Dann sind die $r+1$ Vektoren

$$\operatorname{grad} f(\boldsymbol{a}), \quad \operatorname{grad} g_1(\boldsymbol{a}), \quad \operatorname{grad} g_2(\boldsymbol{a}), \quad \dots, \quad \operatorname{grad} g_r(\boldsymbol{a})$$

linear abhängig.

Der *Beweis* erfordert keine wesentlich neuen Ideen gegenüber dem Fall einer einzigen Nebenbedingung und sei daher nur kurz skizziert: Falls die Gradienten der g_i im Punkt $\boldsymbol{a}$ bereits untereinander linear abhängig sind, gibt es nichts mehr zu beweisen; nehmen wir also an, sie seien linear unabhängig. Dann gibt es (mindestens) r verschiedene Variablen x_{j_1} bis x_{j_r}, so daß

$$\frac{\partial g_i}{\partial x_{j_i}}(\boldsymbol{a}) \neq 0$$

ist. Also kann nach dem Satz über implizite Funktionen jede Nebenbedingung zur Elimination einer anderen Variablen benutzt werden, und im wesentlichen dieselbe Rechnung wie im Fall einer Nebenbedingung zeigt die Behauptung.

■

Die lineare Abhängigkeit der Vektoren

$$\operatorname{grad} f(\boldsymbol{a}), \quad \operatorname{grad} g_1(\boldsymbol{a}), \quad \operatorname{grad} g_2(\boldsymbol{a}), \quad \dots, \quad \operatorname{grad} g_r(\boldsymbol{a})$$

bezeichnet man als KUHN-TUCKER-Bedingung; sie ist eine offensichtliche Verallgemeinerung der Bedingung von LAGRANGE, ist allerdings deutlich jünger: Sie erschien 1951 in einer gemeinsamen Arbeit von H.W. KUHN und A.W. TUCKER, vier Jahre, nachdem G. DANTZIG den Simplex-Algorithmus entwickelt hatte, und fast zweihundert Jahre, nachdem LAGRANGE seine Multiplikatoren zur Bestimmung von Extrema unter einer Nebenbedingung eingeführt hatte.

§4: Vektorwertige Funktionen mehrer Veränderlicher

Eine Funktion $f\colon D \to \mathbb{R}^m$ auf einer Teilmenge $D \subseteq \mathbb{R}^n$ läßt sich für $n, m \ge 2$ nicht mehr durch einen Graphen veranschaulichen, denn der Graph

$$\Gamma_f = \bigl\{(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^n \times \mathbb{R}^n \mid |\boldsymbol{y} = f(\boldsymbol{x})\bigr\}$$

liegt dann im mindestens vierdimensionalen $\mathbb{R}^{n+m}$.

Im für uns wichtigsten Fall $n = m = 2$ oder 3 gibt es aber eine andere Möglichkeit der Veranschaulichung: Für $\boldsymbol{x} \in D$ läßt sich $f(\boldsymbol{x})$ als Vektor auffassen, und f kann veranschaulicht werden, indem man für hinreichend viele Punkte $\boldsymbol{x} \in D$ diesen Vektor (oder ein geeignetes konstantes Vielfaches davon) im Punkt $\boldsymbol{x}$ anträgt; siehe dazu die Abbildungen 46 und 47.

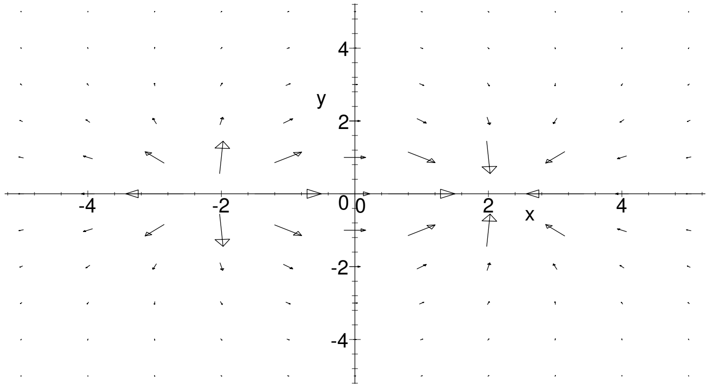


Abb. 46: Das elektrische Feld zweier entgegengesetzt gleicher Punktladungen

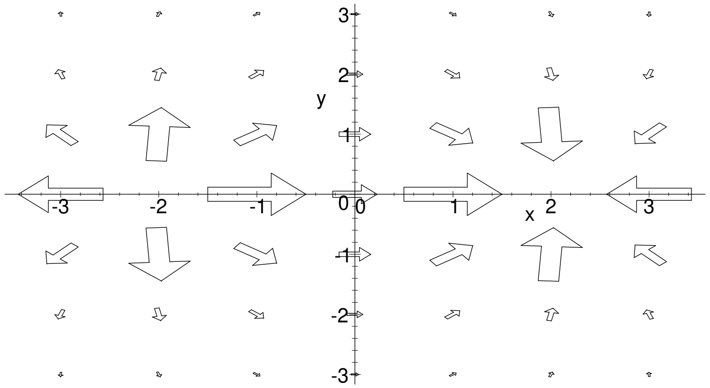


Abb. 47: Dasselbe Feld in einem kleineren Bereich mit „flächigen" Vektoren

Wie Abbildung 46 zeigt, kann es dabei bei betragsmäßig stark variierenden Funktionswerten zu Problemen kommen: Wie man die Konstante auch wählt, mit der alle Vektoren multipliziert werden, hat man immer entweder einige extrem lange Vektoren oder aber Vektoren, die so kurz sind, daß man nichts mehr erkennen kann: Ab einer gewissen Entfernung vom Zentrum sind die Richtungen des Felds praktisch nicht mehr

wahrnehmbar. Eine leichte Verbesserung ergibt sich, wenn man wie in Abbildung 47 die Vektoren „flächig" zeichnet, wobei die Länge des Vektors proportional zur *Fläche* des eingezeichneten Pfeils ist, aber auch das hilft nicht sehr viel.

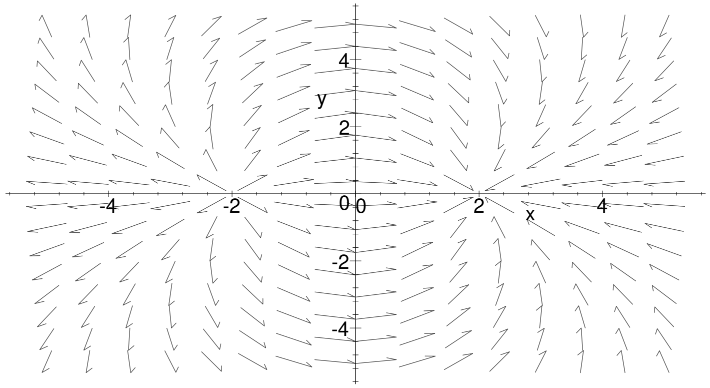


Abb. 48: Das Richtungfeld zum obigen Feld

Da gerade die Richtung sehr wesentlich für die Konstruktion von Feldlinien ist, begnügt man sich oft damit, nur die Richtungen darzustellen, ansonsten aber sämtlichen Vektoren dieselbe Länge zu geben; siehe dazu etwa Abbildung 48. Dadurch geht zwar ein Teil der Information verloren, aber das entsprechende Bild kann trotzdem oft mehr aussagen, als eines, bei dem fast alle Pfeile visuell zu Punkten degeneriert sind. Auch bei der Darstellung als Richtungsfeld müssen allerdings natürlich alle Punkte $\boldsymbol{x}$, für die $f(\boldsymbol{x}) = (0, \ldots, 0)$ ist, durch Punkte oder überhaupt nicht dargestellt werden, denn dort gibt es keine Richtung.

a) Der Begriff des Vektorfelds

Nicht nur zur Veranschaulichung kann es nützlich sein, die Werte einer Funktion als Vektoren zu betrachten: Die für uns wichtigsten Beispiele von Funktionen mehrerer Veränderlicher, deren Wert nicht in $\mathbb{R}$ sondern in einem $\mathbb{R}^n$ liegt, sind die aus der Physik bekannten Felder wie etwa ein

Kraftfeld, ein elektrisches Feld oder ein Magnetfeld. Da solche Felder typischerweise als Vektoren geschrieben werden, schreiben wir deshalb in diesem Paragraphen die *Funktionswerte* vektoriell, wohingegen die Argumente weiterhin Punkte sind, die als Tupel geschrieben werden. Wir betrachten somit Funktionen

$$\vec{V}\colon D \to \mathbb{R}^m; \quad \boldsymbol{x} \mapsto \vec{V}(\boldsymbol{x}) = \begin{pmatrix} V_1(\boldsymbol{x}) \\ \vdots \\ V_n(\boldsymbol{x}) \end{pmatrix},$$

wobei die Komponentenfunktionen $V_i\colon D \to \mathbb{R}$ genau die im vorigen Paragraphen betrachteten skalarwertigen Funktionen mehrerer Veränderlicher sind.

Der mathematische Begriff, der physikalische Felder beschreibt, ist der des Vektorfelds:

Definition: Ein Vektorfeld auf der Teilmenge $D \subseteq \mathbb{R}^n$ ist eine Funktion $\vec{V}\colon D \to \mathbb{R}^n$.

