Vektoranalys

Olli Martio

Svensk översättning: Stefan Michelsson

Förord

Detta kompendium grundar sig på de föreläsningar i flera variablers differential- och integralkalkyl som författaren hållit vid Helsingfors och Jyväskylä universitet. Författaren har hållit motsvarande föreläsningar även utomlands. Innehållsförteckningen ger en uppfattning om det inkluderade materialet.

Analys i flera variabler är en fortsättning på analys i en variabel med realvärda funktioner, och grundar sig på behärskning av differential- och integralkalkyl för funktioner med en variabel. Kursen avvläggs vanligtvis under det andra studieåret. På grund av fysikaliska tillämpningar vore den användbar redan tidigare.

Kursen består vanligtvis av 4-5 föreläsningstimmar per vecka under ett läsår, vilket orsakar påfrestningar på att gallra materialet. Kompendiet är planerat så, att det är möjligt att lära sig det under en läsårslång kurs. En del av bevisen har förbigåtts. Av dessa kan nämnas

· satsen om inversa funktioner,
· flera satser angående integraler,
· Stokes sats.

För flera av satserna presenteras endast idén för beviset med ett specialfall. I dessa strävas det efter att poängtera så kallat “infinitesimalt” tänkande, som inte är väldigt viktigt i kursen om funktioner med en variabel, men vars betydelse i tillämpningar med funktioner med flera variabler är central.

Kompendiet har som fördel att använda en endimensionell situation som modell i en flerdimensionell situation. Därmed är det en naturlig fortsättning på första årets kurser. Det finns flera svagheter:

· kompendiet har inga övningsuppgifter, så föreläsaren måste göra dem själv eller hämta dem från övrig litteratur,
· det finns inga bilder, så studeranden måste rita dem själv,
· handledning i användning av datorgrafik har inte inkluderats p.g.a. de använda systemens snabba förändringstakt,

· några av definitionerna är vagt givna, exempelvis definitionen för en positivt orienterad rand till en öppen mängd,
· teori om kurvor och ytor behandlas endast ur en elementär synvinkel,
· kompendiet innehåller inga kommentarer om områdets historiska utveckling,
· endast få tillämpningar behandlas,
· Taylors formel förekommer endast som specialfall.

Analys i flera variabler behandlas i många både inhemska och utländska läroböcker. De flesta utländska läroböckerna är rätt ingående och innehåller rikligt med liknande övningsuppgifter. Böckerna [MT] och [PB] är utmärkta framställningar och betydligt mera omfattande än detta kompendium. I dem redogörs också flera tillämpningar matematiskt exakt.

Upplagorna för flera finskspråkiga böcker och kompendier har tagit slut, t.ex. [L], [Le] och [V], men dessa kan man ännu finna i biblioteken. De övriga är vanligtvis kompendier tryckta av högskolors tryckerier och motsvarar detta kompendium.

Författaren tackar student Ville Heino för att skriva detta kompendium i TEX-format. Eventuella (tryck)fel är egna, och har inte uppstått under renskrivningen.

Saate

Tämä luentomoniste koostuu kirjoittajan Helsingin ja Jyväskylän yliopistoissa pitämistä useamman muuttujan funktioiden differentiaali- ja integraalilaskennan luennoista. Kirjoittaja on pitänyt vastaavia luentoja myös ulkomailla. Sisällysluettelo antaa käsityksen mukaan otetusta materiaalista.

Useamman muuttujan analyysi jatkaa yhden muuttujan reaaliarvoisten funktioiden analyysia ja pohjautuu yhden muuttujan funktioiden differentiaali- ja integraalilaskennan hallintaan. Kurssi pyritään suorittamaan yleensä toisena opiskeluvuonna. Fysikaalisten sovellusten vuoksi se olisi hyödyllinen jo aikaisemmin.

Kurssi koostuu yleensä 4-5 viikottaisesta luentotunnista yhden lukukauden ajan, mikä aiheuttaa paineita materiaalin karsimiseen. Luentomoniste on mitoitettu niin, että sen omaksuminen on mahdollista lukukauden mittaisen kurssin aikana. Osa todistuksista on jouduttu sivuuttamaan. Näistä mainittakoon

· käänteiskuvauslause,
· useat integrointia koskevat lauseet,
· Stokesin lause.

Useista lauseista esitetään vain todistuksen idea erikoistapauksessa. Näissä pyritään painottamaan niin sanottua “infinidesimaalista” ajattelua, joka ei yhden muuttujan funktioiden kurssilla ole kovin tärkeää, mutta jonka merkitys useamman muuttujan funktioiden sovelluksissa on keskeinen.

Monisteen etuna on yksiulotteisen tilanteen käyttäminen mallina moniulotteisessa tilanteessa. Siten se jatkaa luontevasti 1. vuoden kursseja. Heikkouksia on useita:

· monisteessa ei ole harjoitustehtäviä, joten luennoitsija joutuu tekemään ne itse tai hakemaan muusta kirjallisuudesta,
· kuvia ei ole, joten opiskelija joutuu piirtämään ne itse,
· tietokonegrafiikan käytössä ei ole opastettu käytettävien järjestelmien tiheästä muuttuistahdista johtuen,
· eräät määritelmistä on annettu epämääräisesti, esimerkkinä avoimen joukon positiivisesti suunnistettu reuna,
· käyrä- ja pintateoria käsitellään ainoastaan alkeellisesta näkökulmasta,
· monisteessa ei ole kommentteja alan historiallisesta kehityksestä,
· vain harvoja sovelluksia käsitellään,
· Taylorin kaava esiintyy vain erikoistapauksessa.

Useamman muuttujan funktioiden analyysia käsitellään useassa sekä koti- että ulkomaisessa oppikirjassa. Useimmat ulkomaiset oppikirjat ovat varsin laajoja ja sisältävät runsaasti samankaltaisia harjoitustehtäviä. Kirjat [MT] ja [PB] ovat erinomaisia esityksiä ja selvästi laajempia kuin tämä moniste. Niissä myös selostetaan useita sovelluksia matemaattisen täsmällisesti.

Useista suomenkielisistä monisteista ja kirjoista ovat painokset loppuneet, esimerkkeinä [L], [Le] ja [V], mutta näitä löytää vielä kirjastoista. Muut ovat yleensä korkeakoulujen kustannustalojen painattamia luentomonisteita ja samankaltaisia kuin tämä.

Kirjoittaja kiittää ylioppilas Ville Heinoa tämän luentomonisteen TEX-kirjoittamisesta. Mahdolliset (paino)virheet ovat omia, eivät kirjoituksessa syntyneitä.

Referenser

[A] Adams, R. A., Calculus, 5th edition, Addison Wesley Longman, 2003.

[Ap] Apostol, T. M., Calculus, vol. II, John Wiley & Sons, Inc., 1969.

[B] Bers, L., Calculus, Holt, Rinehart and Winston, 1969.

[EP] Edwards, C. H. and Penney, D. E., Calculus, Prentice Hall, 1998.

[K] Kahanpää, Lauri, Matematiikan menetelmäkurssi, Matematiikan laitos, Jyväskylän yliopisto, luentomoniste 7, Jyväskylä, 1990.

[LHE] Larson, R., Hostetler, R. and Edwards, B., Calculus, D. C. Heath and Company, 1995.

[Le] Lehto, O., Differentiaali- ja integraalilaskenta II, Offset Oy, 1982.

[L] Lindelöf, E., Differentiaali- ja integraalilasku ja sen sovellukset, osa II, monistettu Helsingin yliopiston toimesta, 1959.

[MT] Marsden, J. E. and Tromba, A. J., Vector Calculus, 4th edition, W. H. Freeman and Company, 1996.

[Pa] Patovaara, T., Usean muuttujan funktiot, differentiaali- ja differenssiyhtälöt, Helsingin yliopisto, Tilastotieteen laitos, opetusmoniste.

[PB] Persson, A. and Böiers, L-C., Analys i flera variabler, Studentliteratur, 1988.

[Pu^1] Purmonen, V. T., Differentiaalilaskentaa euklidisissa avaruuksissa, Jyväskylän yliopisto, luentomoniste 22, Jyväskylä, 1993.

[Pu^2] Purmonen, V. T., Mitta- ja integraaliteoria, Matematiikan laitos, Jyväskylän yliopisto, luentomoniste 14, Jyväskylä, 1990.

[V] Väisälä, K., Vektorianalyysi, WSOY, 1961.

Innehåll

1. Det euklidiska rummet $\mathbb{R}^n$

1.1. $\mathbb{R}^n$:s struktur

Det euklidiska rummet $\mathbb{R}^n$, $n = 1, 2, \ldots$, definieras som mängden

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R},\ i = 1, \ldots, n\},$$

där $(x_1, x_2, \ldots, x_n)$ är en följd av reella tal med längden n. För punkterna, d.v.s. vektorerna i rummet $\mathbb{R}^n$ används beteckningen

$$x = \overline{x} = (x_1, x_2, \ldots, x_n)$$

var x_i kallas x *i:te koordinat*. Vanligtvis används inte ett streck på vektorn. Två vektorer i $\mathbb{R}^n$, $x = (x_1, x_2, \ldots, x_n)$ och $y = (y_1, y_2, \ldots, y_n)$, är lika om och endast om $x_i = y_i$ för varje $i = 1, \ldots, n$. Vi observerar att $\mathbb{R}^1$ kan jämställas med $\mathbb{R}$.

Rummet $\mathbb{R}^n$ är ett *reellt vektorrum*, alltså ett *lineärt rum*, då vi definierar summan av vektorerna x och y som

$$x + y = (x_1 + y_1, \ldots, x_n + y_n)$$

och multiplikation med det reella talet λ som

$$\lambda x = (\lambda x_1, \ldots, \lambda x_n);$$

nollelementet i $\mathbb{R}^n$ är nollvektorn, betecknas $0 = \overline{0} = (0, \ldots, 0)$. I $\mathbb{R}^n$ använder vi kunskaper från linjäralgebra.

Standardenhetsvektorerna i vektorrummet $\mathbb{R}^n$ är

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0), \\ e_2 &= (0, 1, 0, \ldots, 0), \\ &\ \ \vdots \\ e_n &= (0, \ldots, 0, 1). \end{aligned}$$

Varje vektor $x = (x_1, x_2, \ldots, x_n)$ kan på ett entydigt sätt presenteras i formen

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

Multiplikation med vektorer är vanligtvis inte definierad, förutom då $n = 1$, varvid den är vanlig multiplikation med reella tal. I rummet $\mathbb{R}^2$, alltså i planet, kan vi använda multiplikation med komplexa tal som vektormultiplikation. Detta gör vi ändå inte under denna kurs. I senare sammanhang använder vi i $\mathbb{R}^3$ kryssprodukten $x \times y$ för vektorerna $x, y \in \mathbb{R}^3$.

Den i rummet $\mathbb{R}^n$ definierade *skalära produkten* $x \cdot y$ får inte förväxlas med multiplikation. Den skalära produkten förenar vektorparet $x, y \in \mathbb{R}^n$ med ett reellt tal $x \cdot y$ med formeln

$$x \cdot y = x_1 y_1 + \ldots + x_n y_n.$$

Den skalära produkten uppfyller ekvationerna

$$x \cdot y = y \cdot x,$$
$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Längden av vektorn $x \in \mathbb{R}^n$ definieras med formeln

$$|x| = \sqrt{x \cdot x} = \sqrt{x_1^2 + \ldots + x_n^2} \geq 0$$

och *avståndet* mellan två vektorer x och y med formeln

$$d(x, y) = |x - y| = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}.$$

Vi observerar, att $|x| = 0$ exakt då $x = 0$, alltså då x är nollvektorn.

Vinkeln mellan vektorerna x och y $\theta \in [0, \pi]$ är den vinkel, som uppfyller ekvationen

$$\cos\theta = \begin{cases} \dfrac{x \cdot y}{|x||y|} & \text{, då } |x| \neq 0 \neq |y|, \\ 0 & \text{, då } |x| = 0 \text{ eller } |y| = 0. \end{cases} \tag{1.1.1}$$

1.1.2. Anmärkning. För att vinkeln θ skall existera måste det gälla att

$$\frac{x \cdot y}{|x||y|} \in [-1, 1].$$

Vi återkommer till detta senare.

Vi undersöker nu några egenskaper gällande skalär produkt, längd och avstånd.

A) Cauchy-Schwartz olikhet:

$$|x \cdot y| \leq |x||y| \tag{1.1.3}$$

d.v.s.

$$|x_1y_1 + \ldots + x_ny_n| \leq \sqrt{x_1^2 + \ldots + x_n^2}\sqrt{y_1^2 + \ldots + y_n^2}.$$

Bevis. Vi kan anta att $x \neq 0 \neq y$, eftersom formeln (1.1.1) annars är självklar. Låt $t \in \mathbb{R}$. Då gäller att

$$\begin{aligned} 0 \leq |tx + y|^2 &= (tx + y) \cdot (tx + y) \\ &= t^2 x \cdot x + tx \cdot y + ty \cdot x + y \cdot y \\ &= t^2|x|^2 + 2tx \cdot y + |y|^2 \end{aligned}$$

Nu är $t \mapsto t^2|x|^2 + 2tx \cdot y + |y|^2$ ett polynom av andra graden, som är större än eller lika med 0 för alla $t \in \mathbb{R}$. Då är dess graf ovanför t-axeln och polynomet kan ha högst ett nollställe. Därmed uppfyller dess diskriminant $D = b^2 - 4ac$ olikheten $D \leq 0$, och

$$D = b^2 - 4ac = 4(x \cdot y)^2 - 4|x|^2|y|^2 \leq 0.$$

Av detta följer att

$$(x \cdot y)^2 \leq |x|^2|y|^2,$$

vilket i sin tur är ekvivalent med olikheten

$$|x \cdot y| \leq |x||y|$$

□

1.1.4. Anmärkning. I det föregående beviset använde vi oss av att

$$\alpha^2 \leq \beta^2\gamma^2 \quad \Leftrightarrow \quad |\alpha|^2 \leq |\beta|^2|\gamma|^2 \quad \Leftrightarrow \quad |\alpha| \leq |\beta||\gamma|.$$

Av Cauchy-Schwartz olikhet följer att vinkeln θ i formeln (1.1.1) är väl definierad.

B) Följande egenskaper gäller för avstånd:

(i) $|x| \geq 0$ och $|x| = 0$ om och endast om $x = 0$,
(ii) $|\lambda x| = |\lambda||x|$ för varje $\lambda \in \mathbb{R}$,
(iii) triangelolikheten

$$|x + y| \leq |x| + |y|.$$

Bevis. Punkterna (i) och (ii) är självklara, punkt (iii) lämnar vi som övningsuppgift. □

C) Punkt B) (iii) har generaliseringarna

(I) $|x + y + \ldots + w| \leq |x| + |y| + \ldots + |w|$,
(II) triangelolikhetens vänstra sida

$$|x| - |y| \leq ||x| - |y|| \leq |x + y|.$$

Bevis. Punkt (I) är självklar. Den första olikheten i punkt (II) är också självklar, ty för varje reellt tal α gäller att $\alpha \leq |\alpha|$. Av olikheten

$$|x| = |(x + y) - y| \leq |x + y| + |y|$$

följer att $|x| - |y| \leq |x + y|$ och av olikheten

$$|y| = |(y + x) - x| \leq |x + y| + |x|$$

följer att $|y| - |x| \leq |x + y|$. Genom att kombinera dessa får vi att

$$||x| - |y|| \leq |x + y|,$$

vilket avslutar beviset. □

D) Vi upptäcker att det för varje $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ gäller att

$$|x_i| \leq |x| \leq |x_1| + |x_2| + \ldots + |x_n|.$$

Bevis. Det är ubbenbart att $|x_i| \leq \sqrt{x_1^2 + \ldots + x_n^2} = |x|$. För att bevisa högra sidan skriver vi x i formen

$$x = x_1 e_1 + \ldots + x_n e_n,$$

varvid punkt C) (I) och $|e_1| = 1,\ i = 1, \ldots, n$, ger att

$$\begin{aligned} |x| &= |x_1 e_1 + \ldots + x_n e_n| \leq |x_1 e_1| + \ldots + |x_n e_n| \\ &= |x_1||e_1| + \ldots + |x_n||e_n| = |x_1| + \ldots + |x_n|. \end{aligned}$$

□

1.2. Konvergens i $\mathbb{R}^n$

Antag att x_1, x_2, ... är en följd punkter i $\mathbb{R}^n$ (observera, att x_i inte är en koordinat till punkten x) alltså

$$x_i = (x_{i,1}, \ldots, x_{i,n}),\ i = 1, 2, \ldots .$$

Vi betecknar vektorföljden $x_1, x_2, \ldots$ lika som en följd reella tal, d.v.s. (x_i).

1.2.1. Exempel. $x_i = (1/i, 1/(i+1)) \in \mathbb{R}^2,\ i = 1, 2, \ldots$, definierar en följd (x_i) med punkter i $\mathbb{R}^2$.

Låt $x_0 \in \mathbb{R}^n$. Följden (x_i) *konvergerar mot punkten* x_0, om

$$|x_i - x_0| \to 0, \tag{1.2.2}$$

då $i \to \infty$. Märk, att det i formeln (1.2.2) är fråga om vanligt gränsvärde för reella tal. Med andra ord närmar sig avståndet mellan punkterna x_i och x_0 noll, då $i \to \infty$. Vi betecknar detta kortare som $x_i \to x_0,\ i \to \infty$, eller endast som $x_i \to x_0$.

1.2.3. Anmärkning. Då $n = 1$ är följande villkor ekvivalenta:

(i) $x_i \to x_0$,
(ii) $|x_i - x_0| \to 0$.

1.2.4. Sats. *$x_i \to x$, $i \to \infty$ gäller om och endast om $x_{i,j} \to x_j$ för varje $j = 1, \ldots, n$. Med andra ord konvergerar följden (x_i) mot punkten x om och endast om varje följd av koordinater $(x_{i,j})$, $j = 1, \ldots, n$, som bildas av punkterna x_i, konvergerar mot den motsvarande koordinaten x_j i punkten x.*

Bevis. $\boxed{\Rightarrow}$ Vi fixerar $j = 1, \ldots, n$ och visar att $x_{i,j} \to x_j$. Egenskapen d) ger att

$$|x_{i,j} - x_j| \le |x_i - x| \to 0,$$

varur påståendet följer.