(Da die typischen elektromagnetischen Felder wie $\vec{E}, \vec{D}, \vec{B}, \vec{H}$ oder auch das Kraftfeld $\vec{F}$ üblicherweise durch Großbuchstaben bezeichnet werden, verwenden wir auch für Vektorfelder meist Großbuchstaben wie $\vec{V}$ oder $\vec{W}$.)

b) Die Jacobi-Matrix

Wir werden uns in diesem Paragraphen zwar hauptsächlich mit Vektorfeldern im $\mathbb{R}^n$ beschäftigen, und uns selbst da teilweise auf den Fall $n = 3$ beschränken; zumindest die grundlegenden Definitionen sollen aber für beliebige Funktionen auf einer Teilmenge D von $\mathbb{R}^n$ mit Werten in irgendeinem $\mathbb{R}^m$ betrachtet werden.

Fast von selbst versteht sich die

Definition: Eine Funktion $\vec{f}\colon D \to \mathbb{R}^m$ auf $D \subseteq \mathbb{R}^n$ heißt stetig *bzw.* differenzierbar, wenn alle Komponentenfunktionen f_i stetig *bzw.* differenzierbar sind. Entsprechend definieren wir auch die Menge $\mathcal{C}^k(D, \mathbb{R}^m)$ aller k-fach stetig differenzierbarer Funktionen von D nach $\mathbb{R}^m$ als Menge aller Funktionen, deren Komponenten $f_i \in \mathcal{C}^k(D, \mathbb{R})$ k-mal stetig differenzierbar sind.

Wie bei skalarwertigen Funktionen überlegt sich leicht, daß $\mathcal{C}^k(D, \mathbb{R}^m)$ ein $\mathbb{R}$-Vektorraum ist. Damit wissen wir also, was Differenzierbarkeit einer Funktion $\vec{f}\colon D \to \mathbb{R}^n$ bedeutet; was aber ist ihre Ableitung?

Bei skalarwertigen Funktionen einer oder mehrerer Veränderlicher hatten wir die Ableitung jeweils als Linearisierung der Funktion in der Umgebung des betrachteten Punkts definiert; genauso gehen wir auch hier vor: Für jede der Komponentenfunktion f_i ist

$$f_i(\boldsymbol{x} + \vec{h}) = f_i(\boldsymbol{x}) + \operatorname{grad} f_i(\boldsymbol{x}) \cdot \vec{h} + o\big(|\vec{h}|\big),$$

wobei der Punkt hier für das Skalarprodukt von Vektoren steht. Im Hinblick auf vektorwertige Funktionen ist es aber besser, anstelle des Skalarprodukt ein Matrixprodukt zu verwenden; dazu müssen wir den Gradienten als Zeilenvektor schreiben, d.h. wenn $\cdot$ das Matrixprodukt bezeichnet, ist

$$f_i(\boldsymbol{x} + \vec{h}) = f_i(\boldsymbol{x}) + {}^t\operatorname{grad} f_i(\boldsymbol{x}) \cdot \vec{h} + o\big(|\vec{h}|\big).$$

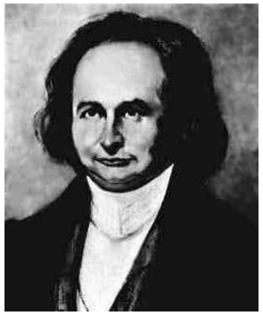
In dieser Notation können wir die als Zeilenvektoren aufgefaßten Gradienten zu einer $n \times n$-Matrix zusammenfassen, der JACOBI-Matrix

$$J_{\vec{f}}(\boldsymbol{x}) = \begin{pmatrix} {}^t\operatorname{grad} f_1(\boldsymbol{x}) \\ \vdots \\ {}^t\operatorname{grad} f_m(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Mit dieser Matrix ist dann

$$\vec{f}(\boldsymbol{x} + \vec{h}) = f(\boldsymbol{x}) + J_{\vec{f}}(\boldsymbol{x}) \cdot \vec{h} + o\big(|\vec{h}|\big),$$

wobei hier das Produkt als Matrixprodukt zu verstehen wird und das Symbol $o\big(|\vec{h}|\big)$ schlampigerweise auch für einen Vektor benutzt wird, dessen sämtliche Komponenten $o\big(|\vec{h}|\big)$ sind.



CARL GUSTAV JACOB JACOBI (1804–51) wurde in Potsdam als Sohn eines jüdischen Bankiers geboren und erhielt den Vornamen Jacques Simon. Im Alter von zwölf Jahren bestand er sein Abitur, mußte aber noch vier Jahre in Abschlußklasse des Gymnasiums bleiben, da die Berliner Universität nur Studenten mit mindestens 16 Jahren aufnahm. 1824 beendete er seine Studien mit dem Staatsexamen für Mathematik, Griechisch und Latein und wurde Lehrer. Außerdem promovierte er 1825 und begann mit seiner Habilitation. Etwa gleichzeitig konvertierte er zum Christentum, so daß er ab 1825 an der Universität Berlin und ab 1826 in Königsberg lehren konnte. 1832 wurde er dort Professor. Zehn Jahre später mußte er aus gesundheitlichen Gründen das rauhe Klima Königsbergs verlassen und war zunächst in Italien, danach für den Rest seines Lebens in Berlin. Er ist vor allem berühmt durch seine Arbeiten zur Zahlentheorie und über elliptische Integrale.

c) Die Divergenz eines Vektorfelds

Für ein Vektorfeld $\vec{V}\colon D \to \mathbb{R}^n$ ist die JACOBI-Matrix quadratisch, wir können daher von symmetrischen und antisymmetrischen Matrizen reden.

(Zur Erinnerung: Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt symmetrisch, wenn sie gleich ihrer transponierten Matrix ist, d.h. $a_{ij} = a_{ji}$ für alle Indexpaare ij, und sie heißt antisymmetrisch, wenn sie gleich dem Negativen ihrer transponierten Matrix ist, d.h. $a_{ij} = -a_{ji}$ und damit insbesondere auch $a_{ii} = -a_{ii} = 0$ für alle i.)

Eine beliebige Matrix $M \in \mathbb{R}^{n\times n}$ läßt sich als Summe einer symmetrischen und einer antisymmetrischen Matrix schreiben:

$$M = S + A \quad \text{mit} \quad S = \tfrac{1}{2}(M + {}^tM) \quad \text{und} \quad A = \tfrac{1}{2}(M - {}^tM)\,.$$

In diesem Abschnitt wollen wir den symmetrischen Anteil S der JACOBI-Matrix genauer untersuchen, im nächsten den antisymmetrischen Anteil A vor allem im Fall $n = 3$.

Für kleine Vektoren $\vec{h}$ ist nach obiger Formel

$$\vec{V}(\boldsymbol{x} + \vec{h}) = V(\boldsymbol{x}) + J_{\vec{V}}(\boldsymbol{x})\vec{h} + o\big(|\vec{h}|\big) = \vec{V}(\boldsymbol{x}) + S\vec{h} + A\vec{h} + o\big(|\vec{h}|\big)\,.$$

Nach **[HM I]**, §7e sind alle Eigenwerte der symmetrischen Matrix S reell, und der $\mathbb{R}^n$ hat eine Basis aus Eigenvektoren von S.

Falls alle Eigenwerte von S positiv sind, hat die vom symmetrischen Teil der JACOBI-Matrix beschriebene Änderung $S\vec{h}$ des Feldes für jeden Eigenvektor $\vec{h}$ dieselbe Richtung wie $\vec{h}$, zeigt also weg vom Punkt $\boldsymbol{x}$, wie dies im linken Teil von Abbildung 49 zu sehen ist. Da es eine Basis aus Eigenvektoren gibt, gilt zumindest letzteres auch für beliebige Vektoren $\vec{h}$; man bezeichnet $\boldsymbol{x}$ daher in so einem Fall als eine *Quelle* des Vektorfelds $\vec{V}$.

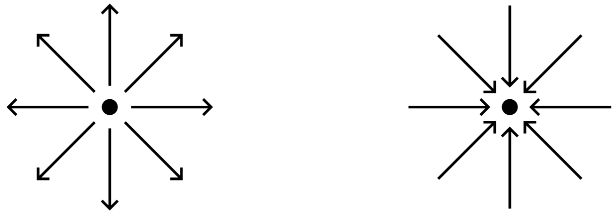


Abb. 49: Quellen und Senken

Falls alle Eigenvektoren negativ sind, gehen die Änderungen jeweils in Gegenrichtung, es sieht also ungefähr so aus wie im rechten Teil von Abbildung 49, und man spricht von einer *Senke.*

Im allgemeinen werden weder *alle* Eigenwerte positiv noch *alle* Eigenwerte negativ sein; als Maßzahl für die Eigenschaft eines Punktes eher eine Quelle oder eher eine Senke zu sein, bietet sich dann die *Summe* aller Eigenwerte von S an:

Definition: Die *Divergenz* $\operatorname{div}\vec{V}(\boldsymbol{x})$ eines Vektorfelds $\vec{V}$ im Punkt $\boldsymbol{x}$ ist die Summe der Eigenwerte des symmetrischen Anteils der JACOBI-Matrix $J_{\vec{V}}(\boldsymbol{x})$. Falls sie im gesamten Definitionsbereich von $\vec{V}$ verschwindet, bezeichnen wir das Feld als *quellenfrei.*

Mit Beispielen wollen wir warten, bis wir diese Definition besser verstanden haben und auch den antisymmetrischen Anteil der JACOBI-Matrix mitbehandeln können.

Gelegentlich, aber selten, liest man statt *Divergenz* auch das Wort *Quellendichte.*

Die obige Definition macht zwar klar, worum es bei der Divergenz geht, sie ist aber nicht sonderlich gut zu deren Berechnung geeignet: Die Bestimmung aller Eigenwerte einer Matrix kann schließlich ziemlich aufwendig sein. Für die *Summe* der Eigenwerte gibt es glücklicherweise ein sehr viel einfacheres Verfahren:

Satz: Die Matrix $M = (m_{ij}) \in k^{n\times n}$ über dem Körper k habe die Eigenwerte $\lambda_1, \dots, \lambda_n$. Dann ist

$$\lambda_1 + \cdots + \lambda_n = m_{11} + \cdots + m_{nn}\,.$$

(Man beachte, daß hier nicht vorausgesetzt wurde, daß die Eigenwerte λ_i in k liegen; für eine beliebige reelle Matrix etwa sind die Eigenwerte ja bekanntlich im allgemeinen *komplexe* Zahlen. Auch bei beliebigen Körpern k gibt es stets einen Erweiterungskörper K, über dem das charakteristische Polynom – mit Vielfachheit gezählt – n Nullstellen hat, allerdings wollen wir hier auf Einzelheiten nicht eingehen.)

Beweis: Die Eigenwerte von M sind die Nullstellen des charakteristischen Polynoms $\det(M - \lambda E)$, wobei $E \in k^{n\times n}$ die Einheitsmatrix bezeichnet. Mit

$$c_{ij} = \begin{cases} m_{ij} & \text{falls } i \neq j \\ m_{ii} - \lambda & \text{falls } i = j \end{cases}$$

ist dann

$$\det(M - \lambda E) = \sum_{\pi\in\mathfrak{S}_n} \varepsilon(\pi) c_{1\pi(1)} \cdots c_{n\pi(n)}\,.$$

Falls π die Identität ist, erhalten wir den Summanden

$$(m_{11} - \lambda)\cdots(m_{nn} - \lambda) = (-1)^n\lambda^n + (-1)^{n-1}\sum_{i=1}^n m_{ii}\lambda^{n-1} + P(\lambda)\,,$$

wobei $P(\lambda)$ ein Polynom vom Grad kleiner oder gleich $n - 2$ ist.

Falls π eine nichtidentische Permutation ist, gibt es mindestens zwei Zahlen i, so daß $\pi(i) \neq i$ ist, denn gäbe es nur eine, so müßte auch diese

wegen der Bijektivität von π auf sich selbst abgebildet werden, d.h. π wäre doch gleich der Identität. Damit gibt es im Produkt

$$c_{1\pi(1)} \cdots c_{n\pi(n)}$$

mindestens zwei Faktoren, in denen $c_{i\pi(i)} = m_{i\pi(i)}$ ist und somit höchstens $n - 2$ Faktoren, in denen λ vorkommen kann, das Produkt ist also ein Polynom vom Grad höchstens $n - 2$ in λ. Damit ist insgesamt

$$\begin{aligned}\det(M - \lambda E) &= \sum_{\pi\in\mathfrak{S}_n} \varepsilon(\pi) c_{1\pi(1)} \cdots c_{n\pi(n)} \\ &= (-1)^n\lambda^n + (-1)^{n-1}\sum_{i=1}^n m_{ii}\lambda^{n-1} + Q(\lambda)\,,\end{aligned}$$

wobei auch $Q(\lambda)$ ein Polynom vom Grad kleiner oder gleich $n - 2$ ist.

Wir können das charakteristische Polynom aber auch aus den Eigenwerten berechnen: Jedes Polynom n-ten Grades mit Nullstellen $\lambda_1, \dots, \lambda_n$ hat die Form

$$\alpha(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)\,,$$

wobei α der Koeffizient von λ^n ist. Somit ist

$$\begin{aligned}\det(M - \lambda E) &= (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n) \\ &= (-1)^n\lambda^n + (-1)^{n-1}\left(\sum_{i=1}^n \lambda_i\right)\cdot\lambda^{n-1} + R(\lambda)\,,\end{aligned}$$

wobei auch $R(\lambda)$ wieder ein Polynom vom Grad höchstens $n-2$ ist. Also ist der Koeffizient von λ^{n-1} wirklich gleich der Summe der Eigenwerte; da wir aber oben auch gezeigt haben, daß er gleich der Summe der Diagonalelemente von M ist, folgt

$$\sum_{i=1}^n \lambda_i = \sum_{i=1}^n m_{ii}\,,$$

wie behauptet. ■

Somit ist die Summe aller Eigenwerte einer Matrix gleich der Summe ihrer Diagonalelemente. Da sich Eigenwerte bei einem Basiswechsel nicht ändern, kann sich auch die Summe der Diagonalelemente bei einem Basiswechsel nicht ändern, sie ist daher – wie auch die Determinante und die Eigenwerte – nur von der durch die Matrix vermittelten linearen Abbildung abhängig und hat daher als wichtige Invariante einer linearen Abbildung Anspruch auf einen eigenen Namen:

Definition: Die Summe der Diagonalelemente einer quadratischen Matrix $M = (m_{ij})$ heißt *Spur* der Matrix, in Zeichen

$$\operatorname{Spur} M = \sum_{i=1}^{n} m_{ii} .$$

Auf englisch heißt Spur *trace* (auf französisch, abgesehen von der etwas anderen Aussprache auch), daher findet man in vielen (teilweise sogar deutschen) Texten auch die Abkürzung tr M. Leute, die zu faul sind, das Wort Spur auszuschreiben, kürzen es auch gelegentlich ab mit Sp.

Nach obigem Satz ist die Spur einer Matrix gleich der Summe ihrer Eigenwerte, genau wie die Determinante gleich deren Produkt ist. Beide sind invariant unter Basiswechsel, d.h. für eine invertierbare Matrix $C \in k^{n\times n}$ ist

$$\det C^{-1}MC = \det M \quad \text{und} \quad \operatorname{Spur} C^{-1}MC = \operatorname{Spur} M .$$

Nach diesem Ausflug zurück zur linearen Algebra wenden wir uns wieder dem augenblicklichem Thema zu, der JACOBI-Matrix eines Vektorfelds. Nach dem gerade bewiesenen Satz ist die Divergenz des Vektorfelds gleich der Spur des symmetrischen Anteils S der JACOBI-Matrix, was ihre Berechnung deutlich vereinfacht. Tatsächlich ist die Situation sogar noch ein bißchen besser: Da alle Diagonaleinträge einer antisymmetrischen Matrix verschwinden, sind die Diagonaleinträge des symmetrischen Anteils gleich denen der JACOBI-Matrix selbst, d.h. die Divergenz ist einfach gleich der Spur der JACOBI-Matrix. Damit läßt sie sich leicht explizit berechnen:

$$\operatorname{div} \vec{V}(\boldsymbol{x}) = \operatorname{Spur} J_{\vec{V}}(\boldsymbol{x}) = \frac{\partial V_1}{\partial x_1}(\boldsymbol{x}) + \cdots + \frac{\partial V_n}{\partial x_n}(\boldsymbol{x}) .$$

d) Die Rotation eines dreidimensionalen Vektorfelds

Im vorigen Abschnitt haben wir den antisymmetrischen Anteil der JACOBI-Matrix einfach vernachlässigt; dies ist natürlich im allgemeinen nicht zulässig. Hier wollen wir diesen Anteil genauer untersuchen.

Für $n = 1$ ist jede $n \times n$-„Matrix“ symmetrisch, es gibt als keinen antisymmetrischen Anteil.

Für $n = 2$ hat jede antisymmetrische Matrix A die Form

$$\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} ,$$

und im Fall des antisymmetrischen Anteils der JACOBI-Matrix ist

$$a = \frac{1}{2}\left(\frac{\partial f_1}{\partial y} - \frac{\partial f_2}{\partial x}\right) .$$

Damit ist der antisymmetrische Anteil durch diese eine Zahl bestimmt.

Für $n > 3$ gibt es keine einfachere Darstellung des antisymmetrischen Anteils als die schiefsymmetrische Matrix selbst, bleibt also noch der für die Anwendungen besonders wichtige Fall $n = 3$.

Eine antisymmetrische 3×3-Matrix A hat die Form

$$A = \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix} .$$

Für einen Vektor $\vec{x} \in \mathbb{R}^3$ ist daher

$$A\vec{x} = \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha y + \beta z \\ -\alpha x + \gamma z \\ -\beta x - \gamma y \end{pmatrix} .$$

Diese spezielle Form des Matrix-Vektor-Produkts sollte Erinnerungen an das Vektorprodukt im $\mathbb{R}^3$ wecken: Nach **[HM1]**, §1, ist

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} bz - cy \\ cx - az \\ ay - bx \end{pmatrix} ,$$

d.h.
$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
und damit
$$\begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\gamma \\ \beta \\ -\alpha \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} .$$
Somit entspricht im Dreidimensionalen die Multiplikation mit einer schiefsymmetrischen Matrix genau dem Vektorprodukt (das im übrigen auch nur im Dreidimensionalen definierbar ist).

Der Vektor $\vec{a}$ läßt sich im vorliegenden Fall leicht ausrechnen: Die JACOBI-Matrix
$$J_{\vec{V}}(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial V_1}{\partial x} & \frac{\partial V_1}{\partial y} & \frac{\partial V_1}{\partial z} \\ \frac{\partial V_2}{\partial x} & \frac{\partial V_2}{\partial y} & \frac{\partial V_2}{\partial z} \\ \frac{\partial V_3}{\partial x} & \frac{\partial V_3}{\partial y} & \frac{\partial V_3}{\partial z} \end{pmatrix}$$
eines dreidimensionalen Vektorfelds $\vec{V}$ hat antisymmetrischen Anteil
$$A = \frac{1}{2}\left(A - {}^tA\right) = \frac{1}{2}\begin{pmatrix} 0 & \frac{\partial V_1}{\partial y} - \frac{\partial V_2}{\partial x} & \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} & 0 & \frac{\partial V_2}{\partial z} - \frac{\partial V_3}{\partial y} \\ \frac{\partial V_3}{\partial x} - \frac{\partial V_1}{\partial z} & \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} & 0 \end{pmatrix} ,$$
und wie wir uns gerade überlegt haben, ist für jeden Vektor $\vec{v}$
$$A\vec{v} = \vec{a} \times \vec{v} \quad \text{mit} \quad \vec{a} = \frac{1}{2}\begin{pmatrix} \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} \\ \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \end{pmatrix} .$$
Da der Faktor $\frac{1}{2}$ in konkreten Rechnungen nur hinderlich wäre, betrachten wir den Vektor $2\vec{a}$ und definieren:

Definition: $\vec{V}: D \to \mathbb{R}^3$ sei ein Vektorfeld auf der offenen Teilmenge $D \subseteq \mathbb{R}^3$. Dann bezeichnen wir das Vektorfeld
$$\operatorname{rot} V: D \to \mathbb{R}^3; \quad \boldsymbol{x} \mapsto \begin{pmatrix} \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} \\ \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \end{pmatrix}$$

als Rotation von $\vec{V}$.