$\boxed{\Leftarrow}$ Som känt gäller det för talföljder att det av villkoren $x_n \to 0$, $y_n \to 0$, ..., $z_n \to 0$ följer att $x_n + y_n + \ldots + z_n \to 0$. Vi antar att

$x_{i,j} \to x_j$ för varje j, varvid $|x_{i,j} - x_j| \to 0$, då $i \to \infty$. Eftersom detta gäller för varje $j = 1, \ldots, n$, så är

$$|x_{i,1} - x_1| + |x_{i,2} - x_2| + \ldots + |x_{i,n} - x_n| \to 0,$$

då $i \to \infty$. Därmed följer av egenskapen D) att

$$|x_i - x| \leq |x_{i,1} - x_1| + \ldots + |x_{i,n} - x_n| \to 0,$$

då $i \to \infty$. Påståendet följer av detta. □

Med stöd av Sats 1.2.4 kan konvergensen för punktföljden (x_i) i $\mathbb{R}^n$ reduceras till konvergens för reella tal. För talföljden i exemplet 1.2.1, $x_i = (1/i, 1/(i+1))$, $i = 1, 2, \ldots$, gäller att $x_i \to 0 = (0, 0)$, då $i \to \infty$, ty $1/i \to 0$ och $1/(i+1) \to 0$. Vi observerar att gränsvärdet för följden (x_i) är entydigt (p.g.a. Sats 1.2.4 och gränsvärdets entydighet i $\mathbb{R}$).

Punktföljden (x_i), $x_i \in \mathbb{R}$, *divergerar*, ifall den inte konvergerar. Punktföljden divergerar alltså exakt när åtminstone en av realtalsföljderna

$$(x_{i,j}), \ j = 1, \ldots, n,$$

divergerar.

Konvergensbegreppet fördjupas i kursen i topologi.

1.3. Mängder i $\mathbb{R}^n$

Ett av grundproblemen i rummet $\mathbb{R}^n$, $n \geq 2$, är att där inte finns en likadan "naturlig" ordning som på den reella axeln. Av detta följer att det i $\mathbb{R}^n$ inte finns en lika naturlig mängd som motsvaras av ett intervall i $\mathbb{R}$.

Med hjälp av avstånd är det möjligt att definiera några elementära mängder i $\mathbb{R}^n$. Om $x \in \mathbb{R}^n$ och $r > 0$, så kallas mängden

$$B(x, r) = B^n(x, r) = \{y \in \mathbb{R}^n : |y - x| < r\}$$

en *öppen kula* i $\mathbb{R}^n$ med medelpunkten x och radien r. Mängden $B(0, 1)$ kallas en öppen enhetskula. Då $n = 1$, så ger definitionen att $B(x, r)$

är ett öppet intervall $(x-r, x+r)$ och $B(0,1)$ är det öppna intervallet $(-1,1)$. Då $n=2$, så är $B(0,1) = \{y \in \mathbb{R}^2 : y_1^2 + y_2^2 < 1\}$.

Randen för kulan $B(x,r)$,

$$\begin{aligned}\partial B(x,r) = S(x,r) = S^{n-1}(x,r) &= \{y \in \mathbb{R}^n : |y-x| = r\} \\ &= \{y \in \mathbb{R}^n : (y_1-x_1)^2 + \ldots + (y_n-x_n)^2 = r^2\}\end{aligned}$$

kallas en ($n-1$ -dimensionell) boll i $\mathbb{R}^n$ med medelpunkten x och radien r. Mängden $S(x,r)$ kallas ofta även *sfär*.

För effektiv användning av mängderna i rummet $\mathbb{R}^n$ krävs några grundbegrepp i topologi.

Låt $A \subset \mathbb{R}^n$ vara en godtycklig mängd. Mängden A är *sluten*, om det för varje konvergent följd (x_i), där $x_i \in A$ och $x_i \to x_0$, gäller att $x_0 \in A$. Vi säger att A *innehåller alla sina anhopningspunkter.*

1.3.1. Exempel. En sluten kula i $\mathbb{R}^n$ definieras som mängden

$$\overline{B}(x,r) = \overline{B}^n(x,r) = \{y \in \mathbb{R}^n : |y-x| \leq r\}.$$

Detta är en sluten delmängd i $\mathbb{R}^n$. För att bevisa detta antar vi att (x_i) är en följd, där $x_i \in \overline{B}(x,r)$ och $x_i \to x_0$. Vi visar att $x_0 \in \overline{B}(x,r)$.

Eftersom $x_i \in \overline{B}(x,r)$, så gäller att $|x_i - x| \leq r$ och därmed är $|x_i - x|^2 \leq r^2$. Detta betyder att $\alpha_i = (x_{i,1}-x_1)^2 + \ldots + (x_{i,n}-x_n)^2 \leq r^2$. Eftersom $x_{i,j} \to x_{0,j}$ då $i \to \infty$ för varje $j = 1, 2, \ldots, n$, så har talföljden (α_i) gränsvärdet $\alpha_0 = (x_{0,1}-x_1)^2 + \ldots + (x_{0,n}-x_n)^2$ och $\alpha_0 \leq r^2$. Detta betyder att $|x_0 - x|^2 \leq r^2$ och därmed är $|x_0 - x| \leq r$.

1.3.2. Anmärkning. I räkningarna ovan använde vi principen om olikheters invarians, d.v.s. om $\alpha_i \to \alpha_0$ och $\alpha_i \leq r^2$, så är $\alpha_0 \leq r^2$

Mängden $A \subset \mathbb{R}^n$ definieras att vara *öppen*, om $\mathbb{R}^n \setminus A$ är sluten.

1.3.3. Sats. *Mängden $A \subset \mathbb{R}^n$ är öppen om och endast om det för varje $x \in A$ existerar ett sådant $r > 0$ att $B(x,r) \subset A$.*

Bevis. $\boxed{\Leftarrow}$ Vi skall visa att det av villkoret följer att $\mathbb{R}^n \setminus A$ är sluten. Antag att (x_i) är en följd punkter s.a. $x_i \in \mathbb{R}^n \setminus A$ och $x_i \to x_0$. Vi skall visa att $x_0 \in \mathbb{R}^n \backslash A$. Det finns två möjligheter: $x_0 \in \mathbb{R}^n \backslash A$, varvid påståendet är bevisat, eller $x_0 \in A$, som skall visas vara omöjligt.

Om $x_0 \in A$, så finns det enligt antagandet ett sådant $r > 0$ att $B(x_0, r) \subset A$. Å andra sidan är $x_i \in \mathbb{R}^n \backslash A$ och därmed gäller att $x_i \notin B(x_0, r)$. Av detta följer att $|x_i - x_0| \geq r$. Nu är $x_i \not\to x_0$, eftersom $x_i \to x_0$ betyder att $|x_i - x_0| \to 0$. Vi har därmed fått den sökta motsägelsen.

$\boxed{\Rightarrow}$ Övningsuppgift. $\square$

Låt $A \subset \mathbb{R}^n$ vara en mängd. Mängden A:s *rand* ∂A är mängden

$$\partial A = \{x \in \mathbb{R}^n : \forall r > 0 \ B(x, r) \cap A \neq \emptyset \text{ och } B(x, r) \cap (\mathbb{R}^n \backslash A) \neq \emptyset\}.$$

Randen för mängden A är allstå den mängd av punkter, vars varje kulomgivning möter både A och A:s komplement. Om A är öppen, så är punkterna i ∂A aldrig punkter i A. Detta följer av Sats 1.3.3. Å andra sidan, om A är sluten, så är $\partial A \subset A$ (övningsuppgift).

1.3.4. Exempel. Antag att A är en *enpunktsmängd*, alltså $A = \{x_0\}$. Då är $A = \partial A = \{x_0\}$. Punkten x_0 tillhör ∂A, ty för varje $r > 0$ innehåller kulan $B(x_0, r)$ punkter i A (punkten x_0), och uppenbart även punkter i $\mathbb{R}^n \backslash A$.

Om $x \neq x_0$, så gäller att $x \notin \partial A$, ty om $r = |x - x_0| > 0$, så innehåller kulan $B(x, r)$ inte x_0 och därmed inte heller punkter i A. Därmed är x_0 den enda punkten i ∂A och det gäller alltså att $\partial A = \{x_0\}$.

1.3.5. Sats. *Låt $A \subset \mathbb{R}^n$. Då är*

(i) *∂A sluten och*
(ii) *$A \cup \partial A$ sluten.*

Bevis. Övningsuppgift. $\square$

Mängden $A \cup \partial A$ kallas det *slutna höljet* av A och betecknas $\overline{A}$. Med stöd av Sats 1.3.5 är det en sluten mängd. I själva verket är det den minsta slutna mängden, som innehåller A.

1.3.6. Exempel. Låt $B(x_0, r)$ vara en öppen kula i $\mathbb{R}^n$. Vi upptäcker, att $\partial B(x_0, r) = S(x_0, r) = \{y : |y - x_0| = r\}$. Med stöd av det slutna höljets definition är

$$\overline{B(x_0, r)} = B(x_0, r) \cup S(x_0, r) = \{y : |y - x_0| \leq r\}.$$

Därmed är $\overline{B}(x_0, r) = \overline{B(x_0, r)}$ verkligen en "sluten kula".

1.3.7. Anmärkning. I fallet $n = 1$ gäller att

$$\begin{aligned} B(x_0, r) &= (x_0 - r, x_0 + r), \\ S(x_0, r) &= \{x_0 - r, x_0 + r\}, \\ \overline{B}(x_0, r) &= [x_0 - r, x_0 + r]. \end{aligned}$$

Ett av topologins viktigaste begrepp är kompakthet. I det euklidiska rummet $\mathbb{R}^n$ kan detta definieras på följande sätt: Mängden $A \subset \mathbb{R}^n$ är *kompakt*, ifall det för varje följd (x_i), $x_i \in A$, finns en sådan delföljd (x_{i_j}), att $x_{i_j} \to x_0$ och $x_0 \in A$. Med andra ord kan vi för varje följd i A välja en delföljd, som konvergerar mot en punkt i A.

1.3.8. Anmärkning. I fallet $n = 1$ gäller: Om $x_i \in [a, b]$, så existerar det en delföljd $x_{i_j} \to x_0$ och $x_0 \in [a, b]$. Ett slutet intervall $[a, b]$ är alltså en kompakt delmängd till den reella axeln $\mathbb{R} = \mathbb{R}^1$.

Vi säger att mängden $A \subset \mathbb{R}^n$ är *begränsad*, om $A \subset B(0, r)$ för något $r > 0$.

1.3.9. Sats. *Låt $A \subset \mathbb{R}^n$. Då är A kompakt om och endast om A är sluten och begränsad.*

Bevis. $\boxed{\Leftarrow}$ Låt (x_i) vara en följd av punkter i A. Vi måste först visa, att denna följd har en konvergerande delföljd. Varje följd av reella tal $(x_{i,j})$, $j = 1, 2, \ldots, n$, är begränsad ($|x_{i,j}| \leq |x_i| < r$). Vi väljer först en konvergent delföljd $(x_{i_j,1})$ till följden $(x_{i,1})$ och sedan en konvergent delföljd till följden $(x_{i_j,2})$. Vi fortsätter på detta sätt och väljer slutligen en konvergent delföljd $(x_{k_i,n})$ till följden $(x_{i,n})$. Nu konvergerar alla följder $(x_{k_i,j})$, $j = 1, 2, \ldots, n$, (eftersom varje delföljd till en konvergent följd konvergerar), d.v.s. vi kan finna sådana tal $x_j \in \mathbb{R}$, att

$$x_{k_i,j} \to x_j, \ i \to \infty, \ j = 1, 2, \ldots, n.$$

Vi betecknar $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Då gäller att $x_{k_i} \to x$ då $i \to \infty$ (ty $x_i \to x$ betyder att varje koordinat i x_i konvergerar mot en motsvarande koordinat i x, jämför med Sats 1.2.4).

Vi har visat, att följden (x_i) har en delföljd (x_{k_i}) som konvergerar mot x. Vi skall ännu bevisa, att $x \in A$. Av påståendet följer att A är sluten, d.v.s. gränsvärdet x för följden (x_{k_i}) tillhör A.

$\boxed{\Rightarrow}$ Övningsuppgift. $\square$

1.3.10. Exempel. Den slutna kulan $\overline{B}(x_0, r)$ är kompakt, ty den är sluten och begränsad.

Öppna, slutna och kompakta mängder behandlas närmare i kursen i topologi.

2. Realvärda funktioner i $\mathbb{R}^n$

2.1. Funktioner i $\mathbb{R}^n$

Låt $A \subset \mathbb{R}^n$ vara en mängd och $u : A \to \mathbb{R}$ vara en *realvärd funktion* alltså en *avbildning* i mängden A, med andra ord är ett entydigt reellt tal $u(x) \in \mathbb{R}$ förenat med varje $x \in A$. Ofta skriver vi $u(x) = u(x_1, \ldots, x_n)$, alltså $x = (x_1, \ldots, x_n)$. Visualisering av funktionerna, alltså den grafiska representationen, är svårare i $\mathbb{R}^n$, $n \geq 2$, än på den reella axeln $\mathbb{R}$. En graf av en funktion i rummet $\mathbb{R}^n$ är alltid en delmängd i rummet $\mathbb{R}^{n+1}$. I det endimensionella fallet är alltså grafen av en realvärd funktion i planet, varvid den är enkel att gestalta.

2.1.1. Exempel. Antag att $u : \mathbb{R}^n \to \mathbb{R}$, $u(x) = |x|$, är avståndet från punkten punkten x till origo. Då $n = 2$ är grafen av funktionen u en delmängd av $\mathbb{R}^3$

$$\{(x_1, x_2, u(x_1, x_2) \in \mathbb{R}^3 : \ (x_1, x_2) \in \mathbb{R}^2\}.$$

Detta är en kon i $\mathbb{R}^3$, vars spets är i origo.

2.1.2. Exempel. Antag att $u : \mathbb{R}^n \to \mathbb{R}$, $u(x) = u(x_1, \ldots, x_n)$, är den första decimalen i talet x_1 Då är u inte en funktion, ty om $x_1 = 1,00\ldots = 0.99\ldots$, så är $u(x) = 0$ eller $u(x) = 9$, och därmed är u:s värde i punkten x inte entydigt bestämt.

2.1.3. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^2$ samt $c \in \mathbb{R}$. Mängden

$$C = C_c = \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}$$

kallas en *nivåkurva* för f. Mängden C behöver inte vara en "kurva"; för en allmän avbildning f kan den vara en godtycklig mängd i $\mathbb{R}^2$. I detta fall är C lätt att få reda på:

$$f(x, y) = x^2 - y^2 = c \quad \Leftrightarrow \quad (x - y)(x + y) = c.$$

Om $c = 0$, så följer att $x - y = 0$ eller $x + y = 0$, vilket betyder att C_0 består av linjen $x = y$ eller $x = -y$. Om $c \neq 0$, så består C_c av två hyperboliska kurvor $x^2 - y^2 = c$.

Märk, att vi i exemplet 2.1.3 använde beteckningen (x, y) istället för beteckningen (x_1, x_2). Detta är ett vanligt beteckningssätt i $\mathbb{R}^2$. På motsvarande sätt är (x, y, z) det vanliga beteckningssättet för punkter i $\mathbb{R}^3$.

2.1.4. Exempel. Låt $f : \mathbb{R}^3 \to \mathbb{R}$. Nu är det omöjligt att rita grafen av f, eftersom denna är en delmängd i $\mathbb{R}^4$

$$\{(x, y, z, f(x, y, z)) \in \mathbb{R}^4 : (x, y, z) \in \mathbb{R}^3\}.$$

Låt $f(x, y, z) = x^2 - z$. Vi kan nu skissera *nivåytorna* av f, alltså mängderna

$$\begin{aligned} C_c &= \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = c\} \\ &= \{(x, y, z) \in \mathbb{R}^3 : x^2 - z = c\} \\ &= \{(x, y, z) \in \mathbb{R}^3 : z = x^2 - c\}. \end{aligned}$$

Mängden C_c är en parabolisk cylinder.

2.2. GRÄNSVÄRDEN OCH KONTINUITET

Antag att $A \subset \mathbb{R}^n$ och $f : A \to \mathbb{R}$ samt att $x_0 \in \mathbb{R}^n$ och $a \in \mathbb{R}$. Funktionen f har gränsvärdet a i punkten x_0, om det för varje följd (x_i), för vilken $x_i \in A$, $x_i \to x_0$ och $x_i \neq x_0$, gäller att

$$f(x_i) \to a. \tag{2.2.1}$$

Vi betecknar $\lim_{x \to x_0} f(x) = a$. Ibland används även beteckningen

$$\lim_{\substack{x \to x_0 \\ x \in A \setminus \{x_0\}}} f(x) = a.$$

Följden $(f(x_i))$ är en följd av reella tal och i formeln (2.2.1) förekommer vanligt gränsvärde för följder av reella tal.

2.2.2. Anmärkning. Definitionen är meningsfull endast då $x_0 \in \overline{A}$, ty om $x_0 \notin \overline{A}$, så är $x_0 \in \mathbb{R}^n \setminus \overline{A}$. Mängden $\mathbb{R}^n \setminus \overline{A}$ är öppen, så det finns ett $r > 0$ s.a. $B(x_0, r) \subset \mathbb{R}^n \setminus \overline{A}$ och eftersom $\overline{A} = A \cup \partial A$, så gäller $\overline{A} \supset A$ och därmed är $B(x_0, r) \subset \mathbb{R}^n \setminus A$. Därmed existerar det ingen följd (x_i), för vilken $x_i \in A$ och $x_i \to x_0$. Detsamma gäller om x_0 är en enskild punkt för A, m.a.o. för något $r > 0$, $B(x_0, r) \cap A = \{x_0\}$.

2.2.3. Exempel. Låt $f : \mathbb{R}^2\backslash\{0\} \to \mathbb{R}$,

$$f(x, y) = \frac{2xy}{x^2 + y^2}.$$

Vi undersöker, ifall gränsvärdet $\lim_{x\to 0} f(x)$ existerar. Frågan är meningsfull, eftersom $0 \in \overline{\mathbb{R}^2\backslash\{0\}} = \mathbb{R}^2$. Vi undersöker gränsvärdet genom att se hur f beter sig i punkterna $f(x_i)$ då följden (x_i) väljs på ett lämpligt sätt. Vi väljer först

$$x_i = (1/i, 0) \in \mathbb{R}^2\backslash\{0\}.$$

Nu gäller att $x_i \to 0 = (0, 0)$, eftersom

$$|x_i - 0| = \sqrt{(1/i - 0)^2 + (0 - 0)^2} = 1/i \to 0.$$

Å andra sidan är

$$f(x_i) = f(1/i, 0) = \frac{2 \cdot 1/i \cdot 0}{1/i^2 + 0^2} = 0,$$

d.v.s. $f(x_i) \to 0$. Vi väljer härnäst

$$x_i = (1/i, 1/i) \in \mathbb{R}^2\backslash\{0\}.$$

Nu är även $x_i \to 0$ och

$$f(x_i) = f(1/i, 1/i) = \frac{2 \cdot 1/i \cdot 1/i}{1/i^2 + 1/i^2} = \frac{2}{1 + 1} = 1.$$

Därmed uppfyller $(f(x_i))$ som en konstant följd $f(x_i) \to 1$. Av detta kan vi dra slutsatsen, att f inte har ett gränsvärde i origo, eftersom gränsvärdet annars skulle vara samma för varje följd.