Für ein Vektorfeld $\vec{V} \in \mathcal{C}^1(D, \mathbb{R}^3)$, dessen JACOBI-Matrix symmetrischen Anteil S hat, ist also
$$\vec{V}(\boldsymbol{x} + \vec{h}) = \vec{V}(\boldsymbol{x}) + S\vec{h} + \tfrac{1}{2}(\operatorname{rot}\vec{V}) \times \vec{h} + o\big(|\vec{h}|\big) .$$

Der Name *Rotation* bedarf einer Erklärung, und dazu betrachten wir wieder den Vektor
$$\vec{a} = 2 \operatorname{rot} \vec{V}(\boldsymbol{x}) .$$

Nach Definition des Vektorprodukts in **[HM1]**, §1, steht $\vec{a} \times \vec{h}$ senkrecht sowohl auf $\vec{a}$ als auch auf $\vec{h}$. Bedenkt man nun noch, daß Vektoren Äquivalenzklassen von Pfeilen sind, die man sich nicht unbedingt so vorstellen muß, daß sie alle im selben Punkt beginnen, bietet sich folgende Interpretation an: Wenn wir den Vektor $\vec{h}$ um die von $\vec{a}$ aufgespannte Achse drehen, zeigt $\vec{a} \times \vec{h}$ in Richtung der Tangente der Bahn des Endpunkts von $\vec{h}$. (Damit dies stimmt, muß die Achse so orientiert werden, daß sich $\vec{h}$ im mathematisch positiven, d.h. Gegenuhrzeigersinn, um $\vec{a}$ dreht. Von der Ebene der Uhr aus gesehen zeigt dann $\vec{a}$ nach oben, $\vec{h}$ zu dem sich drehenden Punkt, und $\vec{a} \times \vec{h}$ in Richtung des Gegenuhrzeigersinns, wie man sich z.B. mit der Dreifingerregel leicht veranschaulicht.)

Der antisymmetrische Anteil der JACOBI-Matrix sollte also etwas mit einer Drehbewegung zu tun habe, genauer gesagt mit deren Geschwindigkeitsvektor, denn dieser zeigt in Richtung der Tangente der Bahnkurve.

Um dies genauer zu untersuchen, betrachten wir zunächst eine einfache Drehung eines Vektors um die z-Achse mit Drehwinkel φ. Diese ist aufgrund der Definition von Sinus und Cosinus ausdrückbar durch die lineare Abbildung
$$\begin{aligned} x &\mapsto x\cos\varphi - y\sin\varphi \\ y &\mapsto x\sin\varphi + y\cos\varphi \quad . \\ z &\mapsto z \end{aligned}$$

Als nächstes wollen wir einen Punkt $\boldsymbol{x} \in \mathbb{R}^3$ um eine beliebige Achse drehen; der Drehwinkel sei weiterhin φ. Dazu sei $\vec{a}$ der Einheitsvektor in

Richtung der Achse; er sei so orientiert, daß der Drehsinn (bei positivem Winkel φ) gleich dem Gegenuhrzeigersinn ist. Dann gelten in den auf $\vec{a}$ senkrecht stehenden Ebenen im wesentlichen dieselbe Formeln, falls wir nur die Koordinateneinheitsvektoren der x- und der y-Achse ersetzen durch zwei aufeinander und auf $\vec{a}$ senkrecht stehende gleichlange Vektoren. Als ein solcher Vektor bietet sich der Verbindungsvektor $\vec{r}$ von der Achse zum Punkt $\boldsymbol{x}$ an, d.h. jener Vektor, der den Abstand zwischen $\boldsymbol{x}$ und der Achse beschreibt. Ein sowohl auf $\vec{r}$ als auch auf $\vec{a}$ senkrecht stehender Vektor ist $\vec{s} = \vec{a} \times \vec{r}$; da $\vec{a}$ die Länge eins hat, ist $\vec{s}$ genauso lang wie $\vec{r}$.

In der von $\vec{r}$ und $\vec{s}$ aufgespannten Ebene bildet die Drehung um den Winkel φ somit den Vektor $\alpha\vec{r} + \beta\vec{s}$ ab auf

$$(\alpha \cos\varphi - \beta \sin\varphi)\,\vec{r} + (\alpha \sin\varphi + \beta \cos\varphi)\,\vec{s}\,.$$

Nun zerlegen wir den Ortsvektor $\vec{x}$ eines Punktes $\boldsymbol{x}$ in einen Vektor $\vec{x}_a$ auf der Achse und den radialen, d.h. senkrecht auf der Achse stehenden, Anteil $\vec{r}$.

$\vec{x}_a$ ist die Projektion von $\vec{x}$ auf die Drehachse; da wir $\vec{a}$ als Einheitsvektor vorausgesetzt haben, ist dies nach den allgemeinen Regeln der Vektorrechnung einfach

$$\vec{x}_a = (\vec{a} \cdot \vec{x})\,\vec{a}\,,$$

denn die Länge des Vektors wird bei der Projektion mit dem Cosinus des Winkels zur Achse multipliziert, und die Richtung des auf die Achse projezierten Vektors ist natürlich gleich der Richtung der Achse.

Damit läßt sich auch die radiale Komponente problemlos berechnen: Da beide Komponenten zusammen den Vektor $\vec{x}$ ergeben müssen, ist

$$\vec{r} = \vec{x} - \vec{x}_a = \vec{x} - (\vec{a} \cdot \vec{x})\,\vec{a}\,.$$

Nun ist $\vec{x}_a$ parallel zu $\vec{a}$, und das Vektorprodukt zweier paralleler Vektoren verschwindet. Deshalb können wir auch schreiben

$$\vec{s} = \vec{a} \times \vec{r} = \vec{a} \times \big(\vec{x} - (\vec{x} \cdot \vec{a})\vec{a}\big) = \vec{a} \times \vec{x}\,.$$

Da die Axialkomponente eines Vektors bei Drehungen erhalten bleibt, wird speziell der Vektor $\vec{x} = \vec{x}_a + \vec{r}$ also abgebildet auf

$$\begin{aligned}
&\vec{x}_a + \cos\varphi\;\vec{r} + \sin\varphi\;\vec{s}\\
&= (\vec{x} \cdot \vec{a})\,\vec{a} + \cos\varphi\;\big(\boldsymbol{x} - (\vec{x} \cdot \vec{a})\,\vec{a}\big) + \sin\varphi\;\vec{a} \times \vec{x}\\
&= (1 - \cos\varphi)(\vec{x} \cdot \vec{a})\,\vec{a} + \cos\varphi\;\vec{x} + \sin\varphi\;\vec{a} \times \vec{x}\,.
\end{aligned}$$

Um den Geschwindigkeitsvektor einer Drehung auszurechnen, müssen wir diese zunächst zu einem dynamischen Prozeß machen, d.h. anstelle einer Drehung mit konstantem Drehwinkel φ betrachten wir eine Drehung mit Winkelgeschwindigkeit ω. Wir gehen wieder aus von einem festen Punkt $\boldsymbol{x}$ und betrachten dessen Position zum Zeitpunkt t, d.h. wir betrachten die Funktion

$$\vec{f}\colon \begin{cases} \mathbb{R} & \to \mathbb{R}^3 \\ t & \mapsto (1 - \cos\omega t)(\vec{x} \cdot \vec{a})\;\vec{a} + \cos\omega t\;\vec{x} + \sin\omega t\;\vec{a} \times \vec{x}\,. \end{cases}$$

Der Geschwindigkeitsvektor im Punkt $\boldsymbol{x}$ ist die Ableitung dieser Funktion im Nullpunkt, also der Vektor

$$\frac{d\vec{f}}{dt}(0) = \big(\omega \sin\omega t\;(\vec{x} \cdot \vec{a})\vec{a} - \omega \sin\omega t\;\vec{x} + \omega\cos\omega t\;\vec{a} \times \vec{x}\big)_{|t=0} = \omega\vec{a} \times \vec{x}\,.$$

Mit anderen Worten: Bei der Drehung mit Winkelgeschwindigkeit ω um die von $\vec{a}$ aufgespannte Achse ist der Geschwindigkeitsvektor im Punkt $\boldsymbol{x}$ gleich $\omega\vec{a} \times \vec{x}$.

Der Vektor $\omega\vec{a}$ läßt sich auffassen als eine vektorielle Winkelgeschwindigkeit (die in der Physik in der Tat genau so definiert wird), und für einen beliebigen Vektor $\vec{\omega} \in \mathbb{R}^3$ ist das Vektorfeld

$$\vec{V}\colon \mathbb{R}^3 \to \mathbb{R}^3; \quad \boldsymbol{x} \mapsto \vec{\omega} \times \vec{x}$$

das Feld der Geschwindigkeitsvektoren zu einer Rotation mit vektorieller Winkelgeschwindigkeit $\vec{\omega}$.