2.2.4. Anmärkning. Gränsvärdets definition är detsamma som i fallet $n = 1$. Den enda skillnaden är att f nu kan vara definierad i en godtycklig mängd A i $\mathbb{R}$. Märk, att vi med denna definition samtidigt behandlar de ensidiga gränsvärdena i intervallets ändpunkter.

2.2.5. Exempel. Låt $f : \mathbb{R}^2\backslash\{0\} \to \mathbb{R}$,

$$f(x, y) = \frac{x^2 y^2}{x^2 + y^2}.$$

Vi visar, att $\lim_{x\to 0} f(x) = 0$. Antag att $(x_i, y_i) \to (0, 0)$, $(x_i, y_i) \in \mathbb{R}^2\backslash\{0\}$. Vi skall visa att $f(x_i, y_i) \to 0$.

Bevis 1.

$$|f(x_i, y_i) - 0| = f(x_i, y_i) = \frac{x_i^2 y_i^2}{x_i^2 + y_i^2}$$

Låt $\varepsilon > 0$. Om $x_i \to 0$, så är $x_i^2 \to 0$. Därmed existerar det ett sådant i_ε, att $x_i^2 < \varepsilon$ då $i > i_\varepsilon$. Av detta följer att

$$|f(x_i, y_i) - 0| = \frac{x_i^2 y_i^2}{x_i^2 + y_i^2} \le \frac{\varepsilon\, y_i^2}{y_i^2} = \varepsilon, \text{ då } i > i_\varepsilon,$$

varvid påståendet gäller. □

I det förgående beviset finns ett fel. Det sista estimatet är ogiltigt, eftersom y_i i detta fall kan vara 0. Detta kan dock korrigeras på följande sätt.

Bevis 2. Eftersom $\alpha^2 + \beta^2 - 2\alpha\beta = (\alpha - \beta)^2 \ge 0$, så är $2\alpha\beta \le \alpha^2 + \beta^2$. Därmed gäller $2x_i y_i \le x_i^2 + y_i^2$. Av detta följer att

$$\frac{x_i^2 y_i^2}{x_i^2 + y_i^2} = \frac{2x_i y_i \cdot x_i y_i}{2(x_i^2 + y_i^2)} \le \frac{x_i y_i}{2} \to 0, \text{ då } i \to \infty,$$

ty om $x_i \to 0,\ y_i \to 0$, så gäller att $x_i y_i \to 0$. □

Vi undersöker härnäst kontinuitetsbegreppet i rummet $\mathbb{R}^n$. Antag att $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}$ och att $x_0 \in A$. Avbildningen f är *kontinuerlig i punkten* x_0, om

$$\lim_{x \to x_0} f(x) = f(x_0).$$

2.2.6. Anmärkning. Akademiskt fall: Om x_0 är en enskild punkt i f:s definitionsmängd A, så är f automatiskt kontinuerlig i punkten x_0.

För kontinuitet får vi en ekvivalent karakterisering även med hjälp av punktföljder. Avbildningen f är kontinuerlig i punkten $x_0 \in A$ exakt då, när det för alla följder (x_i), för vilka $x_i \in A$ och $x_i \to x_0$ gäller att $f(x_i) \to f(x_0)$. Märk, att x_i här även får vara x_0, men det räcker att granska de följder (x_i), för vilka det gäller att $x_i \neq x_0$ för varje i.

2.2.7. Exempel. Vi undersöker funktionen i exempel 2.2.5

$$f : \mathbb{R}^2 \backslash \{0\} \to \mathbb{R},\ f(x, y) = \frac{x^2 y^2}{x^2 + y^2}.$$

Det går inte att utreda ifall funktionen är kontinuerlig i origo, eftersom avbildningen f inte är definierad i origo. Att undersöka kontinuiteten i punkterna $(x,y) \neq (0,0)$ är meningsfullt.

Antag att $(x_i, y_i) \to (x,y)$, $(x_i,y_i) \neq (0,0)$. Eftersom $x_i \to x$ och $y_i \to y$, så är $x_i^2 \to x^2$ och $y_i^2 \to y^2$. Därmed gäller $x_i^2 y_i^2 \to x^2 y^2$. Likaså gäller att $x_i^2 + y_i^2 \to x^2 + y^2 \neq 0$. Av detta följer att

$$\frac{x_i^2 y_i^2}{x_i^2 + y_i^2} \to \frac{x^2 y^2}{x^2 + y^2},$$

och därmed är

$$|f(x_i,y_i) - f(x,y)| = \left| \frac{x_i^2 y_i^2}{x_i^2 + y_i^2} - \frac{x^2 y^2}{x^2 + y^2} \right| \to 0,$$

varvid f är kontinuerlig i punkterna $(x,y) \neq (0,0)$.

Vi definierar $f(0,0) = 0$. Nu är f kontinuerlig även i origo, eftersom det enligt exempel 2.2.5 gäller att

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0 = f(0,0).$$

Om vi definierar $f(0,0) = a \neq 0$, så är den på så sätt erhållna funktionen inte kontinuerlig i origo.

2.2.8. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = ax^n y^n + bx^n y^{n-1} + \ldots + cx + dy + \alpha$. En sådan funktion kallas ett *polynom* i $\mathbb{R}^2$. Denna är uppenbart kontinuerlig i varje punkt i $\mathbb{R}^2$.

Funktionen $f : A \to \mathbb{R}$ sägs vara *kontinuerlig i mängden* A, om f är kontinuerlig i varje punkt i A.

Gränsvärden och kontinuitet behandlas närmare i kursen i topologi.

2.3. Partiella derivator

Antag att $D \subset \mathbb{R}^n$ är öppen och att $f : D \to \mathbb{R}$. Vi fixerar en punkt $x_0 \in D$. Eftersom D är öppen, så existerar det ett sådant $r > 0$, att $B(x_0, r) \subset D$. Därmed är f definierad i punkterna $x_0 + y$ då $|y| < r$. Vi fixerar $j = 1, 2, \ldots, n$. Nu är funktionen $h \mapsto f(x_0 + he_j)$ definierad åtminstone då $h \in (-r, r)$. Om gränsvärdet

$$\begin{aligned}&\lim_{h\to 0} \frac{f(x_0 + he_j) - f(x_0)}{h}\\ &= \lim_{h\to 0} \frac{f(x_{0,1}, x_{0,2}, \ldots, x_{0,j} + h, \ldots, x_{0,n}) - f(x_0)}{h}\end{aligned}$$

existerar, så kallas det den *j:te partiella derivatan av funktionen f* i punkten x_0 och betecknas

$$\partial_j f(x_0) = \frac{\partial f}{\partial x_j}(x_0) = \mathrm{D}_j f(x_0) = \mathrm{D}_j f(x_{0,1}, \ldots, x_{0,n}).$$

Av definitionen följer att $\partial_j f(x_0) \in \mathbb{R}$, ifall det existerar.

Den partiella derivatans geometriska betydelse är enkel att tolka. För enkelhetens skull granskar vi fallet i planet, alltså då $n = 2$, och den partiella derivatan $\partial_1 f(x_0)$. Funktionen f begränsas till snittet av linjen

$$L = \{x \in \mathbb{R}^2 : x = x_0 + te_1, t \in \mathbb{R}\},$$

och mängden D, med andra ord undersöker vi den realvärda funktionen

$$s \overset{g}{\mapsto} f(s, x_{0,2}).$$

Denna är definierad åtminstone i något intervall $(x_{0,1} - r, x_{0,1} + r)$. Vi bestämmer denna funktions derivata, ifall den existerar, i punkten $x_{0,1}$. Vi erhåller att

$$\begin{aligned}g'(x_{0,1}) &= \lim_{h\to 0} \frac{g(x_{0,1} + h) - g(x_{0,1})}{h}\\ &= \lim_{h\to 0} \frac{f(x_{0,1} + h, x_{0,2}) - f(x_{0,1}, x_{0,2})}{h}\\ &= \lim_{h\to 0} \frac{f(x_0 + he_1) - f(x_0)}{h}\\ &= \partial_1 f(x_0).\end{aligned}$$

Att beräkna den partiella derivatan reduceras alltså till att räkna alldeles vanlig derivata.

Beteckningarna ovan är onödigt komplicerade. Genom att övergå till att använda beteckningen (x, y) för punkter i planet $\mathbb{R}^2$, granskar vi funktionen

$$(x, y) \mapsto f(x, y)$$

i närheten av punkten $(x_0, y_0) \in D$. Vi väljer såsom ovan $g(s) = f(s, y_0)$, men vi använder x istället för s, $g(x) = f(x, y_0)$. Nu är $\partial_1 f(x_0, y_0) = g'(x_0)$ och på motsvarande sätt är $\partial_2 f(x_0, y_0) = h'(y_0)$, där $h(y) = f(x_0, y)$.

Vi upptäcker alltså, att beräkning av den partiella derivatan i punkten

$$x_0 = (x_{0,1}, x_{0,2}, \ldots, x_{0,n})$$

med avseende på variabeln x_j, $j = 1, \ldots, n$, reduceras till det nedanstående. Vi undersöker den i intervallet $(x_{0,j} - r, x_{0,j} + r)$ definierade funktionen

$$s \mapsto f(x_{0,1}, \ldots, x_{0,j-1}, s, x_{0,j+1}, \ldots, x_{0,n})$$

och beräknar dess derivata i punkten $x_{0,j}$.

2.3.1. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$. Vi beräknar $\partial_1 f(x_0, y_0)$. Vi skall derivera funktionen $s \mapsto f(s, y_0)$ i punkten (x_0, y_0). Denna funktion är $g(s) = sy_0$. Nu är $g'(s) = y_0$ och därmed är $\partial_1 f(x_0, y_0) = y_0$. Detta kan göras enklare genom att ändra beteckningarna. För att bestämma den partiella derivatan $\partial_1 f(x_0, y_0)$ skall vi derivera $x \mapsto f(x, y) = xy$. Derivering med avseende på x ger att $\partial_1 f(x, y) = y$. På motsvarande sätt får vi att $\partial_2 f(x, y) = x$, ty derivatan av funktionen $s \mapsto xs$ i punkten y är x.

2.3.2. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = |x|$. Vad är $\partial_2 f(x, y)$? För att bestämma detta skall vi derivera funktionen $y \mapsto f(x, y) = |x|$, då x är fixerad. Detta är en konstant funktion. Alltså är $\partial_2 f(x, y) = 0$ för varje $(x, y) \in \mathbb{R}^2$.

Vi bestämmer härnäst $\partial_1 f(x, y)$. Vi skall alltså derivera funktionen $x \mapsto f(x, y) = |x|$, då y är fixerad.

$$\partial_1 f(x, y) = \begin{cases} \dfrac{x}{|x|}, & \text{då } x \neq 0, \\ \text{ej definierad,} & \text{då } x = 0. \end{cases}$$

Därmed existerar inte den partiella derivatan $\partial_1 f(0, y)$.

2.3.3. Exempel. Antag att

$$f(x, y) = \frac{1}{x} + \cos(xy) + e^{\sin(x+y)}.$$

Detta uttryck definierar funktionen $f : \underbrace{\mathbb{R}^2 \setminus \{(x,y) \in \mathbb{R}^2 : x = 0\}}_{\text{öppen mängd}} \to \mathbb{R}$.

Vi skall bestämma $\partial_2 f(x,y)$ då $x \neq 0$, med andra ord utanför y-axeln. För att göra detta fixerar vi $x \neq 0$ och deriverar funktionen

$$y \mapsto \frac{1}{x} + \cos(xy) + e^{\sin(x+y)}.$$

Dennas derivata är

$$\begin{aligned}
&0 - \sin(xy) \cdot x + e^{\sin(x+y)} \cos(x+y) \cdot 1 \\
=& - x \sin(xy) + \cos(x+y) e^{\sin(x+y)} \\
=& \partial_2 f(x,y).
\end{aligned}$$

I högre dimensioner räknar vi partiella derivator på helt motsvarande sätt.

2.3.4. Exempel. Låt $f : \mathbb{R}^4 \to \mathbb{R}$, $f(x) = x_1 x_2 + x_3 e^{x_4}$. Vi skall bestämma $\partial_3 f(x)$. Vi fixerar x_1, x_2, x_4 och deriverar funktionen

$$x_3 \mapsto x_1 x_2 + x_3 e^{x_4}.$$

Dennas derivata är $e^{x_4} = \partial_3 f(x)$.

2.3.5. Anmärkning. I tillämpningar är det vanligt att f inte är definierad i en öppen mängd i $\mathbb{R}^n$ (i $\mathbb{R}^n$ förekommer derivatan även i intervallets ändpunkter). Till exempel kan funktionen vara definierad i en "sluten" kvadrat i $\mathbb{R}^2$. Då är det meningsfullt att undersöka partiella derivator även på kvadratens kanter. Definitionen påminner om ensidiga derivator i ändpunkter.

2.4. DERIVERBARHET OCH TANGENTPLAN

I fallet $n = 1$ gäller att $\partial_1 f(x) = f'(x)$. Om f har en derivata $f'(x)$ i punkten x, så är f kontinuerlig i punkten x.

2.4.1. Exempel. Antag att $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} 0 & , \text{ då } x = 0 \text{ eller } y = 0, \\ 1 & , \text{ annars} \end{cases}$$

Helt klart är $\partial_1 f(0,0) = 0 = \partial_2 f(0,0)$, ty f får värdet 0 i korset som bildas av koordinataxlarna, men f är inte kontinuerlig i origo, eftersom den inte har något gränsvärde i origo.

Det föregående exemplet visar, att de partiella derivatornas existens i fallet $n \geq 2$ inte implicerar funktionens kontinuitet i de ifrågavarande punkterna. Situationen går dock att korrigera.

Vi behöver ett begrepp så att funktionen f kan, nära en given punkt, approximeras med en affin avbildning. Ofta uttrycks detta med att grafen av funktionen f kan, nära en given punkt x_0, approximeras med sitt *tangentplan*. Detta motsvarar approximationsformeln i fallet $n = 1$,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + |x - x_0|\varepsilon(x - x_0),$$

där $\varepsilon(x - x_0) \to 0$ då $x \to x_0$. Uttrycket ovan skrivs ofta i formen

$$f(x_0 + h) = f(x_0) + f'(x_0)h + |h|\varepsilon(h),\ h \to 0.$$

Grafen av avbildningen $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ är en linje, som går genom punkten $(x_0, f(x_0))$ och vars riktningskoefficient är $f'(x_0)$. Detta är f:s tangentplan i punkten $(x_0, f(x_0))$. Approximation av funktionen f med sin tangent betyder alltså att approximera f med den affina avbildningen

$$x \mapsto f(x_0) + f'(x_0)(x - x_0).$$

Denna företeelse har en naturlig motsvarighet i $\mathbb{R}^n$. Linjen i $\mathbb{R}^2$ motsvarar nu ett n-plan i $\mathbb{R}^{n+1}$. Ett sådant plan (som inte står vinkelrätt mot planet $x_{n+1} = 0$) är alltid grafen

$$G = \{(x_1, \ldots, x_n, a_0 + a_1 x_1 + \ldots + a_n x_n) \in \mathbb{R}^{n+1} : (x_1, \ldots, x_n) \in \mathbb{R}^n\}$$

av avbildningen $\mathbb{R}^n \to \mathbb{R}$,

$$(x_1, \ldots, x_n) \mapsto a_0 + a_1 x_1 + \ldots + a_n x_n. \tag{2.4.2}$$

Ett behändigt sätt att framföra ett plan är att använda den skalära produkten av vektorn $a = (a_1, \ldots, a_n)$ och $\mathbb{R}^n$. Nu får avbildningen (2.4.2) formen

$$x \mapsto a_0 + a \cdot x.$$

En sådan avbildning kallas *affin*. Dess graf är ett n-plan i rummet $\mathbb{R}^{n+1}$. En affin avbildning erhålls genom att kombinera en linjär funktion $L : \mathbb{R}^n \to \mathbb{R}$, $L(x) = a \cdot x$, och förskjutningen $t \mapsto a_0 + t$ i $\mathbb{R}$. Naturligtvis är även avbildninen

$$x \mapsto a_0 + a \cdot (x - x_0)$$

affin, även om det vanligtvis inte är samma avbildning.

Låt $D \subset \mathbb{R}^n$ vara öppen, $f : D \to \mathbb{R}$ och $x_0 \in D$. Vi sätter ett plan att gå genom punkten $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$ och tolkar det som grafen av funktionen

$$x \mapsto a_0 + a \cdot (x - x_0),$$

där $a_0 \in \mathbb{R}$ och $a \in \mathbb{R}^n$. Denna går genom punkten $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$ exakt då, när $a_0 = f(x_0)$, alltså när

$$x \mapsto f(x_0) + a \cdot (x - x_0).$$

Detta leder till följande begrepp: Funktionen $f : D \to \mathbb{R}$ är *deriverbar d.v.s. differentierbar i punkten* $x_0 \in D$, om det existerar ett $a \in \mathbb{R}^n$, för vilket det gäller att

$$f(x) = f(x_0) + a \cdot (x - x_0) + |x - x_0| \varepsilon(x - x_0), \tag{2.4.3}$$

där $\varepsilon(x - x_0) \to 0$ då $x \to x_0$. Vektorn a kallas f:s *derivata i punkten* x_0 och betecknas $a = f'(x_0)$.

2.4.4. Anmärkning. Formeln (2.4.3) existerar alltid för vilket som helst $a \in \mathbb{R}^n$, ty vi kan välja

$$\varepsilon(x - x_0) = \begin{cases} \dfrac{f(x) - f(x_0) - a \cdot (x - x_0)}{|x - x_0|} & \text{, då } x \neq x_0,\ x \in A, \\ 0 & \text{, då } x = x_0. \end{cases}$$

Å andra sidan stämmer det vanligtvis inte att $\varepsilon(x - x_0) \to 0$ då $x \to x_0$.

2.4.5. Sats. *Om f är deriverbar i x_0, så är $f'(x_0)$ entydigt bestämd.*

Bevis. Övningsuppgift. □

2.4.6. Sats. *Låt f vara deriverbar i x_0. Då gäller att*

$$f'(x_0) = (\partial_1 f(x_0), \ldots, \partial_n f(x_0)).$$

Med andra ord existerar alla partiella derivator i x_0 och de ger $f'(x_0)$.

Bevis. Vi räknar den första koordinaten a_1 för vektorn $f'(x_0) = a$. De andra koordinaterna räknas på motsvarande sätt. Låt $x = x_0 + he_1$. Nu är $x \in D$ då $h \in \mathbb{R}$ är liten. Då $h \neq 0$, så substituerar vi x i formeln (2.4.3). Detta ger att

$$f(x_0 + he_1) - f(x_0) = a \cdot (he_1) + |h||e_1|\varepsilon(he_1).$$

Av detta följer att

$$\frac{f(x_0 + he_1) - f(x_0)}{h} = a \cdot e_1 + \frac{|h|}{h}\,\varepsilon(he_1).$$

Genom att ta gränsvärdet, då $h \to 0$, får vi att $\partial_1 f(x_0) = a \cdot e_1 = a_1$. □

2.4.7. Anmärkning. Ofta skrivs formeln (2.4.3) i formen

$$f(x_0 + h) = f(x_0) + f'(x_0) \cdot h + |h|\varepsilon(h), \tag{2.4.8}$$

där h är en vektor i $\mathbb{R}^n$ (vanligtvis är $|h|$ så litet att $f(x_0 + h)$ är definierad) och $\varepsilon(h) \to 0$ då $h \to 0$.