Es sei noch darauf hingewiesen, daß $\operatorname{rot}\vec{V}$ in der englischsprachigen Literatur meist als $\operatorname{curl}\vec{V}$ geschrieben wird nach dem englischen Wort *curl* = Locke.

e) Erste Beispiele

Zum besseren Verständnis von Divergenz und Rotation und um den Umgang mit den beiden Begriffen zu üben, wollen wir sie für zwei einfache Felder ausrechnen: Für das elektrische Feld einer Punktladung und für das Magnetfeld eines stromdurchflossenen Leiters.

1) Das elektrische Feld einer Punktladung: Eine Punktladung q im Nullpunkt erzeugt nach dem COULOMBschen Gesetz ein elektrisches Feld $\vec{E}$ der Stärke

$$\vec{E}(\boldsymbol{x}) = \frac{q}{4\pi\varepsilon_0}\,\frac{\vec{x}}{\left|\vec{x}\right|^3}\,,$$

denn die Feldstärke nimmt mit der Entfernung vom Nullpunkt quadratisch ab und hat die Richtung (oder – je nach Vorzeichen von q – auch Gegenrichtung) des Ortsvektors., d.h. der Einheitsvektor der Richtung ist $\pm\vec{x}/\left|\vec{x}\right|$, und er muß noch durch $\left|\vec{x}\right|^2$ dividiert werden.

Das Feld $\vec{E}$ ist nur definiert auf $\mathbb{R}^3 \setminus \{(0,0,0)\}$; im Nullpunkt haben wir einen Nenner Null, und auch physikalisch ist es nicht sinnvoll, nach dem Feld *im Innern* einer Punktladung zu fragen: Punktladungen sind nun einmal mathematische Idealisierungen, und wirklich anwendbar ist das COULOMBsche Gesetz ohnehin erst ab einem gewissen Mindestabstand, jenseits dessen andere, stärkere Wechselwirkungen keine Rolle mehr spielen.

In Koordinaten ist also

$$\vec{E}(\boldsymbol{x}) = \begin{pmatrix} E_1(\boldsymbol{x}) \\ E_2(\boldsymbol{x}) \\ E_3(\boldsymbol{x}) \end{pmatrix} = \frac{q}{4\pi\varepsilon_0\left|\vec{x}\right|^3}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\,.$$

Zur Berechnung der Divergenz müssen wir die erste Komponente nach x, die zweite nach y und die dritte nach z ableiten und dann die drei Ableitungen addieren. Beachten wir, daß

$$\left|\vec{x}\right|^3 = (x^2+y^2+z^2)^{3/2}$$

ist, folgt sofort aus der Quotientenregel, daß

$$\begin{aligned}\frac{\partial E_1}{\partial x}(\boldsymbol{x}) &= \frac{q}{4\pi\varepsilon_0}\,\frac{\left|\vec{x}\right|^3 - x\cdot\frac{3}{2}\cdot 2x\left|\vec{x}\right|}{\left|\vec{x}\right|^6}\\ &= \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{\left|\vec{x}\right|^3} - \frac{3x^2}{\left|\vec{x}\right|^5}\right)\\ &= \frac{q}{4\pi\varepsilon_0\left|\vec{x}\right|^5}(-2x^2+y^2+z^2)\,.\end{aligned}$$

Entsprechend ist

$$\frac{\partial E_2}{\partial y}(\boldsymbol{x}) = \frac{q}{4\pi\varepsilon_0\left|\vec{x}\right|^5}(x^2-2y^2+z^2)$$

und

$$\frac{\partial E_3}{\partial z}(\boldsymbol{x}) = \frac{q}{4\pi\varepsilon_0\left|\vec{x}\right|^5}(x^2+y^2-2z^2)\,,$$

also

$$\operatorname{div}\vec{E}(\boldsymbol{x}) = \frac{q}{4\pi\varepsilon_0\left|\vec{x}\right|^5}(-2x^2+y^2+z^2+x^2-2y^2+z^2+x^2+y^2-2z^2) = 0\,.$$

Zur Berechnung der Rotation benötigen wir die Einträge der JACOBI-Matrix, die nicht in der Diagonale stehen; dabei werden die Komponenten von $\vec{E}$ nach jenen Variablen abgeleitet, die *nicht* im Zähler stehen. Die Situation ist also stets dieselbe, und es genügt, wenn wir beispielsweise die Ableitung von E_1 nach y ausrechnen:

$$\frac{\partial E_1}{\partial y} = \frac{q}{4\pi\varepsilon_0}\left(-\frac{3}{2}\right)\frac{x\cdot 2y}{\left|\vec{x}\right|^5}\,.$$

Entsprechend ist

$$\frac{\partial E_2}{\partial x} = \frac{q}{4\pi\varepsilon_0}\left(-\frac{3}{2}\right)\frac{y\cdot 2x}{\left|\vec{x}\right|^5}\,,$$

die x-Komponente von $\operatorname{rot}\vec{E}$ als Differenz dieser beiden Ableitungen verschwindet also, und genauso überzeugt man sich auch vom Verschwinden der anderen Komponenten.

Somit ist also das Feld einer Punktladung außerhalb des Punktes selbst sowohl quellen- als auch wirbelfrei, genau wie es nach den MAXWELLschen Gleichungen auch sein muß.

2) Das Magnetfeld eines stromdurchflossenen Leiters: Das Magnetfeld eines von einem Strom der Stärke I durchflossenen geradlinigen Drahts ist nach dem BIOT-SAVARTschen Gesetz

$$\vec{H}(\boldsymbol{x}) = \frac{I}{2\pi\left|\vec{r}\right|^2}\,\vec{a}\times\vec{x}\,,$$

wobei $\vec{a}$ der Einheitsvektor in Richtung des Stroms und $\vec{r}$ der Abstandsvektor zwischen dem Draht und dem Punkt $\vec{x}$ ist. $\vec{H}$ ist überall im $\mathbb{R}^3$ erklärt, außer auf der von $\vec{a}$ aufgespannten Geraden, wo $\vec{r}$ der Nullvektor ist.

Zum einfacherer Rechnen wählen wir das Koordinatensystem so, daß $\vec{a}$ der Einheitsvektor in Richtung der positiven z-Achse ist; für den Punkt mit Koordinaten (x, y, z) ist dann

$$\vec{r} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a}\times\vec{x} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix},$$

d.h.

$$\vec{H}(\boldsymbol{x}) = \frac{I}{4\pi(x^2+y^2)}\begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}.$$

Dieses Vektorfeld ist erklärt auf $\mathbb{R}^3$ minus der z-Achse; auf der z-Achse, d.h. im Innern des Leiters, bekommen wir eine Null in den Nenner.

Die drei Komponenten dieses Felds könnten wir nun wie oben nach den drei Variablen x, y, z ableiten, um so Rotation und Divergenz zu berechnen; es gibt allerdings auch eine andere Möglichkeit, die JACOBI-Matrix eines solchen Felds zu bestimmen: Wie auch das $\vec{E}$-Feld im obigen Beispiel läßt sich $\vec{H}$ schreiben als Produkt einer skalaren Funktion mit einer *linearen* Abbildung $\mathbb{R}^3 \to \mathbb{R}^3$. Solche Funktionen, bei denen alle Nichtlinearität in einer skalaren Funktion steckt, bezeichnen wir als *quasilinear:*

Definition: Eine Funktion $\vec{f}: D \to \mathbb{R}^m$ auf einer Teilmenge $D \subseteq \mathbb{R}^n$ heißt *quasilinear,* wenn es eine Funktion $\varphi: D \to \mathbb{R}$ und eine Matrix $A \in \mathbb{R}^{n\times m}$ gibt, so daß $\vec{f}(\boldsymbol{x}) = \varphi(\boldsymbol{x}) \cdot A\vec{x}$ ist.

In diesem Sinne ist $\vec{H}(x)$ quasilinear mit

$$\varphi(\boldsymbol{x}) = \frac{I}{4\pi(x^2+y^2)} \quad \text{und} \quad A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und das obige Feld $\vec{E}$ ist quasilinear mit

$$\varphi(\boldsymbol{x}) = \frac{1}{4\pi\varepsilon_0\left|\vec{r}\right|^3} \quad \text{und} \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Lemma: Für eine quasilineare Funktion $\vec{f}: D \to \mathbb{R}^m$ auf einer Teilmenge $D \subseteq \mathbb{R}^n$ ist

$$J_{\vec{f}}(\boldsymbol{x}) = \varphi(\boldsymbol{x}) \cdot A + (A\vec{x}) \cdot {}^t\!\big(\operatorname{grad}\varphi(\boldsymbol{x})\big)\,.$$

In dieser Formel steht der erste Malpunkt für die Multiplikation eines Skalars mit einer Matrix, der zweite für die Matrixmultiplikation des *Spaltenvektors* $A\vec{x}$ mit dem zum *Zeilenvektor* transponierten Gradienten; das Produkt ist eine $n \times m$-Matrix. Ausgeschrieben wird die Formel zu

$$\frac{\partial f_i}{\partial x_j} = \varphi(\boldsymbol{x}) \cdot a_{ij} + \left(\sum_{\nu=1}^{n} a_{i\nu}x_\nu\right) \cdot \frac{\partial\varphi}{\partial x_j}\,.$$

Zum *Beweis* leiten wir $f_i(\boldsymbol{x}) = \varphi(\boldsymbol{x})\sum_{\nu=1}^{n} a_{i\nu}x_\nu$ nach x_j ab; nach der Produktregel ist, wie gewünscht,

$$\begin{aligned}\frac{\partial f_i}{\partial x_j}(\boldsymbol{x}) &= \varphi(\boldsymbol{x})\frac{\partial}{\partial x_j}\left(\sum_{\nu=1}^{n} a_{i\nu}x_\nu\right) + \frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\sum_{\nu=1}^{n} a_{i\nu}x_\nu \\ &= \varphi(\boldsymbol{x}) \cdot a_{ij} + \frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\sum_{\nu=1}^{n} a_{i\nu}x_\nu\,.\end{aligned}$$

■

Dieses Lemma wenden wir gleich an auf das Feld $\vec{H}(\boldsymbol{x})$; hier ist

$$\frac{\partial\varphi}{\partial x}(\boldsymbol{x}) = \frac{-2x\cdot I}{4\pi(x^2+y^2)^2} \quad\text{und}\quad \frac{\partial\varphi}{\partial y}(\boldsymbol{x}) = \frac{-2y\cdot I}{4\pi(x^2+y^2)^2}\,,$$

d.h. $\operatorname{grad}\varphi(\boldsymbol{x}) = \dfrac{-2I}{4\pi(x^2+y^2)^2}\begin{pmatrix} x\\ y\\ 0\end{pmatrix}$. Somit ist

$$\begin{aligned}
J_{\vec{H}}(\boldsymbol{x}) &= \frac{I}{4\pi(x^2+y^2)}\begin{pmatrix} 0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\\
&\quad + \begin{pmatrix} 0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} x\\ y\\ z\end{pmatrix}\cdot\frac{-2I}{4\pi(x^2+y^2)^2}\begin{pmatrix} x & y & 0\end{pmatrix}\\
&= \frac{I}{4\pi(x^2+y^2)^2}\left(\begin{pmatrix} 0 & -x^2-y^2 & 0\\ x^2+y^2 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\right.\\
&\quad \left. + \begin{pmatrix} -y\\ x\\ 0\end{pmatrix}\cdot\begin{pmatrix} -2x & -2y & 0\end{pmatrix}\right)\\
&= \frac{I}{4\pi(x^2+y^2)^2}\left(\begin{pmatrix} 0 & -x^2-y^2 & 0\\ x^2+y^2 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\right.\\
&\quad \left. + \begin{pmatrix} 2xy & 2y^2 & 0\\ -2x^2 & -2xy & 0\\ 0 & 0 & 0\end{pmatrix}\right)\\
&= \frac{I}{4\pi(x^2+y^2)^2}\begin{pmatrix} 2xy & y^2-x^2 & 0\\ y^2-x^2 & -2xy & 0\\ 0 & 0 & 0\end{pmatrix}.
\end{aligned}$$

Diese Matrix ist symmetrisch und hat Spur Null, d.h auch hier ist

$$\operatorname{div}\vec{H}(\boldsymbol{x}) = 0 \quad\text{und}\quad \operatorname{rot}\vec{H}(\boldsymbol{x}) = \vec{0}\,,$$

und auch dies natürlich wieder in Übereinstimmung mit den MAXWELLschen Gleichungen.

Da die Feldlinien des hier betrachteten Felds Kreise sind, mag es auf den ersten Blick seltsam erscheinen, daß die Rotation von $\vec{H}$ verschwindet.

Der Grund liegt darin, daß alle Kreise ihren Mittelpunkt auf der z-Achse haben, die wir in der obigen Rechnung ausschließen mußten: Die Rotation in einem Punkt $\boldsymbol{x}$ mißt die Rotation um $\boldsymbol{x}$, und beim betrachteten Feld rotiert eben alles um die z-Achse, so daß die Rotation überall sonst verschwindet.

f) Allgemeine Rechenregeln

Nachdem wir gerade mit relativ großem Aufwand für zwei Beispiele Divergenz und Rotation berechnet haben, wollen wir nun einige Rechenregeln herleiten, die uns unter anderem zeigen werden, daß zumindest ein Teil der obigen Rechnungen hätte vermieden werden können.

Zunächst sei auf eine oftmals nützliche Abkürzung hingewiesen: Wir hatten den Gradienten einer Funktion f nach HAMILTON auch kurz als ∇f geschrieben. Die Definition

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1}\\ \vdots\\ \frac{\partial f}{\partial x_n}\end{pmatrix}$$

können wir *formal* auch so auffassen, daß hier der *Vektor*

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1}\\ \vdots\\ \frac{\partial}{\partial x_n}\end{pmatrix}$$

mit dem Skalar f multipliziert wird. Diese Interpretation hat zwar ihre Grenzen, denn beispielsweise ist $f\nabla$ sicherlich etwas völlig anderes als ∇f, aber sie ist doch sehr gebräuchlich, da sie sich genauso auch für Divergenz und Rotation anwenden läßt: Für eine Vektorfeld $\vec{V}$ mit Komponenten $\vec{V}_1,\dots,\vec{V}_n$ ist, wieder rein formal, das „Skalarprodukt“

$$\nabla\vec{V} = \begin{pmatrix} \frac{\partial}{\partial x_1}\\ \vdots\\ \frac{\partial}{\partial x_n}\end{pmatrix}\cdot\begin{pmatrix} V_1\\ \vdots\\ V_n\end{pmatrix} = \sum_{i=1}^{n}\frac{\partial V_i}{\partial x_i} = \operatorname{div}\vec{V}\,,$$

und im Falle $n = 3$ ist das „Vektorprodukt“

$$\nabla \times \vec{V} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \cdot \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial V_3}{\partial x_2} - \frac{\partial V_2}{\partial x_3} \\ \frac{\partial V_1}{\partial x_3} - \frac{\partial V_3}{\partial x_1} \\ \frac{\partial V_2}{\partial x_1} - \frac{\partial V_1}{\partial x_2} \end{pmatrix} = \operatorname{rot} \vec{V} .$$

Solange man sich der Grenzen der Analogie zwischen ∇ und einem „gewöhnlichen“ Vektor bewußt ist, können diese Abkürzungen recht nützlich sein.

Wir wollen uns als nächstes überlegen, was passiert, wenn wir zwei der drei Operatoren grad, div und rot hintereinanderausführen – sofern dies möglich ist.

Der Gradient einer skalaren Funktion $f\colon D \to \mathbb{R}$ auf einer offenen Teilmenge $D \subseteq \mathbb{R}^n$ ist ein Vektorfeld; es ist also möglich, dessen Divergenz und Rotation zu berechnen. Um keine Probleme mit der Existenz und der Reihenfolge von Ableitungen zu haben, setzen wir voraus, daß f in $\mathcal{C}^2(D, \mathbb{R})$ liegt; dann ist

$$\operatorname{div} \operatorname{grad} f = \operatorname{div} \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2} .$$

Die rechte Seite schreiben wir auch kurz als

$$\Delta f \underset{\text{def}}{=} \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2}$$

und bezeichnen

$$\Delta \underset{\text{def}}{=} \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$$

als LAPLACE-Operator.

(Der LAPLACE-Operator ist älter als Divergenz, Gradient, Rotation und so weiter; er spielt eine große Rolle sowohl bei Wellen als auch bei Potentialfeldern. HAMILTON wählte das Zeichen ∇ in Anlehnung an diesen Operator; formal kann man $\Delta = \nabla \cdot \nabla$ schreiben.)

Auch die Rotation eines Gradientenfelds läßt sich leicht berechnen, allerdings natürlich nur im $\mathbb{R}^3$, da wir für andere Dimensionen keine

Rotation definiert haben.

$$\operatorname{rot} \operatorname{grad} f = \operatorname{rot} \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 f}{\partial z \partial y} - \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial x \partial z} - \frac{\partial^2 f}{\partial z \partial x} \\ \frac{\partial^2 f}{\partial y \partial x} - \frac{\partial^2 f}{\partial x \partial y} \end{pmatrix} = \vec{0} .$$

Die Rotation eines stetig differenzierbaren Gradientenfelds ist also stets Null; da das elektrische Feld der negative Gradient des elektrischen Potentials ist, hätten wir oben also für

$$\vec{E}(\boldsymbol{x}) = \frac{1}{4\pi\varepsilon_0} \frac{\vec{x}}{|\vec{x}|^3} = -\operatorname{grad} \frac{1}{4\pi\varepsilon_0 |\vec{x}|}$$

auf die Berechnung der Rotation verzichten können.

Die Divergenz eines Vektorfelds ist eine skalare Funktion; der einzige Operator, der darauf angewandt werden kann, ist daher der Gradient:

$$\operatorname{grad} \operatorname{div} \vec{V} = \operatorname{grad} \sum_{i=1}^{n} \frac{\partial V_i}{\partial x_i} = \begin{pmatrix} \sum_{i=1}^n \frac{\partial^2 V_i}{\partial x_1 \partial x_i} \\ \vdots \\ \sum_{i=1}^n \frac{\partial^2 V_i}{\partial x_n \partial x_i} \end{pmatrix} .$$

Die Rotation eines Vektorfelds im Dreidimensionalen schließlich ist wieder ein Vektorfeld, wir können also Divergenz und Rotation davon berechnen:

$$\begin{aligned} \operatorname{div} \operatorname{rot} \vec{V} &= \operatorname{div} \begin{pmatrix} \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} \\ \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \end{pmatrix} \\ &= \frac{\partial^2 V_3}{\partial x \partial y} - \frac{\partial^2 V_2}{\partial x \partial z} + \frac{\partial^2 V_1}{\partial y \partial z} - \frac{\partial^2 V_3}{\partial y \partial x} + \frac{\partial^2 V_2}{\partial z \partial x} - \frac{\partial^2 V_1}{\partial z \partial y} = 0 \end{aligned}$$

für ein zweimal stetig differenzierbares Vektorfeld. Da ein Magnetfeld als Rotation eines Vektorpotentials geschrieben werden kann, erklärt dies, warum im obigen Beispiel $\operatorname{div} \vec{H} = 0$ war.