2.4.9. Sats. *Antag att $D \subset \mathbb{R}^n$ och att $x_0 \in D$. Om funktionen $f : D \to \mathbb{R}$ är differentierbar i punkten x_0, så är f kontinuerlig i punkten x_0.*

Bevis. Vi skall visa, att $\lim_{x \to x_0} f(x) = f(x_0)$, med andra ord skall vi visa att det av villkoret $x_i \to x_0,\ x_i \in D$ följer att $f(x_i) \to f(x_0)$. Nu är

$$f(x_i) = f(x_0) + f'(x_0) \cdot (x_i - x_0) + |x_i - x_0|\ \varepsilon(x_i - x_0)$$

och därmed gäller att

$$\begin{aligned} |f(x_i) - f(x_0)| &\leq |f'(x_0) \cdot (x_i - x_0)| + |x_i - x_0||\varepsilon(x_i - x_0)| \\ &\leq |f'(x_0)||(x_i - x_0)| + |x_i - x_0||\varepsilon(x_i - x_0)| \to 0. \end{aligned}$$

Här använde vi Cauchy-Schwartz olikhet, formel (1.1.3). Av detta följer att $f(x_i) \to f(x_0)$, vilket var det vi krävde. □

2.4.10. Anmärkning. Fastän funktionen f har de partiella derivatorna

$$\partial_1 f(x_0), \ldots, \partial_n f(x_0),$$

så är f inte nödvändigtvis deriverbar i x_0, ty de partiella derivatorna kan existera utan att f ens är kontinuerlig i x_0.

Vektorn $(\partial_1 f(x_0), \ldots, \partial_n f(x_0))$ kallas f:s *gradient* i punkten x_0 och betecknas $\nabla f(x_0)$. Formeln (2.4.3) ändras således till formen

$$f(x_0 + h) = f(x_0) + \nabla f(x_0) \cdot h + |h|\varepsilon(h).$$

Ifall vi vet att f är deriverbar i x_0, så finns det inga svårigheter med att bestämma derivatan $f'(x_0)$, alltså gradienten $\nabla f(x_0)$, eftersom

$$f'(x_0) = \nabla f(x_0) = (\partial_1 f(x_0), \ldots, \partial_n f(x_0)).$$

Man måste alltså endast räkna de partiella derivatorna. Däremot är det svårare att bestämma i vilka fall f är deriverbar i x_0.

2.4.11. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sin(xy)$. Då är $\partial_1 f(x, y) = \partial_x f(x, y) = y\cos(xy)$ och $\partial_2 f(x, y) = \partial_y f(x, y) = x\cos(xy)$. Om f är deriverbar i punkten (x, y), så är alltså

$$f'(x, y) = (y\cos(xy), x\cos(xy)).$$

Vi ger snart ett villkor för f:s deriverbarhet i punkten x_0. Villkoret visar bl.a., att funktionen f i exempel 2.4.11 är deriverbar i varje punkt i $\mathbb{R}^2$.

Om f är deriverbar i x_0, så kallas planet

$$x_{n+1} = f(x_0) + \partial_1 f(x_0)(x_1 - x_{0,1}) + \ldots + \partial_n f(x_0)(x_n - x_{0,n})$$

f:s *tangentplan i punkten* $(x_0, f(x_0))$. Detta är ett n-plan i $\mathbb{R}^{n+1}$. Det lönar sig vanligen att tänka sig detta plan som grafen av funktionen $\mathbb{R}^n \to \mathbb{R}$

$$x \mapsto f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

i $\mathbb{R}^{n+1}$. Grafen är ett plan, som går genom punkten $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$ och står vinkelrätt mot vektorn

$$(\nabla f(x_0), -1) = (\partial_1 f(x_0), \ldots, \partial_n f(x_0), -1)$$

i $\mathbb{R}^{n+1}$.

2.4.12. Anmärkning. Fallet $n = 1$:

$$y = f(x_0) + f'(x_0)(x - x_0),$$

där $y = x_{n+1}$. Normalen för denna linje är vektorn $(f'(x_0), -1)$ i $\mathbb{R}^2$.

Om f är deriverbar i x_0, så kan f approximeras i en liten omgivning av punkten x_0 med den affina avbildningen

$$x \mapsto f(x_0) + \nabla f(x_0) \cdot (x - x_0), \tag{2.4.13}$$

se formel (2.4.2). Detta är den viktigaste egenskapen för funktioner f som är deriverbara i punkten x_0. Vi säger att funktionen f lokalt kan approximeras nära punkten x_0 med en affin avbildning bestämd av formel (2.4.13). Observera, att avbildningen (2.4.13) är definierad i hela $\mathbb{R}^n$.

2.4.14. Anmärkning. Avbildningen $h \mapsto \nabla f(x_0) \cdot h$ är en linjär avbildning $\mathbb{R}^n \to \mathbb{R}$. I själva verket kan vi med kunskaper från lineäralgebran visa att varje linjär avbildning $L : \mathbb{R}^n \to \mathbb{R}$ kan skrivas i formen $L(h) = a \cdot h$ för något $a \in \mathbb{R}^n$.

2.5. Deriveringsregler

Räkning med partiella derivator sker med samma regler som i det endimensionella fallet. I fallet av sammansatta funktioner förutsätter dessa regler dock ofta deriverbarhet (inte enbart partiella derivatornas existens), så vi granskar först följande fråga. När kan man dra slutsatsen, att funktionen $f : D \to \mathbb{R}$ är deriverbar i punkten $x_0 \in D$?

2.5.1. Definition. Låt $D \subset \mathbb{R}^n$ vara en öppen mängd och $f : D \to \mathbb{R}$. Funktionen f *tillhör klassen* $\mathcal{C}^1(D)$ (en gång kontinuerligt deriverbara funktioner i D), om

a) de partiella derivatorna $\partial_1 f(x_0), \ldots, \partial_n f(x_0)$ existerar i varje punkt i D och
b) dessa är kontinuerliga i D.

I definitionen ovan antar vi ofta, att även f är kontinuerlig i mängden D. Detta är dock onödigt. Med stöd av följande sats ger villkoren a) och b) att funktionen f är kontinuerlig, ty en deriverbar funktion är alltid kontinuerlig, såsom Sats 2.5.6 visar.

2.5.2. Sats. *Om $f \in \mathcal{C}^1(D)$, så är f deriverbar i varje punkt i D.*

Bevis då $n = 2$. Låt $f : D \to \mathbb{R}$ och $x_0 = (a,b) \in D$. Vi skall visa, att f är deriverbar i punkten x_0, med andra ord att

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + |x - x_0|\, \varepsilon(x - x_0), \tag{2.5.3}$$

där $\varepsilon(x - x_0) \to 0$ då $x \to x_0$. Märk, att $\nabla f(x_0)$ existerar, eftersom de partiella derivatorna $\partial_1 f(x_0)$ och $\partial_n f(x_0)$ existerar enligt antagandet. Såsom vi tidigare konstaterade, så gäller alltid formel (2.5.3); det enda problemet är funktionens ε gränsvärde.

Vi fixerar en kvadrat $Q \subset D$, vars sidor har längden $2r$ och vars medelpunkt är i $x_0 = (a,b)$. Låt $x = (a+h, b+k) \in Q$, där $(h,k) \neq 0$. Då är $|h|, |k| < r$. Nu gäller att

$$\begin{aligned} f(x) - f(x_0) &= f(a+h, b+k) - f(a,b) \\ &= f(a+h, b+k) - f(a, b+k) + f(a, b+k) - f(a,b). \end{aligned}$$

Med ett fixerat k är funktionen $\varphi(t) = f(a+t, b+k) - f(a, b+k)$ definierad i intervallet $(-r, r)$ och deriverbar i det ifrågavarande intervallet, eftersom $\varphi'(t) = \partial_1 f(a+t, b+k)$. Genom att tillämpa den vanliga medelvärdessatsen i intervallet $[0,h]$ (eller i intervallet $[h,0]$) får vi att

$$\begin{aligned} f(a+h, b+k) - f(a, b+k) &= \varphi(h) - \varphi(0) \\ &= \varphi'(\xi)(h-0) = \partial_1 f(a+\xi, b+k)h. \end{aligned}$$

På samma sätt får vi genom att tillämpa medelvärdessatsen på funktionen

$$\psi(s) = f(a, b+s) - f(a,b)$$

att

$$\begin{aligned} f(a, b+k) - f(a,b) &= \psi(k) - \psi(0) \\ &= \psi'(\theta)(k-0) = \partial_2 f(a, b+\theta)k. \end{aligned}$$

Genom att kombinera dessa med den föregående ekvationen har vi att

$$\begin{aligned} f(x) - f(x_0) &= \partial_1 f(a+\xi, b+k)h + \partial_2 f(a, b+\theta)k \\ &= (\partial_1 f(a+\xi, b+k), \partial_2 f(a, b+\theta)) \cdot (h,k). \end{aligned}$$

Talet ξ är mellan 0 och h och talet θ är på motsvarande sätt mellan 0 och k; märk att h och k även kan vara negativa. Genom att tillämpa detta på formel (2.5.3) får vi följande uttryck för funktionen ε

$$\varepsilon(x-x_0) = \frac{1}{|x-x_0|}\,[(\partial_1 f(a+\xi, b+k),\, \partial_2 f(a, b+\theta)) - (\partial_1 f(x_0),\, \partial_2 f(x_0))]\cdot(h,k)$$

$$= \frac{1}{|x-x_0|}\,[(\partial_1 f(a+\xi,b+k) - \partial_1 f(x_0),\, \partial_2 f(a,b+\theta) - \partial_2 f(x_0))\cdot(h,k)].$$

Eftersom $x - x_0 = (h,k)$, följer av detta att

$$|\varepsilon(x-x_0)| \le \frac{1}{|(h,k)|}\,|(\partial_1 f(a+\xi, b+k) - \partial_1 f(x_0),\, \partial_2 f(a,b+\theta) - \partial_2 f(x_0)||(h,k)|$$

$$= |(\partial_1 f(a+\xi,b) - \partial_1 f(x_0), \partial_2 f(a,b+\theta) - \partial_2 f(x_0)| \to 0,$$

då $(h,k) \to 0$. Detta gäller, eftersom $\partial_1 f$ och $\partial_2 f$ är kontinuerliga i punkten $x_0 = (a,b)$, varvid $\partial_1 f(a+\xi, b) \to \partial_1 f(x_0)$ och $\partial_2 f(a, b+\theta) \to \partial_2 f(x_0)$, då $(h,k) \to 0$. Påståendet följer av detta. $\square$

2.5.4. Exempel. Vi bestämmer gradienten $\nabla f(x)$ för funktionen $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = |x| = \sqrt{x_1^2 + \ldots + x_n^2}$ och undersöker deriverbarheten för funktionen f. Partiella derivator existerar inte i punkten $x = 0$ (övningsuppgift; detta motsvarar det endimensionella fallet, där funktionen $x \mapsto |x|$ inte har någon derivata i 0). Således är funktionen f inte deriverbar i punkten 0. Vi räknar $\nabla f(x)$ då $x \neq 0$. Vi fixerar $i = 1, 2, \ldots, n$ och undersöker funktionen x_i med en variabel

$$x_i \mapsto \sqrt{x_1^2 + \ldots + x_i^2 + \ldots + x_n^2}.$$

Genom att räkna dennas derivata får vi att

$$\partial_i f(x) = 1/2\,(x_1^2 + \ldots + x_n^2)^{-1/2}\,2x_i = \frac{x_i}{|x|}.$$

Nu är funktionen f deriverbar i mängden $\mathbb{R}^n \backslash \{0\}$. Orsaken till detta är följande. Funktionen $x \mapsto |x|$ är kontinuerlig i $\mathbb{R}^n$, av vilket det enkelt följer (såsom i fallet $n = 1$), att $x \mapsto 1/|x|$ är kontinuerlig i mängden $\mathbb{R}^n \backslash \{0\}$. Funktionen $x \mapsto x_i$ är uppenbart kontinuerlig i $\mathbb{R}^n$ och produkten av två kontinuerliga funktioner är kontinuerlig, så funktionen $x \mapsto x_1/|x|$ är kontinuerlig i mängden $\mathbb{R}^n \backslash \{0\}$. Detta gäller för varje $i = 1, 2, \ldots, n$. Med stöd av Sats 2.5.2 är f således deriverbar i mängden $\mathbb{R}^n \setminus \{0\}$. För gradienten $\nabla f(x)$ av funktionen f får vi i punkter $x \neq 0$ uttrycket

$$\begin{aligned}\nabla f(x) = (\partial_1 f(x), \ldots, \partial_n f(x)) &= \left(\frac{x_1}{|x|}, \ldots, \frac{x_n}{|x|}\right) \\ = \frac{1}{|x|}(x_1, \ldots, x_n) &= \frac{x}{|x|}.\end{aligned}$$

2.5.5. Anmärkning. Fallet $n = 1$: Låt $f : \mathbb{R} \to \mathbb{R},\ f(x) = |x|$. Då är

$$f'(x) = \nabla f(x) = \frac{x}{|x|} = \begin{cases} = 1 & , \text{då } x > 0 \\ = -1 & , \text{då } x < 0 \end{cases}$$

2.5.6. Sats. *Låt $D \subset \mathbb{R}^n$ och $x_0 \in D$. Om funktionen $f : D \to \mathbb{R}$ är deriverbar i punkten x_0, så är f kontinuerlig i punkten x_0.*

2.5.7. Anmärkning. Detta motsvarar den kända satsen i fallet $n = 1$. Kom ihåg, att kontinuiteten för f i punkten x_0 inte följer av att de partiella derivatorna i x_0 existerar.

Bevis för Sats 2.5.6. Eftersom f är deriverbar i punkten x_0, så gäller det att

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + |x - x_0|\ \varepsilon(x - x_0).$$

Av detta följer att

$$\begin{aligned}|f(x) - f(x_0)| &= |\nabla f(x_0) \cdot (x - x_0) + |x - x_0|\ \varepsilon(x - x_0)| \\ &\leq |\nabla f(x_0)||x - x_0| + |x - x_0||\varepsilon(x - x_0)|.\end{aligned}$$

Vi låter nu följden (x_i) vara sådan, att $x_i \to x_0$. Eftersom det gäller att $|\varepsilon(x_i - x_0)| \to 0$, ser vi genast ur estimatet att $f(x_i) \to f(x_0)$, alltså att f är kontinuerlig i x_0. $\square$

Vi observerar alltså, att om $f \in \mathcal{C}^1(D)$, så är f kontinuerlig och deriverbar i varje punkt i D.

Vi granskar härnäst härledningen av f:s partiella derivator då f är en sammansatt funktion. Vi härleder alltså de så kallade *kedjereglerna* för derivering av sammansatta funktioner. Vi undersöker först vissa specialfall. Det allmänna fallet behandlas i kapitel 2.6.

A) Låt $D \subset \mathbb{R}^n$ vara öppen, $\Delta \subset \mathbb{R}$ ett intervall, $h : D \to \Delta$ och $g : \Delta \to \mathbb{R}$. Vi granskar derivering av funktionen $f = g \circ h$.

Detta är enkelt, eftersom man räknar partiella derivator genom att begränsa funktionen till linjer riktade längs $\mathbb{R}^n$:s koordinataxlar. Vi bestämmer i detta fall $\partial_1 f$. Enligt definitionen är

$$\partial_1 f(x_0) = \lim_{t\to 0} \frac{g \circ h(x_0 + te_1) - g \circ h(x_0)}{t}.$$

Eftersom $x_0 \in D$ och D är öppen, är funktionen

$$\varphi(t) = g \circ h(x_0 + te_1)$$

definierad i något intervall $(-\delta, \delta)$, $\delta > 0$. Genom att räkna derivatan i 0 för denna funktion med den vanliga endimensionella kedjeregeln får vi att

$$\begin{aligned}\varphi'(0) &= g'(h(x_0 + 0e_1)) \cdot \partial_1 h(x_0 + 0e_1)\\ &= g'(h(x_0))\partial_1 h(x_0).\end{aligned}$$

Å andra sidan är

$$\begin{aligned}\varphi'(0) &= \lim_{t\to 0} \frac{g \circ h(x_0 + te_1) - g \circ h(x_0)}{t}\\ &= \partial_1 f(x_0).\end{aligned}$$

Vi upptäcker alltså att $\partial_1 f(x_0) = g'(h(x_0))\partial_1(x_0)$, om $\partial_1 h(x_0)$ och $g'(h(x_0))$ existerar. Generellt får vi kedjeregeln

$$\partial_i f(x_0) = g'(h(x_0))\, \partial_i f(x_0),\ i = 1, \ldots, n.$$

2.5.8. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sin(xy)$. Nu är $f = g \circ h$, där $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x, y) = xy$ och $g : \mathbb{R} \to \mathbb{R}$, $g(s) = \sin s$. Vi bestämmer $\partial_2 f(x, y)$. Formeln ovan ger att

$$\begin{aligned}\partial_2 f(x, y) &= g'(h(x, y)) \cdot \partial_2 h(x, y)\\ &= \cos(h(x, y))x\\ &= \cos(xy)x.\end{aligned}$$

Samma resultat får vi förstås direkt genom att fixera x och derivera med avseende på y.

B) Antag att $D, D' \subset \mathbb{R}^n$ är öppna, $w : D \to D'$ och $h : D' \to \mathbb{R}$. Vi undersöker derivering av funktionen $f = h \circ w$.

Avbildningar, alltså funktioner w, vars värden är i $\mathbb{R}^n$, har inte behandlats tidigare. Dessa reduceras till det realvärda fallet på följande sätt. Enligt definitionen för avbildningen $w : D \to D'$ är det för varje $x \in D$ kopplat ett entydigt $w(x) \in D'$, med andra ord $w(x) = (y_1, \ldots, y_n)$, där y_i är en koordinat i $w(x)$. Således definierar w n stycken funktioner $y_i = w_i(x),\ x \in D$. Dessa är avbildningar $D \to \mathbb{R}$, med andra ord är $w(x) = (w_1(x), \ldots, w_n(x))$, där w_i är den i:te *koordinatfunktionen* för w. Således är

$$f(x) = h(w(x)) = h(w_1(x), \ldots, w_n(x)).$$

2.5.9. Sats. *Antag att funktionerna w_i är deriverbara i punkten $x_0 \in D$ och att h är deriverbar i punkten $w(x_0)$. Då är f deriverbar i punkten x_0 och*

$$\partial_i f(x_0) = \sum_{j=1}^{n} \partial_j h(w(x_0)) \partial_i w_j(x_0),\ i = 1, 2, \ldots, n.$$

Bevis. Vi bevisar formeln i fallet $i = 1$. De övriga indexen i går på samma sätt. Vi skall alltså visa, att

$$\partial_1 f(x_0) = \sum_{j=1}^{n} \partial_j h(w(x_0)) \partial_1 w_j(x_0).$$

Enligt definitionen är

$$\partial_1 f(x_0) = \lim_{t \to 0} \frac{f(x_0 + te_1) - f(x_0)}{t},$$

då $t \neq 0$ är så litet, att $x_0 + te_1 \in D$. Nu är

$$\frac{f(x_0 + te_1) - f(x_0)}{t} = \frac{h(w(x_0 + te_1)) - h(w(x_0))}{t}.$$

Eftersom h är deriverbar i punkten $w(x_0)$, gäller det för varje $y \in D'$ att

$$h(y) - h(w(x_0)) = \nabla h(w(x_0)) \cdot (y - w(x_0)) + |y - w(x_0)|\ \varepsilon(y - w(x_0),$$

jämför med formel (2.4.8). Genom att välja $y = w(x_0 + te_1) \in D'$ och dividera med t får vi att

$$\begin{aligned}
&\frac{h(w(x_0+te_1)) - h(w(x_0))}{t} \\
&= \nabla h(w(x_0)) \cdot \frac{w(x_0+te_1) - w(x_0)}{t} \\
&\qquad + \frac{|w(x_0+te_1) - w(x_0)|}{t}\, \varepsilon(w(x_0+te_1) - w(x_0)).
\end{aligned}$$

Genom att uttrycka detta med koordinatfunktioner får vi att

$$\begin{aligned}
&\frac{w(x_0+te_1) - w(x_0)}{t} \\
&= \left(\frac{w_1(x_0+te_1) - w_1(x_0)}{t}, \dots, \frac{w_n(x_0+te_1) - w_n(x_0)}{t} \right) \\
&\to (\partial_1 w_1(x_0), \dots, \partial_1 w_n(x_0)),
\end{aligned}$$

då $t \to 0$. Genom att kombinera detta med den tidigare formeln och låta $t \to 0$ märker vi att

$$\begin{aligned}
&\partial_1 f(x_0) \\
=&\nabla h(w(x_0)) \cdot (\partial_1 w_1(x_0), \dots, \partial_1 w_n(x_0)) + |\partial_1 w_1(x_0), \dots, \partial_1 w_n(x_0))| \cdot 0 \\
=&\sum_{j=1}^{n} \partial_j h(w(x_0)) \partial_1 w_j(x_0).
\end{aligned}$$

Detta är summaformeln vi krävde. □

2.5.10. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \sin(\sin(x)e^{xy}).$$

Det är möjligt att bestämma den partiella derivatan $\partial_1 f(x,y)$ på vanligt sätt genom att fixera y och derivera med avseende på x. Vi använder trots det kedjeregeln genom att sätta

$$\begin{aligned}
h(x,y) &= \sin(xy) \\
w(x,y) &= (w_1(x,y), w_2(x,y)) = (\sin x, e^{xy}).
\end{aligned}$$

Då är

$$\begin{aligned}
(h \circ w)(x,y) &= h(w(x,y)) = h(w_1(x,y), w_2(x,y)) \\
&= \sin(\sin(x)e^{xy}) = f(x,y).
\end{aligned}$$

Med en direkt räkning ser vi att gradienten för funktionen h är

$$\begin{aligned}\nabla h(x,y) &= (\partial_1 h(x,y), \partial_2 h(x,y)) \\ &= (y\cos(xy), x\cos(xy))\end{aligned}$$

och de partiella derivatorna för funktionen w är

$$\partial_1 w_1(x,y) = \cos x,\ \ \partial_1 w_2(x,y) = ye^{xy}.$$

Alla dessa funktioner är kontinuerliga, så vi kan använda kedjeregeln. Den ger att

$$\partial_1 f(x,y) = e^{xy}\cos(\sin x e^{xy}) \cdot \cos x + \sin x \cos(\sin x e^{xy}) \cdot ye^{xy}.$$

C) Följande situation kommer ofta till mötes. Antag att $\Delta \subset \mathbb{R}$ är ett intervall, $D \subset \mathbb{R}^n$ en öppen mängd och att $g : D \to \mathbb{R}$ samt $\gamma : \Delta \to D$ är avbildningar. Funktionen γ är alltså en avbildning från intervallet Δ till den öppna mängden $D \subset \mathbb{R}^n$. $\gamma(t)$ är för varje $t \in D$ en vektor i $\mathbb{R}^n$

$$(\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t)),$$

med andra ord definierar γ koordinatfunktionerna $\gamma_i : \Delta \to \mathbb{R}$, $i = 1, \ldots, n$. Avbildningen γ sägs vara kontinuerlig, om varje γ_i är kontinuerlig, och en kontinuerlig avbildning $\gamma : \Delta \to D$ kallas en *stig* i mängden D. Derivatan $\gamma'(t)$ av avbildningen γ i punkten $t \in \Delta$ definieras som vektorn

$$\gamma'(t) = (\gamma_1'(t), \ldots, \gamma_n'(t)),$$

där $\gamma_i'(t)$ är vanlig derivata. Detta förutsätter förstås, att varje derivata $\gamma_i'(t)$ existerar i punkten t. Vi skall nu bestämma derivatan av den vanliga realvärda funktionen $u : \Delta \to \mathbb{R}$ i punkten $x_0 \in \Delta$, då u är av formen $u = g \circ \gamma$.

2.5.11. Sats. *Antag att $t_0 \in \Delta$, att $\gamma'(t_0)$ existerar och att $g : D \to \mathbb{R}$ är deriverbar i punkten $\gamma(t_0)$. Då är*

$$h'(t_0) = \nabla g(\gamma(t_0)) \cdot \gamma'(t_0).$$

Bevis. Övningsuppgift. □

Stigar och övriga vektorvärda funktioner behandlas närmare i kapitel 3.

2.6. Derivata och kedjeregeln i det allmänna fallet

Antag att $D \subset \mathbb{R}^n$ och att $f : D \to \mathbb{R}^p,\ p \geq 1$ är en avbildning. För varje $x \in D$ gäller då att

$$f(x) = (f_1(x), \ldots, f_p(x)),$$

där $f_j : D \to \mathbb{R},\ j = 1, \ldots, p$, är *koordinatfunktioner* till f. Avbildningen f sägs vara *kontinuerlig i punkten* $x_0 \in D$, om det för varje följd (x_i) i mängden D, för vilken $x_i \to x_0$, gäller att $f(x_i) \to f(x_0)$. Såsom vi tidigare konstaterade, är detta ekvivalent med att varje koordinatfunktion $f_j,\ j = 1, \ldots, p$, är kontinuerlig i punkten x_0. Avbildningen är *kontinuerlig i mängden* D, om den är kontinuerlig i varje punkt i D.

Vi antar att D är en öppen delmängd i $\mathbb{R}^n$ och att $x_0 \in D$. Avbildningen $f : D \to \mathbb{R}^p$ är *deriverbar*, alltså *differentierbar* i punkten x_0, om det existerar en sådan lineärfunktion

$$L : \mathbb{R}^n \to \mathbb{R}^p,$$

att

$$f(x_0 + h) - f(x_0) = L(h) + |h|\varepsilon(h), \tag{2.6.1}$$

där $\varepsilon(h) \to 0$, då $h \to 0$. Detta betyder alltså, att om (h_i) är en sådan följd vektorer, att $h_i \to 0$ och $h_i \neq 0$, så gäller att $\varepsilon(h_i) \to 0$. Detta är ekvivalent med att det för varje $\varepsilon > 0$ existerar ett sådant $\delta > 0$, att

$$|\varepsilon(h)| < \varepsilon, \text{ då } |h| < \delta. \tag{2.6.2}$$

Av öppenheten för mängden D följer att ε är definierad åtminstone i någon kula $B(0, r),\ r > 0$. Genom att sätta $\varepsilon(0) = 0$ upptäcker vi att villkoret (2.6.2) är ekvivalent med att ε är kontinuerlig i origo. Såsom tidigare är det skäl att observera, att formeln (2.6.1) gäller för varje lineärfunktion $L : \mathbb{R}^n \to \mathbb{R}^p$; deriverbarhet beror däremot på villkoret $\varepsilon(h) \to 0$ då $h \to 0$.

Om f är deriverbar i punkten x_0, så betecknar vi den lineära funktionen L i formeln (2.6.1) med $f'(x_0)$. Följande sats är inte svår att bevisa; dess bevis är likadant som i fallet $p = 1$, jämför Sats 2.5.6.

2.6.3. Sats. *Låt $f : D \to \mathbb{R}^p$ vara deriverbar i punkten $x_0 \in D$. Då är*

(i) *f kontinuerlig i punkten x_0,*

(ii) *lineärfunktionen $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^p$ entydigt bestämd,*

(iii) *koordinatfunktionerna f_j, $j = 1, \ldots, p$, deriverbara i punkten x_0.*

Avbildningen $f : D \to \mathbb{R}^p$ kan uttryckas med hjälp av sina koordinatfunktioner, $f = (f_1, \ldots, f_p)$, där $f_j : D \to \mathbb{R}$, $j = 1, \ldots, n$. Vi antar, att f är deriverbar i punkten x_0. Då är $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^p$ en lineärfunktion och varje lineärfunktion har en avbildningsmatris, som i detta fall är en $(p \times n)$-matris (naturligtvis i standardbasen för $\mathbb{R}^n$ och $\mathbb{R}^p$)

$$f'(x_0) = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & \ddots & \\ a_{p1} & a_{p2} & \ldots & a_{pn} \end{bmatrix}.$$

Detta betyder, att då $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ så är $f'(x_0)\, h$ en vektor i $\mathbb{R}^p$

$$\begin{aligned} f'(x_0)\, h &= \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \\ a_{p1} & \ldots & a_{pn} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \\ &= (a_{11}\, h_1 + \ldots + a_{1n}\, h_n,\ \ldots,\ a_{p1}\, h_1 + \ldots + a_{pn}\, h_n). \end{aligned}$$

I det föregående betecknades, såsom det är vanligt i samband med linjära avbildningar, $f'(x_0)(h) = f'(x_0)\, h$, med andra ord utelämnas parenteserna runt avbildningens argument

Om f är deriverbar i punkten x_0 och linjärfunktionen $f'(x_0)$ har den ovannämnda avbildningsmatrisen, så bestämmer vi nu talen a_{ji}, $1 \le j \le p$, $1 \le i \le n$, skrivna med hjälp av de partiella derivatorna av f:s koordinatfunktioner f_j. Vi väljer i formeln (2.6.1) $h = te_1$, $t \neq 0$ och $L = f'(x_0)$. Nu gäller p.g.a. lineäriteten hos avbildningen $f'(x_0)$ att

$$f'(x_0)(te_1) = t\, f'(x_0)\, e_1$$

och å andra sidan är

$$f(x_0 + te_1) - f(x_0) = (f_1(x_0 + te_1) - f_1(x_0), \ldots, f_p(x_0 + te_1) - f_p(x_0)).$$

Vi får alltså formeln (2.6.1), genom att dividera med t, i formen

$$\begin{aligned}
&\left(\frac{1}{t}\left(f_1(x_0+te_1)-f_1(x_0)\right),\ldots,\frac{1}{t}\left(f_p(x_0+te_1)-f_p(x_0)\right)\right)\\
&=f'(x_0)\,e_1+\frac{|t|}{t}\,\varepsilon(te_1)\\
&=\begin{bmatrix} a_{11} & \ldots & a_{1n}\\ a_{21} & \ldots & a_{2n}\\ \vdots & \ddots & \\ a_{p1} & \ldots & a_{pn}\end{bmatrix}\begin{bmatrix}1\\0\\ \ldots\\0\end{bmatrix}+\frac{|t|}{t}\,\varepsilon(te_1)\\
&=(a_{11},a_{21},\ldots,a_{p1})+\frac{|t|}{t}\,\varepsilon(te_1).
\end{aligned}$$

Eftersom $\varepsilon(te_1)\to 0$, då $t\to 0$, och eftersom koordinatfunktionerna f_j har de partiella derivatorna

$$\frac{1}{t}\left(f_j(x_0+te_1)+f_j(x_0)\right)\to \partial_1 f_j(x_0),$$

så ser vi genom att låta $t\to 0$, att vektorerna

$$(\partial_1 f_1(x_0),\partial_1 f_2(x_0),\ldots,\partial_1 f_p(x_0))$$

och $(a_{11},a_{21},\ldots,a_{p1})$ är lika och därmed är

$$a_{j1}=\partial_1 f_j(x_0)$$

för varje $j=1,\ldots,p$. På motsvarande sätt får vi för varje $i=1,\ldots,n$ att

$$a_{ji}=\partial_i f_j(x_0),$$

där $1\le i\le n$ och $1\le j\le p$. Således är avbildningsmatrisen för $f'(x_0)$ i formen

$$f'(x_0)=\begin{bmatrix}\partial_1 f_1(x_0) & \partial_2 f_1(x_0) & \ldots & \partial_n f_1(x_0)\\ \partial_1 f_2(x_0) & \partial_2 f_2(x_0) & \ldots & \partial_n f_2(x_0)\\ \vdots & & \ddots & \\ \partial_1 f_p(x_0) & \partial_2 f_p(x_0) & \ldots & \partial_n f_p(x_0)\end{bmatrix}. \tag{2.6.4}$$

Ovan har vi i själva verket bevisat ett tidigare resultat, Sats 2.4.6, i en allmännare form. I formeln i Sats 2.4.6 fanns det endast en horisontell rad.

Såsom tidigare i ett enklare fall, så gäller att det av alla koordinatfunktioners deriverbarhet följer deriverbarhet för f. Dessutom gäller, att om alla funktioner $f_j \in \mathcal{C}^1(D),\ 1 \le j \le p$, så är f deriverbar i varje punkt i D. Att bevisa dessa påståenden är inte väsentligt svårare än de tidigare resultaten.

Om $p = 1$, alltså om $f : D \to \mathbb{R}$, så har vi definierat deriverbarheten i punkten $x_0 \in D$ med formel (2.4.3), där vektorn $a = \nabla f(x_0)$. Deriverbarhet är i detta fall exakt detsamma som det ovan definierade. Orsaken är att lineärfunktionen $f'(x_0) : \mathbb{R}^n \to \mathbb{R}$ ändras till formen

$$f'(x_0)\, h = \nabla f(x_0) \cdot h, \tag{2.6.5}$$

vilket vi ser direkt från uttrycket (2.6.4). Märk, att varje lineärfunktion $L : \mathbb{R}^n \to \mathbb{R}$ kan skrivas i formen $Lh = a \cdot h$, $h \in \mathbb{R}^n$, för något fixerat $a \in \mathbb{R}^n$.

Vi undersöker till slut kedjeregeln i allmänt fall. Antag att $D \subset \mathbb{R}^n$ och $D' \subset \mathbb{R}^p$ är öppna mängder samt att $f : D \to D'$ och $g : D' \to \mathbb{R}^q$ är avbildningar. Följande sats är inte svårare att bevisa är de tidigare kedjereglerna.

2.6.6. Sats. *Om f är deriverbar i punkten x_0 och g är deriverbar i punkten $f(x_0)$, så är den sammansatta avbildningen $g \circ f : D \to \mathbb{R}^q$ deriverbar i punkten x_0 och*

$$(g \circ f)'(x_0) = g'(f(x_0))\, f'(x_0). \tag{2.6.7}$$

Bevis. Förbigås. □

Formel (2.6.7) betyder, att vi erhåller lineärfunktionen $(g \circ f)'(x_0) : \mathbb{R}^n \to \mathbb{R}^q$ genom att kombinera lineärfunktionerna $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^p$ och $g'(f(x_0)) : \mathbb{R}^p \to \mathbb{R}^q$; som vanligt använder vi inte beteckningen $\circ$ mellan två lineärfunktioner. Avbildningsmatrisen för linjärfunktionen $(g \circ f)'(x_0)$ får vi genom att multiplicera avbildningsmatriserna för $g'(f(x_0))$ och $f'(x_0)$. Alla tidigare kedjeregler är följdsatser av Sats 2.6.6.

2.6.8. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (\sin x, xy)$ och $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x, y) = xy$. Då är $g \circ f : \mathbb{R}^2 \to \mathbb{R}$,

$$(g \circ f)(x, y) = g(f(x, y)) = g(\sin x, xy) = xy \sin x.$$

Nu är g, f, $g \circ f \in \mathcal{C}^1(\mathbb{R}^2)$, och således är alla avbildningar deriverbara. Vi räknar gradienten $\nabla(g \circ f)(x, y)$ med formeln i Sats 2.6.6

$$(g \circ f)'(x, y) = \underbrace{\underbrace{g'(f(x, y))}_{\text{lineärfunktion } \mathbb{R}^2 \to \mathbb{R}} \circ \underbrace{f'(x, y)}_{\text{lineärfunktion } \mathbb{R}^2 \to \mathbb{R}^2}}_{\text{lineärfunktion sammansatt av två lineärfunktioner } \mathbb{R}^2 \to \mathbb{R}}$$

genom att använda sambandet mellan derivatan och gradienten (2.6.5),

$$(g \circ f)'(x, y)h = \nabla(g \circ f)(x, y) \cdot h,$$

där $h \in \mathbb{R}^2$. Ur formeln 2.6.4 får vi avbildningsmatrisen av funktionen $f'(x, y)$

$$f'(x, y) = \begin{bmatrix} \partial_1 f_1(x, y) & \partial_2 f_1(x, y) \\ \partial_1 f_2(x, y) & \partial_2 f_2(x, y) \end{bmatrix} = \begin{bmatrix} \cos x & 0 \\ y & x \end{bmatrix}.$$

Genom att tillämpa formel (2.6.5) på nytt på avbildningen g får vi att

$$\begin{aligned} g'(x, y)h &= \nabla g(x, y) \cdot h \\ &= (g_1(x, y), g_2(x, y)) \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \\ &= (y, x) \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = y h_1 + x h_2. \end{aligned}$$

Vi räknar sedan den sammansatta lineärfunktionen $g'(f(x, y)) \circ f'(x, y)$ tillämpad på vektorn $h \in \mathbb{R}^2$; detta får vi direkt ur de föregående formlerna. Vi har att

$$\begin{aligned} g'(f(x, y))\, f'(x, y)\, h &= g'(f(x, y))(f'(x, y)\, h) \\ &= (xy, \sin x) \left(\begin{bmatrix} \cos x & 0 \\ y & x \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \right) \\ &= (xy, \sin x) \begin{bmatrix} (\cos x)\, h_1 \\ y\, h_1 + x\, h_2 \end{bmatrix} \\ &= (xy\, (\cos x)\, h_1 + (\sin x)(y\, h_1 + x\, h_2) \\ &= (xy\ \cos x + y\ \sin x) h_1 + x\, (\sin x)\, h_2 \\ &= (xy\ \cos x + y\ \sin x, x\ \sin x) \cdot (h_1, h_2). \end{aligned}$$

Således är $\nabla(g \circ f)(x, y) = (xy\ \cos x + y\ \sin x, x\ \sin x)$. Samma resultat fås enklare från de partiella derivatorna med direkt räkning.

2.7. Riktad derivata och gradientens geometriska betydelse

Vi behandlar saken endast då $n = 2$. Generalisering till fallet $n \geq 3$ är enkel. De partiella derivatorna i punkten x_0 av realvärda funktioner är tagna i koordinataxlarnas riktning. Lika väl kan vi definiera en "partiell derivata" i vilken riktning som helst.

Antag att $f : D \to \mathbb{R}$, D öppen i $\mathbb{R}^2$ och att $x_0 \in D$ samt att $v \in \mathbb{R}^2$, $|v| = 1$ d.v.s. v är en "enhetsvektor". Den *riktade derivatan i punkten* x_0 för funktionen f *i riktningen* v är

$$f'_v(x_0) = \partial_v f(x_0) = D_v f(x_0) = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t},$$

ifall det ifrågavarande gränsvärdet existerar. Vi observerar, att

$$\partial_1 f(x_0) = \partial_{e_1} f(x_0), \; \partial_2 f(x_0) = \partial_{e_2} f(x_0).$$

2.7.1. Sats. *Låt f vara deriverbar i punkten x_0. Då är*

$$\partial_v f(x_0) = \nabla f(x_0) \cdot v$$

och

$$|\partial_v f(x_0)| \leq |\nabla f(x_0)|.$$

2.7.2. Anmärkning. I Sats 2.7.1 är olikhetens vänstra sida absolutbetoppet av ett reelt tal och den högra sidan längden av en vektor.

Bevis för Sats 2.7.1. Eftersom f är deriverbar i punkten x_0, gäller det att

$$f(x_0 + h) - f(x_0) = \nabla f(x_0) \cdot h + |h| \, \varepsilon(h).$$

Vi väljer $h = tv, \; t \neq 0$. Då är

$$\begin{aligned} \frac{f(x_0 + tv) - f(x_0)}{t} &= \frac{\nabla f(x_0) \cdot (tv)}{t} + \frac{|tv|}{t} \, \varepsilon(tv) \\ &= \nabla f(x_0) \cdot v + \frac{|tv|}{t} \, \varepsilon(tv). \end{aligned}$$

Nu är

$$\frac{|tv|}{t} \, \varepsilon(tv) \to 0,$$

då $t \to 0$, eftersom

$$\left| \frac{|tv|}{t} \right| = |v|$$

och $\varepsilon(tv) \to 0$, då $t \to 0$. Därmed är $\partial_v f(x_0) = \nabla f(x_0) \cdot v$. Dessutom, eftersom $|v| = 1$, får vi med stöd av Cauchy-Schwartz olikhet, formel (1.1.3), att

$$|\partial_v f(x_0)| = |\nabla f(x_0) \cdot v| \leq |\nabla f(x_0)||v| = |\nabla f(x_0)|.$$

□

2.7.3. Exempel. Vi räknar den riktade derivatan i origo för funktionen $f(x, y) = \sin x + y$ i riktningen $\frac{1}{\sqrt{2}}(e_1 + e_2) = v$. Observera, att $|v| = 1$. Avbildningen $f \in \mathcal{C}^1(\mathbb{R}^2)$, så f är deriverbar i origo. Dessutom gäller att

$$\nabla f(x) = (\partial_1 f(x), \partial_2 f(x)) = (\cos x, 1),$$

varvid $\nabla f(0) = (1, 1)$. Sats 2.7.1 ger att

$$\partial_v f(0) = (1, 1) \cdot (\frac{1}{\sqrt{2}}(1, 1)) = \frac{1}{\sqrt{2}}(1 + 1) = \frac{2}{\sqrt{2}} = \sqrt{2}.$$

Detta kan även räknas direkt från definitionen.

Härnäst utreder vi gradientens ∇f geometriska betydelse. Låt $D \subset \mathbb{R}^2$ vara öppen och $f : D \to \mathbb{R}$. Vi antar, att $f \in \mathcal{C}^1(D)$. Vi undersöker f:s nivåyta $f(x, y) = c$. Vi antar, att det existerar en stig

$$\gamma(t) = (x(t), y(t)),$$

$x, y : [a, b] \to \mathbb{R}$ så att

$$f(x(t), y(t)) = c,$$

med andra ord "färdas" stigen γ på nivåytan $f(x, y) = c$. Vi antar dessutom, att $x, y \in \mathcal{C}^1([a, b])$.

2.7.4. Anmärkning. Om $(x'(t_0), y'(t_0)) \neq (0, 0)$, $t_0 \in [a, b]$, så säger vi, att stigen γ är tangenten i punkten t_0. Om $y'(t_0) \neq 0$, så är dess ekvation

$$y - y(t_0) = \frac{x'(t_0)}{y'(t_0)}(x - x(t_0)).$$

Vi återkommer till detta senare.

2.7.5. Sats (Gradientens "geometri"). *Låt γ och f vara avbildningar såsom tidigare. Då är*

$$\nabla f(\gamma(t)) \cdot \gamma'(t) = 0, \ t \in [a, b]$$

alltså

$$\nabla f(x(t), y(t)) \cdot (x'(t), y'(t)) = 0, \ t \in [a, b].$$

Med andra ord står ∇f *vinkelrätt mot varje* $\mathcal{C}^1$*-stig som färdas på nivåytan.*

Bevis. Enligt antagandet är $f(x(t), y(t)) = c$ för varje $t \in [a, b]$. Vi använder kedjeregel C) och undersöker funktionen h

$$t \mapsto (f \circ \gamma)(t) = f(x(t), y(t)).$$

Denna är en avbildning från intervallet $[a, b]$ till de reella talen. Eftersom γ innehålls i nivåytan, är denna funktion konstanten c i intervallet $[a, b]$, av vilket det följer att $h'(t) = (f \circ \gamma)'(t) = 0, \ t \in [a, b]$. Vi bestämmer nu $\nabla f(\gamma(t)) \cdot \gamma'(t)$. Kedjeregel C) get att

$$0 = h'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t),$$

vilket är påståendet. □

Som följd av satserna 2.7.1 och 2.7.5 får vi att gradienten ∇f för funktionen f har följande egenskap: Gradienten ∇f står vinkelrätt mot f:s nivåkurvor (nivåytor, då $n \geq 3$) och visar riktningen, i vilken f växer kraftigast.

2.8. Partiella derivator av högre grad och Taylors formel

Antag att $D \subset \mathbb{R}^2$ är en öppen mängd. Vi undersöker funktionen $f : D \to \mathbb{R}$. Generaliseringar till fallet $n \geq 3$ är enkla. Om $\partial_1 f(x)$ existerar för varje punkt i mängden D, får vi funktionen $\partial_1 f : D \to \mathbb{R}$. Nu kan vi försöka derivera med avseende på x eller y, alltså bilda de *partiella dervatorna av andra graden* $\partial_1\partial_1 f$, betecknas $\partial_1^2 f = f_{xx}$, eller $\partial_2\partial_1 f = f_{xy}$. På motsvarade sätt kan vi bilda $\partial_2\partial_2 f$, betecknas $\partial_2^2 f$ eller $\partial_1\partial_2 f = f_{yx}$. Dessa är partiella derivator av andra graden: i alla ovannämnda fall har funktionen f deriverats två gånger. Genom att fortsätta på motsvarande sätt får vi $\partial_1^n f$, $\partial_1\partial_2\partial_1 f$ och så vidare, ifall de ifrågavarande derivatorna existerar. Derivatan $\partial_1^n f$ kallas en partiell derivata av n:te graden och exempelvis $\partial_1\partial_2\partial_1 f$ en partiell derivata av tredje graden. Vi använder beteckningen $f \in \mathcal{C}^p(D)$, med vilken vi avser, att f och alla dess partiella derivator av grad p eller under existerar och är kontinuerliga i D. Då är

$$\mathcal{C}^0(D) = \mathcal{C}(D) = \{f : D \to \mathbb{R} : f \text{ kontinuerlig}\}$$

och $\mathcal{C}^1(D)$ är redan bekant som klassen för en gång kontinuerligt deriverbara funktioner.

2.8.1. Exempel. Låt $f(x, y) = \sin(xy),\ (x, y) \in \mathbb{R}^2$. Nu är

$$\begin{aligned}
\partial_1 f(x, y) &= y\cos(xy), \\
\partial_2 f(x, y) &= x\cos(xy), \\
\partial_2\partial_1 f(x, y) &= \cos(xy) - yx\sin(xy), \\
\partial_1\partial_1 f(x, y) &= -y^2\sin(xy), \\
\partial_1\partial_2 f(x, y) &= \cos(xy) - xy\sin(xy), \\
\partial_2\partial_2 f(x, y) &= -x^2\sin(xy), \\
&\vdots
\end{aligned}$$

Vi upptäcker, att $\partial_1\partial_2 f(x, y) = \partial_2\partial_1 f(x, y)$. Det visar sig, att detta inte är en tillfällighet. Funktionen f är sådan, att den har kontinuerliga partiella derivator av alla grader i $\mathbb{R}^2$. Mängden av sådana funktioner betecknas $\mathcal{C}^\infty(\mathbb{R}^2)$. Generellt betecknar vi $\mathcal{C}^\infty(D)$, om funktionen f har kontinuerliga partiella derivator av alla grader i den öppna mängden $D \subset \mathbb{R}^n$.

2.8.2. Exempel. Antag att $f : \mathbb{R}^2 \to \mathbb{R},\ f(x) = |x|$, alltså att $f(x, y) = \sqrt{x^2 + y^2}$. Nu är $f \in \mathcal{C}(\mathbb{R}^2)$, men $f \notin \mathcal{C}^1(\mathbb{R}^2)$, ty f har inte partiella derivator i punkten 0. Det gäller dock att $f \in \mathcal{C}^\infty(\mathbb{R}^2 \setminus \{0\})$.

2.8.3. Sats. *Låt $f \in \mathcal{C}^2(D)$. Då gäller att $\partial_1\partial_2 f = \partial_2\partial_1 f$.*

Bevis. Idén är att använda den vanliga medelvärdessatsen. Vi skall visa, att då $(x, y) \in D$, så är

$$\partial_1\partial_2 f(x, y) = \partial_2\partial_1 f(x, y).$$

Vi fixerar en (x, y)-centrerad kvadrat Q, som i sin helhet inkluderas i D. Detta är möjligt, eftersom D är öppen. Vi undersöker nu punkterna (x, y), $(x+h, y)$, $(x, y+k)$, $(x+h, y+k)$, där $h,\ k \neq 0$ är så små, att dessa punkter är innanför Q. Vi definierar

$$\begin{aligned}\varphi(t) &= f(x+h, t) - f(x, t),\\ \psi(s) &= f(s, y+k) - f(s, y).\end{aligned}$$

Då är φ definierad i ett y-centrerat intervall och ψ i ett x-centrerat intervall. Vi upptäcker, att

$$\begin{aligned}q(h, k) &= \varphi(y+k) - \varphi(y)\\ &= f(x+h, y+k) - f(x, y+k) - f(x+h, y) + f(x, y)\\ &= \psi(x+h) - \psi(x).\end{aligned}$$

Genom att derivera funktionen φ och tillämpa medelvärdessatsen får vi att

$$\begin{aligned}\varphi'(t) &= \partial_2 f(x+h, t) - \partial_2 f(x, t),\\ q(h, k) &= \varphi(y+k) - \varphi(y) = \varphi'(\xi)k\\ &= [\partial_2 f(x+h, y+\theta k) - \partial_2 f(x, y+\theta k)]k,\end{aligned}$$

där ξ är mellan punkterna y och $y+k$, d.v.s. $\xi = y + \theta k$, $\theta \in (0, 1)$. Sedan tillämpar vi medelvärdessatsen på funktionen $s \mapsto \partial_2 f(s, y+\theta k)$. Detta ger att

$$q(h, k) = \partial_1 \partial_2 f(x + \eta h, y + \theta k)hk,$$

där $\eta \in (0, 1)$. Genom att låta $(h, k) \to 0$ ser vi, att det för varje tal $h, k \neq 0$ gäller att

$$\frac{q(h, k)}{hk} \to \partial_1 \partial_2 f(x, y),$$

eftersom $\partial_1 \partial_2 f$ är kontinuerlig i punkten (x, y). Genom att utgå från uttrycket

$$q(h, k) = \psi(x+h) - \psi(x)$$

kommer vi med en liknande räkneoperation till resultatet

$$\frac{q(h, k)}{hk} \to \partial_2 \partial_1 f(x, y), \text{ då } (h, k) \to (0, 0).$$

Av detta följer att $\partial_2 \partial_1 f(x, y) = \partial_1 \partial_2 f(x, y)$. $\square$

En motsvarande sats till Sats 2.8.3 gäller för alla $n \geq 2$. I själva verket är beviset likadant, eftersom satsens påstående är till sin grund tvådimensionellt. Satsens antaganden är aningen för starka, men Sats 2.7.5 används allmänt taget i den ovannämnda formen.

Vi utvecklar nu *Taylors formel* i flera dimensioner. Vi undersöker endast den i tillämpningar viktiga Taylors formel av andra graden i fallet $n = 2$. Vi strävar inte ens efter att minimera antagandena.

2.8.4. Sats (Taylorutveckling av andra graden)**.** *Antag att $D \subset \mathbb{R}^2$ är öppen, $f \in \mathcal{C}^3(D)$ och $(x_0, y_0) \in D$. Om $(x_0 + h, y_0 + k) \in D$ så är*

$$\begin{aligned} f(x_0 + h, y_0 + k) &= f(x_0, y_0) + \partial_1 f(x_0, y_0)\, h + \partial_2 f(x_0, y_0)\, k \\ &\quad + 1/2\, (\partial_1\partial_1 f(x_0, y_0)\, h^2 + 2\, \partial_1\partial_2 f(x_0, y_0)\, hk + \partial_2\partial_2 f(x_0, y_0)\, k^2) \\ &\quad + (h^2 + k^2)^{3/2} B(h, k), \end{aligned}$$

där $B(h, k)$ är en begränsad funktion i något klot $B(0, r),\ r > 0$.

2.8.5. Anmärkning. Fallet $n = 1$: Låt $f : [a, b] \to \mathbb{R}$, $x_0 \in [a, b]$ och $f \in \mathcal{C}^3[a, b]$. Då är

$$f(x_0 + h) = f(x_0) + f'(x_0)\, h + \frac{1}{2!}\, f''(x_0)\, h^2 + \frac{1}{3!}\, h^3\, f'''(x_0 + \theta h)\, h^3,$$

varvid $x_0 + h \in [a, b]$ och $\theta \in [0, 1]$. Detta är den vanliga Taylorutvecklingen av andra graden för funktionen f, där den tredje derivatan av f har använts som "restterm".

Bevis för Sats 2.8.4. Såsom vi upptäckte då vi undersökte derivatan av en funktion, gäller den givna formeln alltid, då $(x_0 + h, y_0 + k) \in D$. Problemet är begränsning av funktionen B i något klot $B(0, r)$. För att bevisa detta fixerar vi h och k och undersöker funktionen

$$F(t) = f(x_0 + th, y_0 + tk),\ 0 \leq t \leq 1.$$

Denna är definierad i intervallet $[0, 1]$ då h och k har valts tillräckligt små. Efter detta är beviset endimensionellt. Nu är $F \in \mathcal{C}^3([0, 1])$ och den endimensionella formeln ger att

$$F(t) = F(0) + F'(0)t + \frac{F''(0)}{2!}t^2 + \frac{F'''(\theta t)}{3!}t^3,\ 0 \leq \theta \leq t. \tag{2.8.6}$$

Vi tillämpar kedjeregel C), sida 30, på funktionen $F = f \circ \gamma$, där

$$\gamma(t) = (x_0 + th, y_0 + tk).$$

Detta ger att

$$\begin{aligned} F'(t) &= \nabla f(\gamma(t)) \cdot \gamma'(t) \\ &= \partial_1 f(x_0 + th, y_0 + tk)\, h + \partial_2 f(x_0 + th, y_0 + tk)\, k. \end{aligned}$$

Genom att tillämpa denna på nytt får vi att

$$\begin{aligned} F''(t) = \partial_{11} f(x_0 + th, y_0 + tk)\, h^2 &+ 2\, \partial_{12} f(x_0 + th, y_0 + tk)\, hk \\ &+ \partial_{22} f(x_0 + th, y_0 + tk)\, k^2, \end{aligned}$$

där vi i den andra termen har tillämpat bytesregeln för partiella derivator. Genom att substituera dessa i formel (2.8.6) där $t = 1$, får vi att

$$f(x_0 + h, y_0 + k) = F(0) + F'(0) + \frac{1}{2} F''(0) + \frac{1}{6} F'''(\theta),$$

där

$$\begin{aligned} F(0) &= f(x_0, y_0), \\ F'(0) &= \partial_1 f(x_0, y_0)\, h + \partial_2 f(x_0, y_0)\, k, \\ F''(0) &= \partial_{11} f(x_0, y_0)\, h^2 + 2\, \partial_{12} f(x_0, y_0)\, hk + \partial_{22} f(x_0, y_0)\, k^2. \end{aligned}$$

Vi skall ännu bevisa den väsentliga delen av beviset, nämligen att

$$B(h,k) = (h^2 + k^2)^{-3/2} F'''(\theta)/6 \tag{2.8.7}$$

är i den krävda formen. Märk att vi kan definiera $B(0,0)$ att vara noll. Om F deriveras tre gånger, så är den en ändlig summa av f:s partiella derivator av tredje graden, multiplicerade med talen h och k. Summans termer är typiskt i formen

$$\partial_{112} f(x_0 + th, y_0 + tk)\, h^2 k.$$

Genom att susbtituera $t = \theta$, använda estimaten $|h| \le \sqrt{h^2 + k^2}$, $|k| \le \sqrt{h^2 + k^2}$ och upptäcka att de partiella derivatorna av tredje graden är kontinuerliga och därmed begränsade i en omgivning till punkten (x_0, y_0), ser vi ur formel (2.8.7), att $B(h,k)$ är begränsad i en omgivning till origo. Observera att till exempel

$$|(h^2+k^2)^{-3/2}h^2k| \le \frac{|x|^2|x|}{|x|^3} \le 1,$$

då vi betecknar $x = (h,k)$. □

2.8.8. Anmärkning. Då $n \ge 1$ får uttrycket i Sats 2.8.4 följande form: Låt $D \subset \mathbb{R}^n$ vara en öppen mängd, $f \in \mathcal{C}^3(D)$, $x_0 \in D$, $h \in \mathbb{R}^n$ och $x_0 + h \in D$. Då är

$$f(x_0+h) = f(x_0) + \sum_{i=1}^n \partial_i f(x_0)\, h_i + \frac{1}{2}\sum_{i,j=1}^n \partial_{ij} f(x_0)\, h_i h_j + |h|^3\, B(h),$$

där B är begränsad i en omgivning till origo.

För funktioner i $\mathbb{R}^n$ finns även utvecklingar av högre grader. De erhålls på samma sätt som ovan. De ovan härledda utvecklingarna av andra graden räcker dock i många tillämpningar.

2.8.9. Exempel. Vi söker Taylorutvecklingen av andra graden för funktionen

$$f(x,y) = \sin(x^2+y^2)$$

i punkten $(0,0) = 0$. Det är uppenbart att $f \in \mathcal{C}^3(\mathbb{R}^2)$, och således är

$$\begin{aligned}
f(0,0) &= 0,\\
\partial_1 f(x,y) &= 2x\cos(x^2+y^2)\\
\partial_2 f(x,y) &= 2y\cos(x^2+y^2)\\
\partial_{11} f(x,y) &= 2\cos(x^2+y^2) - 4x^2\sin(x^2+y^2)\\
\partial_{22} f(x,y) &= 2\cos(x^2+y^2) - 4y^2\sin(x^2+y^2)\\
\partial_{12} f(x,y) &= -4xy\sin(x^2+y^2).
\end{aligned}$$

Därmed gäller att

$$\partial_1 f(0,0) = \partial_2 f(0,0) = \partial_{12} f(0,0) = \partial_{21} f(0,0) = 0$$

och $\partial_1^2 f(0,0) = \partial_2^2 f(0,0) = 2$. Vi får alltså att

$$\sin(h^2+k^2) = 1/2\,(2h^2+2k^2) + (h^2+k^2)^{3/2}\,B(h,k),$$

där $B(h,k)$ är begränsad i en omgivning till 0. Derivata och Taylorutveckling används för att lokalt approximera funktioner. Om vi vill approximera funktionen f nära punkten 0 så lönar det sig, beroende på approximationens noggrannhetskrav att gå till väga på följande sätt. Först lönar det sig att nöja sig med funktionen $g_0(x,y) \equiv 0$, ty f är kontinuerlig i 0 och $f(0) = 0$. För noggrannare approximering lönar det sig att välja

$$(x,y) \mapsto x^2 + y^2.$$

som approximerande funktion. Taylorutvecklingen av andra graden kan även finnas på ett annat sätt, ty

$$\sin t = t + B(t)\, t^3,$$

där $B(t)$ är begränsad. Detta är den vanliga endimensionella Taylors formel. Genom att här substituera $t = h^2 + k^2$ får vi att

$$\sin(h^2 + k^2) = h^2 + k^2 + B(h^2 + k^2)(h^2 + k^2)^3.$$

Detta är en noggrannare approximation än tidigare med avseende på resttermen.

2.9. EXTREMVÄRDEN

Låt $A \subset \mathbb{R}^n$ och $f : A \to \mathbb{R}$. Ofta antar vi att A är en öppen delmängd i $\mathbb{R}^n$.

2.9.1. Definition. Funktionen f har ett *lokalt maximum* i punkten $x_0 \in A$, om det existerar ett sådant $\delta > 0$, att

$$f(x) \leq f(x_0), \text{ då } x \in A \cap B(x_0, \delta). \tag{2.9.2}$$

I punkten x_0 finns ett *strängt* lokalt maximum, om det i formel (2.9.2) gäller sträng olikhet då $x \neq x_0$. Vi definierar (strängt) lokalt minimum på motsvarande sätt.

2.9.3. Anmärkning. I det föregående finns en liten skillnad från definitionen given i fallet $n = 1$. Låt till exempel $A = [0,1] \subset \mathbb{R}$ och $f(x) = x$. Enligt definitionen ovan är $x_0 = 0$ ett strängt lokalt minimum för f och $x_0 = 1$ ett strängt lokalt maximi för f. Generellt kan lokala maxima och minima för en funktion definierad i intervallet Δ på den reella axeln endast finnas i innerpunkter i Δ. Ibland används detta även då $n \geq 2$.

Vi säger att punkten x_0 i mängden $A \subset \mathbb{R}^n$ är en *innerpunkt*, i A, om det existerar ett sådant $\delta > 0$, att $B(x_0, \delta) \subset A$. Mängden av innerpunkter i mängden A betecknas $\mathrm{int}A$.

Lokala maximi- och minimipunkter kallas även *lokala extremvärden*.

2.9.4. Exempel. Om A är öppen, så är alla punkter i A innerpunkter i A. Om $A = [0, 1] \subset \mathbb{R}$, så är 0 och 1 inte innerpunkter i A.

2.9.5. Sats (grundsats). *Antag att $f : A \to \mathbb{R}$ och att x_0 är en innerpunkt i mängden A och en lokal extrempunkt för funktionen f. Om $\partial_i f(x_0)$ existerar för något $i = 1, \ldots, n$, så är $\partial_i f(x_0) = 0$.*

Bevis. Satsen följer av den motsvarande satsen i fallet $n = 1$. Vi granskar funktionen

$$t \overset{\varphi}{\mapsto} f(x_{0,1}, \ldots, x_{0,i-1}, t, x_{0,i+1}, \ldots, x_{0,n}),$$

$x_0 = (x_{0,1}, \ldots, x_{0,n})$. Eftersom x_0 är en innerpunkt i A, är φ definierad i åtminstone något intervall $(x_{0,i} - \delta, x_{0,i} + \delta) = \Delta,\ \delta > 0$. Eftersom f har ett lokalt maximum i punkten x_0, så gäller det genom att välja ett tillräckligt litet δ att

$$\varphi(t) \leq \varphi(x_{0,i}),$$

då $t \in \Delta$. Således har funktionen φ ett maximum i punkten $x_{0,i}$.

Om funktionen $\varphi : \Delta \to \mathbb{R}$ har en derivata i innerpunkten $x_{0,i}$ i intervallet Δ, så gäller det som känt i de lokala extremvärdespunkterna att $\varphi'(x_{0,i}) = 0$. Å andra sidan ger partiella derivatans definition att $\varphi'(x_{0,i}) = \partial_i f(x_0)$, och således är $\partial_i f(x_0) = 0$. På motsvarande sätt visar vi, att $\partial_i f(x_0) = 0$, ifall f har ett lokalt minimum i punkten x_0. $\square$

2.9.6. Definition. Låt $f \in \mathcal{C}^1(D),\ D \subset \mathbb{R}^n$ öppen. Punkter $x \in D$, för vilka

$$\nabla f(x) = (\partial_1 f(x), \ldots, \partial_n f(x)) = (0, \ldots, 0) = 0,$$

kallas *kritiska punkter* till f. I dessa punkter erhållna värden kallas *kritiska värden* av f.

Av Sats 2.9.5 följer: Låt $D \subset \mathbb{R}^n$ vara öppen och $f \in \mathcal{C}^1(D)$. Då finns de lokala extrempunkterna till f i mängden av de kritiska punkterna till f.

2.9.7. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = (4 - x^2 - y^2)\, e^{x+y}$. Vi skall bestämma de kritiska punkterna för funktionen f. Som partiella derivator har vi

$$\begin{aligned} \partial_1 f(x,y) &= e^{x+y}(-2x + 4 - x^2 - y^2), \\ \partial_2 f(x,y) &= e^{x+y}(-2y + 4 - x^2 - y^2). \end{aligned}$$

Vi bestämmer de punkter (x, y), i vilka

$$\nabla f(x,y) = (\partial_1 f(x,y), \partial_2 f(x,y)) = (0,0).$$

Eftersom $e^{x+y} > 0$ för alla (x, y), så uppfyller dessa punkter ekvationssystemet

$$\begin{cases} -2x + 4 - x^2 - y^2 = 0, \\ -2y + 4 - x^2 - y^2 = 0. \end{cases}$$

Genom att subtrahera ekvationerna med varandra får vi $x = y$ och genom att substituera detta i den första exkvationen följer, att $x = -2$ eller $x = 1$. Därmed är punkterna $(-2, -2)$ och $(1, 1)$ de enda möjliga kritiska punkterna till f. Vi observerar omedelbart, att dessa två punkter verkligen uppfyller $\nabla f(x, y) = (0, 0)$.

Svårare är å sin sida att bestämma, ifall de kritiska punkterna är lokala extrempunkter för f.

2.9.8. Anmärkning. I fallet $n = 1$ erbjuder den andra derivatan av funktionen f ett tillräckligt villkor för lokala extremvärden. Låt $f \in \mathcal{C}^2(a, b)$. Om $f'(x_0) = 0$ och $f''(x_0) > 0$, så finns det ett strängt lokalt maximum i punkten x_0 och på motsvarande sätt om $f'(x_0) = 0$ och $f''(x_0) < 0$, så finns det ett strängt lokalt minimum i punkten x_0.

Vi undersöker fallet $n = 2$ och strävar efter att hitta ett motsvarande kriterium för lokala extrempunkter. Antag att $f \in \mathcal{C}^3(D)$ och att $(x_0, y_0) \in D$ är en kritisk punkt för f, alltså att $\nabla f(x_0, y_0) = 0$. Av Taylors formel, Sats 2.8.4, följer att:

$$\begin{aligned} f(x_0 + h, y_0 + k) &= f(x_0, y_0) \\ &+ 1/2\,\underbrace{\left(\partial_1\partial_1 f(x_0,y_0)\,h^2 + 2\,\partial_1\partial_2 f(x_0,y_0)\,hk + \partial_2\partial_2 f(x_0,y_0)\,k^2\right)}_{\text{motsvarar } f''(x_0)} \\ &+ (h^2+k^2)^{3/2} B(h,k). \end{aligned}$$

Detta leder till att undersöka hur avbildningen $Q : \mathbb{R}^2 \to \mathbb{R}$ i formen

$$Q(h, k) = ah^2 + 2ahk + ck^2, \; a, b, c \in \mathbb{R},$$

uppför sig. Funktionen Q kallas en *kvadratisk form av andra graden.*

Beroende på värdena på konstanterna a, b, c beter sig Q på olika sätt. De fyra följande fallen är möjliga, och genom att välja

$$\begin{aligned} a &= \partial_{11} f(x_0, y_0), \\ b &= \partial_{12} f(x_0, y_0), \\ c &= \partial_{22} f(x_0, y_0) \end{aligned}$$

kan vi dra följande slutsatser.

(i) $Q(h, k)$ är *positivt definit*, alltså $Q(h, k) > 0$ då $(h, k) \neq (0, 0)$. Då är (x_0, y_0) ett strängt lokalt minimum.

(ii) $Q(h, k)$ är *negativt definit*, alltså $Q(h, k) < 0$ då $(h, k) \neq (0, 0)$. Då är (x_0, y_0) ett strängt lokalt maximum.

(iii) $Q(h, k)$ är *indefinit*, med andra ord får den både positiva och negativa värden. Då är (x_0, y_0) inte en extrempunkt, utan en så kallad *terrasspunkt.*

(iv) $Q(h, k)$ är *positivt (negativt) semidefinit*, alltså

$$Q(h, k) \geq 0 \; (Q(h, k) \leq 0)$$

för alla (h, k) och $Q(h, k) = 0$ för något $(h, k) \neq 0$. På basen av detta kan man inte dra några slutsater om extremvärdeskaraktären i punkten (x_0, y_0), utan situationen måste undersökas skilt.

I själva verket visar det sig, att antagandet $f \in \mathcal{C}^2(D)$ räcker för slutsatserna ovan. I tillämpningar gäller dock vanligen att $f \in \mathcal{C}^3(D)$.

2.9.9. Anmärkning. Fallet (iii) förekommer inte då $n = 1$, ty $Q(h) = f''(x_0)\, h^2$.

2.9.10. Exempel. (i) $Q(h,k) = h^2 + k^2$ är positivt definit.

(ii) $Q(h,k) = -(h^2 + k^2)$ är negativt definit.

(iii) $Q(h,k) = hk$ är indefinit, samma sak gäller då $Q(h,k) = h^2 - k^2$.

(iv) $Q(h,k) = h^2$ är positivt semidefinit, eftersom $Q(h,k) = 0$ då $(h,k) = (0,k)$.

Vi lämnar bevisen för dessa som övningsuppgift. Beviset är liknande som i fallet $n = 1$ och grundar sig på Taylors formel.

2.9.11. Anmärkning. I fallet $n \geq 3$ måste vi undersöka den kvadratiska formen

$$Q(h) = \sum_{i,j=1}^{n} a_{ij}\, h_i\, h_j,\ \ h = (h_1, \ldots, h_n),\ \ a_{ij} = \partial_i \partial_j f(x_0).$$

I övrigt är situationen fullständigt motsvarande. Funktionen Q kallas *Hess' kvadratiska form i* x_0 för f.

Vi undersöker härnäst när $Q(h,k)$ är av någon av typerna (i) - (iv). Låt Q alltjämt vara $Q(h,k) = ah^2 + 2bhk + ck^2$.

(i) Q är positivt definit exakt när $a > 0$ och determinanten

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} = ac - b^2 > 0. \tag{2.9.12}$$

(ii) Q är negativt definit exakt när $a < 0$ och olikheten (2.9.12) gäller.

(iii) Q är indefinit exakt när

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} < 0.$$

(iv) Q är semidefinit exakt när

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} = 0.$$

Att bevisa dessa kriterier är inte svårt; i själva verket hänför de sig till karakterisering av andra gradens ytor. Motsvarande, aningen mera komplicerade villkor gäller då $n \geq 3$.

2.9.13. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = 2x^3 - 6xy + 3y^2$. Vi bestämmer de kritiska punkterna och lokala extrempunkerna för f. Funktionen f tillhör mängden $\mathcal{C}^1(\mathbb{R}^2)$, så alla lokala extrempunkter är bland de kritiska punkterna. De kritiska punkterna (x,y) uppfyller ekvationssystemet

$$\begin{cases} \partial_1 f(x,y) = 6x^2 - 6y = 0, \\ \partial_2 f(x,y) = -6x + 6y = 0, \end{cases}$$

ur vilket vi med en enkel härledning får att $x = 0$ och $y = 0$ eller $x = 1$ och $y = 1$. De enda kritiska punkterna är således $(0,0)$ och $(1,1)$.

Härnäst skall vi undersöka ifall dessa punkter är lokala extrempunkter. Eftersom $f \in \mathcal{C}^3(\mathbb{R}^2)$ ($f \in \mathcal{C}^2(\mathbb{R}^2)$ räcker), kan vi använda den ovannämnda teorin. Nu är derivatorna $\partial_1\partial_1 f(x,y) = 12x,\ \partial_1\partial_2 f(x,y) = -6$ och $\partial_2\partial_2 f(x,y) = 6$. I punkten $(0,0)$ är $a = 0,\ b = -6,\ c = 6$, så

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} = -36.$$

Därmed är punkten $(0,0)$ en terrasspunkt, och således inte en lokal extrempunkt. I punkten $(1,1)$ är $a = 12 > 0$, $b = -6$, $c = 6$, så

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0.$$

Enligt punkt (i) är alltså Q positivt definit, av vilket det följer, att punkten $(1,1)$ är en (sträng) lokal minimipunkt.

2.9.14. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = xy\, e^{-(x^2+y^2)/2}$. Funktionens partiella derivator är

$$\begin{aligned} \partial_1 f(x,y) &= y\,(1-x^2)\, e^{-(x^2+y^2)/2}, \\ \partial_2 f(x,y) &= x\,(1-y^2)\, e^{-(x^2+y^2)/2}, \\ \partial_1\partial_1 f(x,y) &= xy\,(x^2-3)\, e^{-(x^2+y^2)/2}, \\ \partial_1\partial_2 f(x,y) &= (1-x^2)\,(1-y^2)\, e^{-(x^2+y^2)/2}, \\ \partial_2\partial_2 f(x,y) &= xy\,(y^2-3)\, e^{-(x^2+y^2)/2}, \end{aligned}$$

och eftersom $f \in \mathcal{C}^3(\mathbb{R}^2)$, ger de första partiella derivatorna att funktionens kritiska punkter är $(0,0)$, $(1,1)$, $(1,-1)$, $(-1,1)$ och $(-1,-1)$. Genom att substituera punkten $(0,0)$ får vi att $a = c = 0$ och $b = 1$, d.v.s. $(0,0)$ är en terrasspunkt. Genom att substituera punkterna $(1,1)$ och $(-1,-1)$ får vi att $a = c = -2/e < 0$ och $b = 0$, d.v.s.

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0.$$

Dessa är därmed stränga lokala maximipunkter. I punkterna $(1,-1)$ och $(-1,1)$ får vi att $a = c = 2/e > 0$ och $b = 0$, d.v.s.

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0.$$

Dessa är alltså stränga lokala minimipunkter.

Vi behandlar härnäst *absoluta d.v.s. globala extremvärden* för reellvärda funktioner. Låt $f : A \to \mathbb{R},\ A \subset \mathbb{R}^n$. Satserna nedan kompletterar de gamla bekanta kunskaperna om beteendet hos funktioner definierade i intervallet $[a,b]$.

2.9.15. Definition. Funktionen f har ett *(absolut) minimum* i punkten $x_0 \in A$, om

$$f(x) \geq f(x_0) \text{ för alla } x \in A. \tag{2.9.16}$$

I punkten x_0 finns ett *strängt* absolut minimum, ifall det gäller sträng olikhet i formel (2.9.16). På motsvarande sätt definierar vi *(strängt) absolut maximum.*

2.9.17. Sats. *Antag att $A \subset \mathbb{R}^n$ är en kompakt, icke-tom mängd och att $f : A \to \mathbb{R}$ är kontinuerlig. Då har f åtmistone en absolut minimi- och maximipunkt.*

Bevis. Vi visar först, att f är begränsad, d.v.s. att det existerar ett sådant $M < \infty$, att

$$|f(x)| \leq M \text{ för alla } x \in A.$$

Vi gör ett motantagande: f är inte begränsad. Då finns det sådana punkter $x_i \in A,\ i = 1,2,\ldots$, att

$$|f(x_i)| \geq i.$$

Eftersom A är kompakt, finns det en sådan delmängd x_{i_j} $(x_i \in A)$, att $x_{i_j} \to x_0$, då $j \to \infty$ och $x_0 \in A$. Nu är $|f(x)|$ kontinuerlig, och därmed är

$$\underbrace{|f(x_{i_j})|}_{\to\infty} \to |f(x_0)| < \infty,$$

vilket är en motsägelse.

Vi visar sedan, att f har t.ex. en maximipunkt. Eftersom f är begränsad, så existerar $\sup\{f(x) : x \in A\} = M_0$. Av supremums definition följer, att det existerar en sådan följd (x_i), att $x_i \in A$, och $f(x_i) \to M_0$. Såsom tidigare följer av mängden A:s kompakthet, att det existerar en delföljd (x_{i_j}) till följden (x_i) samt ett $x_0 \in A$, för vilket det gäller att $x_{i_j} \to x_0$. Av kontinuiteten för f följer att $f(x_0) = M_0$. $\square$

2.9.18. Exempel. Vi undersöker *minsta kvadratmetoden*, som ofta används i statistiken. Vi betraktar en serie observationer, där det gjorts flera mätningar:

$$x = (x_1, x_2, \dots, x_n),$$

där exempelvis värdet x_1 är temperaturen av x, x_2 är diametern av x och så vidare. Antag sedan att $x^1 = (x_1^1, \dots, x_n^1)$ är den första mätningen, $x^2 = (x_1^2, \dots, x_n^2)$ den andra mätningen och $x^k = (x_1^k, \dots, x_n^k)$ den sista, d.v.s. den k:te mätningen.

Vi antar, att mätningarna gjorts under samma förhållanden, men att de till exempel är utförda av olika personer. Vi försöker utreda det "bästa" estimatet för x. Låt

$$x_0 = \frac{1}{k}\sum_{i=1}^{k} x^1 = \left(\frac{1}{k}\sum_i x_1^i, \dots, \frac{1}{k}\sum_i x_n^i\right),$$

vara medeltalet av observationerna x^j, $j = 1, \dots, k$. Märk, att x_0 är en vektor i $\mathbb{R}^n$. Vi strävar efter att finna den punkt $x \in \mathbb{R}^n$, som minimerar uttrycket

$$|x - x^1|^2 + \ldots + |x - x^k|^2.$$

Vi visar, att medelvärdet $x_0 \in \mathbb{R}^n$ har denna egenskap, alltså att funktionen $f : \mathbb{R}^n \to \mathbb{R}$,

$$f(x) = |x - x^1|^2 + \ldots + |x - x^k|^2,$$

har ett strängt absolut minimum i punkten x_0.

Genom att skriva uttrycket för f i koordinatform $x = (x_1, \ldots, x_n)$ får vi att

$$f(x) = \sum_{j=1}^{k} [(x_1 - x_1^j)^2 + \ldots + (x_n - x_n^j)^2].$$

Om funktionen f har ett strängt absolut minimum i punkten x, så är x en kritisk punkt till funktionen f, ty $f \in \mathcal{C}^\infty(\mathbb{R}^n)$. Därmed gäller att

$$\partial_1 f = 2 \sum_{j=1}^{k} (x_1 - x_1^j) = 0$$

$$\vdots$$

$$\partial_n f = 2 \sum_{j=1}^{k} (x_n - x_n^j) = 0.$$

Av dessa ekvationer följer att

$$x_i = \frac{1}{k} \sum_{j=1}^{k} x_i^j, \; i = 1, \ldots, n,$$

så det stämmer faktiskt att $x = x_0$. Vi skall ännu visa, att x_0 är ett väsentligt absolut minimum för f. Vi härleder detta direkt. För det första är $f(x) \geq 0$ för varje $x \in \mathbb{R}^n$ och det gäller att

$$\lim_{|x| \to \infty} f(x) = \infty.$$

Detta är enkelt att se, ty då $|x|$ är stor, är också något x_i stort och $f(x) \geq (x_i - x_i^1)^2$, så då är även $f(x)$ stor. Således är $f(x) \geq 1 + f(x_0)$, då $|x| \geq M$ för något $M > 0$. Mängden $\overline{B}(0, M)$ är kompakt, så f får stt absoluta minimivärde där. Det är ett absolut minimum i hela $\mathbb{R}^n$, eftersom $f(x) \geq 1 + f(x_0)$ är utanför klotet och f får värdet $f(x_0)$ i klotet $\overline{B}(0, M)$, och detta är mindre än $1 + f(x_0)$. I den absoluta minimipunkten är de partiella derivatornas värden 0. Den enda dylika punkt är x_0, alltså får funktionen f sitt absoluta minimum i punkten x_0.

Vi skall ännu visa, att punkten x_0 ger ett strängt absolut minimum. Detta ser man enkelt med ett motantagande: Om $x \in \mathbb{R}^n$ är sådan, att $f(x) = f(x_0)$, så uppnår f sitt absoluta minimum även i punkten x.

Därmed är x en kritisk punkt till f, i vilken det oundvikligen gäller att $\nabla f(x) = 0$. Men x_0 är den enda dylika punkten, så $x = x_0$.

Det, att f uppnår ett strängt lokalt minimum i x_0 kan vi även undersöka genom att använda Hess' kvadratiska form i punkten x_0, varvid vi observerar att den är positivt definit. Det är även möjligt att med direkt räkning visa, att $f(x) > f(x_0)$ för alla $x \neq x_0$.

3. Vektorvärda funktioner

3.1. Stigar

Låt Δ vara vilket som helst intervall på den reella axeln $\mathbb{R}$. Avbildningen $\gamma : \Delta \to \mathbb{R}^n$ är en *stig*, om den är kontinuerlig. Vi antar nu, att $\Delta = [a, b]$. För varje $t \in [a, b]$ är $\gamma(t)$ en vektor i $\mathbb{R}^n$

$$\gamma(t) = (\gamma_1(t), \ldots, \gamma_n(t)),$$

och därmed definierar γ avbildningarna $\gamma_i : [a, b] \to \mathbb{R},\ i = 1, \ldots, n$. Dessa kallas γ:s *koordinatfunktioner*. Det, att γ är kontinuerlig, betyder att varje γ_i är kontinuerlig.

3.1.1. Definition. Vi säger att $\gamma \in \mathcal{C}^1([a, b])$ d.v.s. att γ är en $\mathcal{C}^1$-funktion, om $\gamma_i \in \mathcal{C}^1([a, b])$, $i = 1, 2, \ldots, n$. Derivatan γ' av avbildningen γ i punkten $t \in [a, b]$ definieras med formeln

$$\gamma'(t) = (\gamma_1'(t), \ldots, \gamma_n'(t)) = \lim_{h \to 0} \frac{\gamma(t + h) - \gamma(t)}{h}.$$

Derivatan $\gamma'(t)$ är alltså också en vektor i $\mathbb{R}^n$. Den ovan definierade derivatan är inte ett nytt begrepp. Den motsvarar fallet $p = 1$ i kapitel 2.6. Den enda skillnaden är, att definitionsmängden är ett slutet intervall $[a, b]$, varvid det är nyttigt att även undersöka (ensidig) deriverbarhet i intervallets ändpunkter. I själva verket är $\gamma'(t)$ en lineärfunktion $\gamma' : \mathbb{R} \to \mathbb{R}^n$, som definieras med formeln

$$\gamma'(t)h = (h\gamma_1'(t), \ldots, h\gamma_n'(t)) = h\gamma'(t),$$

jämför med 2.6. Denna linjärfunktion kan identifieras med vektorn $\gamma'(t)$.

Den fysikaliska tolkningen av stigen $\gamma'(t)$ är hastigheten för "den rörliga partikeln" γ i ögonblicket t och $|\gamma'(t)| = (\gamma_1'(t)^2 + \ldots + \gamma_n'(t)^2)^{1/2}$ är partikelns fart i ögonblicket t.

Om $\gamma'(t) \neq 0$, så ger $\gamma'(t)$ stigens tangent (riktning) med parametervärdet t.

Om $\gamma'(t) \neq 0$, så representerar tangenten linjen

$$\lambda \mapsto \gamma(t) + \lambda\gamma'(t),\ \lambda \in \mathbb{R}.$$

Denna är naturligtvis också en stig, vars definitionsintervall är hela $\mathbb{R}$.

Om γ' är kontinuerlig, så definierar den en ny stig $\gamma' : [a, b] \to \mathbb{R}^n$. Om $\gamma' \in \mathcal{C}^2([a, b])$, så kan vi definiera $\gamma'' : [a, b] \to \mathbb{R}^n$,

$$\gamma''(t) = (\gamma_1''(t), \ldots, \gamma_n''(t)).$$

Den fysikaliska tolkningen av stigen $\gamma''(t)$ accelerationen för “partikeln” γ i ögonblicket t.

3.1.2. Exempel. Låt $\gamma : [0, 2\pi] \to \mathbb{R}^2$, $\gamma(t) = (a \cos t, b \sin t)$, där $a, b > 0$. Då är

$$\gamma'(t) = (-a \sin t, b \cos t)$$

och

$$\gamma''(t) = (-a \cos t, b \sin t) = -\gamma(t).$$

Den andra derivatan $\gamma''(t)$ är alltså motsatt riktad i förhållande till vektorn $\gamma(t)$. Då t går igenom intervallet $[0, 2\pi]$, så cirkulerar $\gamma(t)$ längs ellipsen $x^2/a^2 + y^2/b^2 = 1$.

Kom ihåg, att om $\gamma \in \mathcal{C}^1([a, b])$, så är längden $l(\gamma)$ av stigen γ

$$l(\gamma) = \int_a^b |\gamma'(t)| \, \mathrm{d}t = \int_a^b \left(\gamma_1'(t)^2 + \ldots + \gamma_n'(t)^2\right)^{1/2} \mathrm{d}t.$$

Stigens längd, främst i fallet $n = 2$, behandlas vanligtvis i samband med teori om funktioner med en variabel.

3.1.3. Exempel. Låt $f : [a, b] \to \mathbb{R}$ vara kontinuerlig. Då är f naturligtvis också själv en stig, men f definierar även stigen $\gamma : [a, b] \to \mathbb{R}^2$ med ekvationen

$$\gamma(t) = (t, f(t)).$$

Vi observerar, att

$$\gamma([a, b]) = \{(x, y) \in \mathbb{R}^2 : y = f(x), x \in [a, b]\},$$

alltså att bilden av intervallet $[a, b]$ i avbildningen γ är grafen av funktionen f. Om f har en derivata $f'(t)$, så gäller att

$$\gamma'(t) = (1, f'(t)) \neq (0, 0).$$

Tangenten av stigen γ i punkten $t \in [a, b]$ representerar linjen

$$\lambda \overset{T}{\mapsto} \gamma(t) + \lambda\gamma'(t) = (t, f(t)) + \lambda(1, f'(t)) = (t + \lambda, f(t) + \lambda f(t)),$$

där $T : \mathbb{R} \to \mathbb{R}^2$. Punkterna $(x, y) \in T(\mathbb{R})$ på linjen uppfyller ekvationen (erhålls genom att eliminera λ)

$$y = f(t) + f'(t)(x - t).$$

Detta är ekvationen för tangenten till f i punkten t.

3.2. Ytor

Vi granskar en "2-dimensionell" *yta* i $\mathbb{R}^3$. Låt $D \subset \mathbb{R}^2$, $r : D \to \mathbb{R}^3$,

$$r(x, y) = (r_1(x, y), r_2(x, y), r_3(x, y)).$$

Vi antar att D är öppen, fastän detta inte nödvändigtvis gäller i tillämpningar; D kan till exempel även vara en sluten kvadrat. Vi fixerar y_0 och undersöker avbildningen

$$x \overset{\gamma_1}{\mapsto} r(x, y_0).$$

Denna definierar en stig på ett intervall innehållande x_0, då punkten $(x_0, y_0) \in D$. Denna stig färdas "på ytan" $r(D)$. Motsvarande gäller för stigen

$$y \overset{\gamma_2}{\mapsto} r(x_0, y).$$

Vi antar att $r \in \mathcal{C}^1(D)$, med andra ord att $r_i \in \mathcal{C}^1(D),\ i = 1, 2, 3$. Då existerar derivatorna

$$\begin{aligned}
\partial_1 r(x, y) &= (\partial_1 r_1(x, y), \partial_1 r_2(x, y), \partial_1 r_3(x, y)) \\
&= \gamma_1'(x), \\
\partial_2 r(x, y) &= (\partial_2 r_1(x, y), \partial_2 r_2(x, y), \partial_2 r_3(x, y)) \\
&= \gamma_2'(y).
\end{aligned}$$

Om vektorerna $\partial_1 r(x, y)$ och $\partial_2 r(x, y)$ inte är likriktade eller nollvektorer, så spänner de upp ett plan genom punkten (x, y). Detta kallas ytans *tangentplan* i punkten $r(x, y) \in r(D)$. Eftersom $\partial_1 r(x, y) \nparallel \partial_2 r(x, y)$, är $\partial_1 r(x, y) \times \partial_2 r(x, y) \neq 0$ och $\partial_1 r(x, y) \times \partial_2 r(x, y)$ är en vektor, som står vinkelrätt mot tangentplanet, med andra ord är det en *normalvektor*. Ovan har vi använt vanlig kryssprodukt i $\mathbb{R}^3$. Tangentplanet består alltså av de vektorer $\overline{x}$, för vilka

$$(r(x,y)-\overline{x})\cdot(\partial_1 r(x,y)\times\partial_2 r(x,y))=0.$$

3.2.1. Exempel. Vi undersöker en origocentrerad bollyta med radien R i $\mathbb{R}^3$, $r:\mathbb{R}^2\to\mathbb{R}^3$,

$$r=(r_1,r_2,r_3)=(R\sin\theta\cos\varphi,R\sin\theta\sin\varphi,R\cos\theta),$$

där vi har har använt den polära framställningen (φ,θ) för punkterna på bollytan. Vinkeln θ är beräknad på den positiva x_3-axeln och φ är polvinkeln i x_1x_2-planet. Bollytan har egenskapen att $r=r(\varphi,\theta)$ är likriktad med normalvektorn. För att bekräfta detta räknar vi

$$\begin{aligned}\partial_1 r(\varphi,\theta)&=R(-\sin\theta\sin\varphi,\sin\theta\cos\varphi,0),\\ \partial_2 r(\varphi,\theta)&=R(\cos\theta\cos\varphi,\cos\theta\sin\varphi,-\sin\theta).\end{aligned}$$

Av dessa får vi vidare att

$$\begin{aligned}\partial_1 r\times\partial_2 r&=R^2\begin{bmatrix}i&j&k\\-\sin\theta\sin\varphi&\sin\theta\cos\varphi&0\\\cos\theta\cos\varphi&\cos\theta\sin\varphi&-\sin\theta\end{bmatrix}\\ &=R^2(-\sin^2\theta\cos\varphi,-\sin^2\theta\sin\varphi,\\ &\qquad -\sin\theta\cos\theta\sin^2\varphi-\sin\theta\cos\theta\cos^2\varphi)\\ &=R^2(\sin\theta)(\sin\theta\cos\varphi,\sin\theta\sin\varphi,\cos\theta)\\ &=R(-\sin\theta)r(\varphi,\theta),\end{aligned}$$

d.v.s. kryssprodukten $\partial_1 r\times\partial_2 r$ är likriktad med $r(\varphi,\theta)$. Märk, att med detta framställningssätt gäller det att $\partial_1 r\times\partial_2 r=0$ vid nord- och sydpolen ($\theta=0$ och $\theta=\pi$).

3.3. Inversa funktionssatsen

I fallet $n=1$, om $f:(a,b)\to\mathbb{R}$, $f\in\mathcal{C}^1((a,b))$ och $f'(x)\neq 0$ för varje $x\in(a,b)$, så definierar f en bijektion $(a,b)\to f((a,b))=\Delta$ och det existerar en avbildning $g=f^{-1}\in\mathcal{C}^1(\Delta)$. Detta är satsen om inversa funktioner i fallet $n = 1$. Dess bevis grundar sig på, att det av kontinuiteten för f' och villkoret $f'(x) \neq 0$ för alla $x \in (a, b)$ följer, att f är strängt monoton.

Låt $D \subset \mathbb{R}^n$ vara öppen. Vi strävar efter att finna villkor för egenskapen hos funktionen $f : D \to \mathbb{R}^n$ att *lokalt ha en invers*. Att konstatera att avbildningen f är en bijektion i hela D är svårare i fallet $n \geq 2$. Observera, att start- och målmängdens dimension n är densamma. I fallet $n \geq 2$ orsakas svårigheterna av att det i $\mathbb{R}^n$ inte finns någon enkel motsvarighet för strängt monotona avbildningar.

Vi granskar först ett enkelt fall. Låt $f : \mathbb{R}^n \to \mathbb{R}^n$ vara lineär. Vi betecknar $f = A$ och antar att

$$A = \underbrace{\begin{bmatrix} a_{11} & \dots & a_{n1} \\ \vdots & \ddots & \\ a_{1n} & \dots & a_{nn} \end{bmatrix}}_{\text{avbildningsmatris}}.$$

Vi använder oss av kunskaper i lineäralgebra: Avbildningen A har en invers funktion $A^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ (som också är lineär) exakt när $\det A \neq 0$, var

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Det lönar sig att komma ihåg att $\det A = 1/\det(A^{-1})$.

Antag nu att $f : D \to \mathbb{R}^n$ och att $x_0 \in D$. Om f är deriverbar i punkten x_0, så approximeras f "noggrannt" av avbildningen $L : \mathbb{R}^n \to \mathbb{R}^n$,

$$x \overset{L}{\mapsto} f(x_0) + f'(x_0)(x - x_0)$$

i en liten omgivning till punkten x_0. Avbildningen $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ är lineär. Om denna avbildning har en invers, d.v.s. om $\det(f'(x_0)) \neq 0$, så har avbildningen L en invers $L^{-1} : \mathbb{R}^n \to \mathbb{R}^n$, som bestäms av formeln

$$L^{-1}(y) = x_0 + f'(x_0)^{-1}(y - f(x_0)).$$

Att bevisa detta är enkelt. Om $x \in \mathbb{R}^n$, så är

$$\begin{aligned}(L^{-1} \circ L)(x) &= x_0 + f'(x_0)^{-1}(L(x) - f(x_0)) \\ &= x_0 + f'(x_0)^{-1}f'(x_0)(x - x_0) \\ &= x_0 + x - x_0 = x.\end{aligned}$$

På motsvarande sätt ser vi att $(L \circ L^{-1})(y) = y$ för varje $y \in \mathbb{R}^n$, varvid L^{-1} är den inversa funktionen av L. Om funktionen L, som approximerar funktionen f, har en invers, så verkar det sannolikt, att även funktionen f har en invers funktion åtminstone i en liten omgivning till punkten x_0. Med lämpliga tilläggsantaganden gäller detta och följande sats ger ett tillräckligt villkor för att funktionen f lokalt skulle ha en invers funktion.

3.3.1. Sats (inversa funktionssatsen)**.** *Låt $D \subset \mathbb{R}^n$ vara öppen, $f : D \to \mathbb{R}^n$ en $\mathcal{C}^1$-avbildning och $\det(f'(x)) \neq 0$ för alla $x \in D$. Då existerar det för varje $x \in D$ ett sådant $r > 0$, att $f|B(x, r)$ är en injektion och f definierar en bijektion*

$$B(x, r) \to f(B(x, r)).$$

Dessutom är $f(B(x, r))$ öppen, $f^{-1} : f(B(x, r)) \to B(x, r)$ en $\mathcal{C}^1$-avbildning, och

$$f^{-1\,\prime}(f(y)) = f'(y)^{-1} \tag{3.3.2}$$

för varje $y \in B(x, r)$.

Bevis. Förbigås; innehåller ett stort antal detaljer. □

Determinanten $\det(f'(x))$ kallas *Jacobis determinant* d.v.s. *Jacobianen* av funktionen f i punkten x och betecknas

$$\det f'(x) = \tau(x, f).$$

Observera betydelsen av formel (3.3.2): Vi får avbildningsmatrisen för lineärfunktionen $f^{-1\prime}(f(y))$