Bleibt noch die Rotation der Rotation:

$$
\begin{aligned}
\operatorname{rot}\operatorname{rot}\vec{V} &= \operatorname{rot}\begin{pmatrix} \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} \\ \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial^2 V_2}{\partial y\partial x} - \frac{\partial^2 V_1}{\partial y^2} - \frac{\partial^2 V_1}{\partial z^2} + \frac{\partial^2 V_3}{\partial z\partial x} \\ \frac{\partial^2 V_3}{\partial z\partial y} - \frac{\partial^2 V_2}{\partial z^2} - \frac{\partial^2 V_2}{\partial x^2} + \frac{\partial^2 V_1}{\partial x\partial y} \\ \frac{\partial^2 V_1}{\partial x\partial z} - \frac{\partial^2 V_3}{\partial x^2} - \frac{\partial^2 V_3}{\partial y^2} + \frac{\partial^2 V_2}{\partial y\partial z} \end{pmatrix} \\
&= \operatorname{grad}\operatorname{div} f + \begin{pmatrix} \Delta\vec{V}_1 \\ \Delta\vec{V}_2 \\ \Delta\vec{V}_3 \end{pmatrix} .
\end{aligned}
$$

Als letztes wollen wir noch zwei „Produktregeln“ beweisen, die für quasilinearer Funktionen unmittelbare Folgerungen des oben bewiesenen Lemmas sind, für allgemeinere Funktionen aber neu.

Wir gehen aus von einer Funktion $\varphi\colon D \to \mathbb{R}$ und einem Vektorfeld $\vec{V}\colon D \to \mathbb{R}^n$ für eine offene Teilmenge D des $\mathbb{R}^n$; sowohl f als aauch $\vec{V}$ seien stetig differenzierbar. Dann ist

$$
\begin{aligned}
\operatorname{div}(f\vec{V}) &= \frac{\partial(fV_1)}{\partial x} + \frac{\partial(fV_2)}{\partial y} + \frac{\partial(fV_3)}{\partial z} \\
&= f\cdot\left(\frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} + \frac{\partial V_3}{\partial z}\right) + \left(\frac{\partial f}{\partial x}V_1 + \frac{\partial f}{\partial y}V_2 + \frac{\partial f}{\partial z}V_3\right) \\
&= f\operatorname{div}\vec{V} + (\operatorname{grad} f)\cdot\vec{V}
\end{aligned}
$$

und genauso im Falle $n = 3$

$$
\begin{aligned}
\operatorname{rot}(f\vec{V}) &= \begin{pmatrix} \frac{\partial(fV_3)}{\partial y} - \frac{\partial(fV_2)}{\partial z} \\ \frac{\partial(fV_1)}{\partial z} - \frac{\partial(fV_3)}{\partial x} \\ \frac{\partial(fV_2)}{\partial x} - \frac{\partial(fV_1)}{\partial y} \end{pmatrix} \\
&= f\begin{pmatrix} \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} \\ \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \end{pmatrix} + \begin{pmatrix} \frac{\partial f}{\partial y}V_3 - \frac{\partial f}{\partial z}V_2 \\ \frac{\partial f}{\partial z}V_1 - \frac{\partial f}{\partial x}V_3 \\ \frac{\partial f}{\partial x}V_2 - \frac{\partial f}{\partial y}V_1 \end{pmatrix} \\
&= f\operatorname{rot}\vec{V} + (\operatorname{grad} f)\times\vec{V} .
\end{aligned}
$$

Mit dem Operator ∇ geschrieben werden diese beiden Produktregeln völlig analog zur klassischen LEIBNIZ-Regel:

$$\nabla\cdot(f\vec{V}) = f\nabla\cdot\vec{V} + \nabla f\cdot\vec{V}$$

und

$$\nabla\times(f\vec{V}) = f\nabla\times\vec{V} + \nabla f\times\vec{V}$$

Speziell für ein lineares Vektorfeld

$$\vec{V}\colon \mathbb{R}^n \to \mathbb{R}^n; \quad \boldsymbol{x} \mapsto A\vec{x} \quad \text{mit} \quad A \in \mathbb{R}^{n\times n}$$

ist

$$\operatorname{div}\vec{V} = \sum_{i=1}^n \frac{\partial V_i}{\partial x_i} = \sum_{i=1}^n a_{ii} = \operatorname{Spur} A$$

und, im Falle $n = 3$,

$$\operatorname{rot}\vec{V} = \begin{pmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{pmatrix} ;$$

für eine quasilineare Funktion $\vec{f}(\boldsymbol{x}) = \varphi(\boldsymbol{x})A\vec{x}$ ist also nach den beiden gerade bewiesenen Produktregeln

$$\operatorname{div}\vec{f} = \varphi\operatorname{Spur} A + (\operatorname{grad}\varphi)\cdot A\vec{x}$$

und

$$\operatorname{rot}\vec{f} = \varphi\cdot\begin{pmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{pmatrix} + (\operatorname{grad}\varphi)\times A\vec{x} ,$$

was uns in den Beispielen des vorigen Abschnitts manche Rechnung erspart hätte.

g) Polarkoordinaten, Kugelkoordinaten und Zylinderkoordinaten

Auch in diesem letzten Abschnitt von §4 soll es um Möglichkeiten gehen, wie wir bei diesen (und anderen) Beispielen effizienter rechnen können. Das Feld einer Punktladung beispielsweise ist kugelsymmetrisch, jedoch haben wir bei der Berechnung von Divergenz und Rotation dieses Feldes die Kugelsymmetrie mutwillig zerstört, indem wir in einem kartesischen Koordinatensystem rechneten. Hier soll es nun um Koordinatensysteme gehen, die einer solchen Symmetrie angepaßt sind.

1) Polarkoordinaten in $\mathbb{R}^2$***:*** Aus der Analysis I oder aus der Schule ist wohl die Polarkoordinatendarstellung einer komplexen Zahl bekannt; genauso definieren wir für jeden Punkt $(x, y) \neq (0, 0)$ aus $\mathbb{R}^2$ die beiden Polarkoordinaten $r > 0, \varphi$ durch die Bedingung

$$x = r\cos\varphi \quad \text{und} \quad y = r\sin\varphi\,.$$

φ ist dabei, wie üblich, nur modulo 2π bestimmt.

Anstelle des Gitternetzes eines kartesischen Koordinatensystems hat man bei den Poalrkoordinaten ein Netz aus Kreisen um den Nullpunkt, auf denen der Radius r konstant ist, und aus vom Nullpunkt ausgehenden Strahlen, auf denen φ konstant ist.

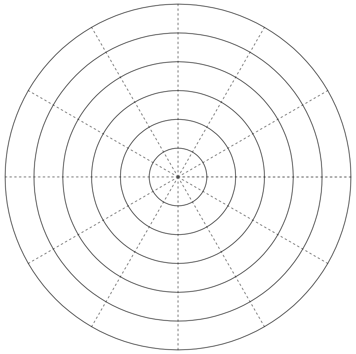


Abb. 50: Das Netz der Polarkoordinaten

Für eine Funktion $f(x, y)$ auf einer Teilmenge des $\mathbb{R}^2$ können wir die Polarkoordinatendarstellung

$$F(r, \varphi) \underset{\text{def}}{=} f(r\cos\varphi, r\sin\varphi)$$

betrachten; wir wollen sehen, wie man mit F allein immer noch den Gradienten oder beispielsweise auch die LAPLACEsche Δf berechnen kann.

Dazu müssen wir zunächst die partiellen Ableitungen von f nach x und y durch die von F nach r und φ ausdrücken. Dies ist auf direkte

Weise etwas schwierig, da wir keine gute Formel für die Berechnung von φ aus x und y haben. Die umgekehrte Aufgabe dagegen ist umso einfacher: Nach der Kettenregel ist

$$\frac{\partial F}{\partial r} = \frac{\partial}{\partial r} f(r\cos\varphi, r\sin\varphi) = \frac{\partial f}{\partial x}\cos\varphi + \frac{\partial f}{\partial y}\sin\varphi$$

und

$$\frac{\partial F}{\partial \varphi} = \frac{\partial}{\partial \varphi} f(r\cos\varphi, r\sin\varphi) = -\frac{\partial f}{\partial x} r\sin\varphi + \frac{\partial f}{\partial y} r\cos\varphi\,.$$

Da $r > 0$ ist, können wir dies auch in der Form

$$\begin{aligned} F_r &= \quad f_x\cos\varphi + f_y\sin\varphi \\ \frac{F_\varphi}{r} &= -f_x\sin\varphi + f_y\cos\varphi \end{aligned}$$

schreiben. Dies ist ein lineares Gleichungssystem für f_x und f_y; wenn wir von der mit $\cos\varphi$ multiplizierten ersten Gleichung die mit $\sin\varphi$ multiplizierte zweite subtrahieren, erhalten wir

$$f_x = F_r\cos\varphi - \frac{F_\varphi}{r}\sin\varphi\,;$$

entsprechend erhalten wir

$$f_y = F_r\sin\varphi + \frac{F_\varphi}{r}\cos\varphi\,,$$

wenn wir die mit $\sin\varphi$ multiplizierte erste Gleichung zur mit $\cos\varphi$ multiplizierten zweiten addieren.

Entsprechend können wir auch die zweiten partiellen Ableitungen von f durch partielle Ableitungen von F ausdrücken, allerdings werden hier die Ausdrücke schon deutlich komplexer, so daß hier auf die detailierte Rechnung verzichtet sei. Wichtig ist vor allem der LAPLACE-Operator, und wenn es auch sehr umständlich wäre, dessen Darstellung in Polarkoordinaten *herzuleiten* können wir die fertige Formel

$$\Delta f = f_{xx} + f_{yy} = F_{rr} + \frac{F_r}{r} + \frac{F_{\varphi\varphi}}{r^2}\,,$$

doch leicht und relativ schnell beweisen kann, indem wir

$$F_{rr} = \big(f_{xx}\cos\varphi + f_{xy}\sin\varphi\big)\cos\varphi + \big(f_{xy}\cos\varphi + f_{yy}\sin\varphi\big)\sin\varphi$$

und

$$F_{\varphi\varphi} = \big(f_{xx} r \sin\varphi - f_{xy} r \cos\varphi\big) r \sin\varphi - f_x r \cos\varphi \\ - \big(f_{xy} r \sin\varphi - f_{yy} r \cos\varphi\big) r \cos\varphi - f_y r \sin\varphi$$

nach der Kettenregel berechnen und einsetzen.

2) Zylinderkoordinaten im $\mathbb{R}^3$: Für den $\mathbb{R}^3$ gibt es zwei naheliegende Verallgemeinerung der Polarkoordinaten: Einmal können wir in der (x, y)-Ebene Polarkoordinaten einführen, die z-Koordinate aber unverändert lassen, zum anderen können wir neben dem Abstand vom Nullpunkt zwei Winkelkoordinaten einführen, wie z.B. die geographische Länge und Breite. Hier soll es zunächst um den ersten Fall gehen, wir haben also drei Koordinaten (r, φ, z) mit $r > 0$, die über die Gleichungen

$$x = r\cos\varphi, \quad y = r\sin\varphi \quad \text{und} \quad z = z$$

mit den kartesischen Koordinaten verbunden sind. Diese Koordinaten heißen *Zylinderkoordinaten,* da die Flächen $r =$ *konstant* Zylinder sind. Insbesondere ist die Koordinate r hier nicht mehr gleich dem Abstand eines Punktes vom Nullpunkt, sondern gleich dem Abstand von der z-Achse.

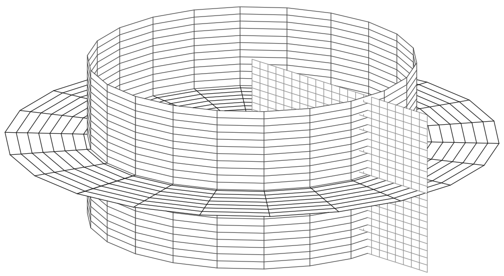


Abb. 51: Flächen mit konstantem r, φ und z für Zylinderkoordinaten

Da die z-Koordinate des kartesischen Koordinatensystems übernommen wird, brauchen wir für diese Koordinaten nicht neu zu rechnen, sondern

können die obigen Ergebnisse praktisch wörtlich übernehmen: Mit

$$F(r, \varphi, z) = f(r\cos\varphi, r\sin\varphi, z)$$

ist

$$f_x = F_r \cos\varphi - \frac{F_\varphi}{r}\sin\varphi$$
$$f_y = F_r \sin\varphi + \frac{F_\varphi}{r}\cos\varphi$$
$$f_z = F_z$$

und

$$\Delta f = f_{xx} + f_{yy} + f_{zz} = F_{rr} + \frac{F_r}{r} + \frac{F_{\varphi\varphi}}{r^2} + F_{zz}\,.$$

3) Kugelkoordinaten: Hier führen wir zusätzlich zu den beiden Polarkoordinaten der Ebenen als dritte Koordinate noch eine *Winkelkoordinate* ϑ ein, wir haben also drei Koordinaten (r, φ, ϑ), wobei r wieder positiv sein muß und den Abstand vom Nulpunkt bezeichnet.

Im Gegensatz zur Konvention der Geographen, wonach der Äquator, d.h. der Schnittkreis einer Kugel um den Nullpunkt mit der (x, y)-Ebenen, die Breite Null hat, ist es in der Mathematik, Physik und Technik üblich, dem Äquator die Koordinate $\vartheta = \pi/2$ zu geben, dem „Nordpol“, d.h. dem Schnittpunkt mit der positiven z-Achse, $\vartheta = 0$ und dem Südpol $\vartheta = \pi$. In der (x, z)-Ebene ist also

$$x = r\sin\vartheta \quad \text{und} \quad z = r\cos\vartheta\,,$$

in der (y, z)-Ebene

$$y = r\sin\vartheta \quad \text{und} \quad z = r\cos\vartheta\,,$$

und ganz allgemein ist

$$R = r\sin\vartheta \quad \text{und} \quad z = r\cos\vartheta\,,$$

wobei R den Abstand eines Punktes von der z-Achse bezeichnet.

Für einen Punkt im Abstand R von der z-Achse ist

$$x = R\cos\varphi \quad \text{und} \quad y = R\sin\varphi\,,$$

zusammen mit den obigen Formeln erhalten wir also als Zusammenhang zwischen Kugelkoordinaten und kartesischen Koordinaten

$$x = r\cos\varphi\sin\vartheta$$
$$y = r\sin\varphi\sin\vartheta$$
$$z = r\cos\vartheta\,.$$

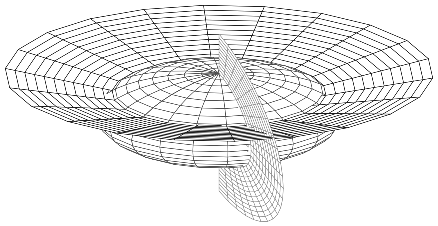


Abb. 52: Flächen mit konstantem r, φ und ϑ für Kugelkoordinaten

Die Flächen mit konstantem r sind hier natürlich Kugeln, konstantes φ führt wie bei den Zylinderkoordinaten auf Halbebenen, und die Flächen mit konstantem ϑ schließlich sind Kegel um die z-Achse.

Wieder können wir einer Funktion $f(x, y, z)$ in kartesischen Koordinaten durch

$$F(r, \varphi, \vartheta) \underset{\text{def}}{=} f(r\cos\varphi\sin\vartheta, r\sin\varphi\sin\vartheta, r\cos\vartheta)$$

eine neue Funktion zuordnen und durch Anwendung der Kettenregel sowie Lösen eines linearen Gleichungssystems die partiellen Ableitungen zueinander in Beziehung setzen; die Vorgehensweise ist im wesentlichen wie bei den ebenen Polarkoordinaten, nur daß wir jetzt ein System aus drei linearen Gleichungen bekommen:

$$\begin{aligned} F_r &= f_x\cos\varphi\sin\vartheta + f_y\sin\varphi\sin\vartheta + f_z\cos\vartheta \\ \frac{F_\varphi}{r} &= -f_x\sin\varphi\sin\vartheta + f_y\cos\varphi\sin\vartheta \\ \frac{F_\vartheta}{r} &= f_x\cos\varphi\cos\vartheta + f_y\sin\varphi\cos\vartheta - f_z\sin\vartheta\,. \end{aligned}$$

Der GAUSS-Algorithmus führt schnell auf die Lösung

$$\begin{aligned} f_x &= F_r\cos\varphi\sin\vartheta - \frac{F_\varphi\sin\varphi}{r\sin\vartheta} + \frac{F_\theta\cos\varphi\cos\vartheta}{r} \\ f_y &= F_r\sin\varphi\sin\vartheta - \frac{F_\varphi\cos\varphi}{r\sin\vartheta} + \frac{F_\theta\sin\varphi\cos\vartheta}{r} \\ f_z &= F_r cos\vartheta - \frac{F_\vartheta\sin\vartheta}{r}\,. \end{aligned}$$

Im Prinzip genauso können auch die zweiten Ableitungen berechnet werden; eine grausame, wenn auch nicht sonderlich schwierige Rechnung ergibt, daß beispielsweise

$$\begin{aligned} \Delta f &= f_{xx} + f_{yy} + f_{zz} \\ &= F_{rr} + \frac{2F_r}{r} + \frac{F_{\varphi\varphi}}{r^2\sin^2\vartheta} + \frac{F_{\vartheta\vartheta}}{r^2} + \frac{F_\vartheta}{r^2\tan\vartheta} \end{aligned}$$

ist.

Als Beispiel wollen wir damit noch einmal die Divergenz des elektrischen Felds einer Ladung q im Nullpunkt berechnen: Diese erzeugt ein Potential

$$U(r, \varphi, \vartheta) = \frac{q}{4\pi\varepsilon_0 r}\,,$$

das Feld ist $\vec{E} = -\operatorname{grad} U$ und

$$\operatorname{div}\vec{E} = -\operatorname{div}\operatorname{grad} U = -\Delta U = -\frac{q}{4\pi\varepsilon_0}\left(\frac{2}{r^3} + \frac{2}{r}\cdot\frac{-1}{r^2}\right) = 0\,.$$

§5: Kurvenintegrale im $\mathbb{R}^n$

Bislang hatten wir uns nur mit der *Differential*rechnung im $\mathbb{R}^n$ beschäftigt; es wird langsam Zeit, uns auch der *Integral*rechnung zuzuwenden. In diesem Paragraphen werden die Mengen, über die integriert wird, noch eindimensional sein, d.h. also Kurven. Entsprechende Integrale benötigt man einerseits, Bogenlängen von Kurven zu bereichen, vor allem aber sind sie wichtig, um die aufzuwendende oder freiwerdende Energie bei der Bewegung eines Objekts in einem Kraftfeld (oder einem elektrisch geladenen Teilchens in einem elektrischen Feld *usw.*) zu berechnen. Wir beginnen mit einer genaueren Klärung der Begriffe: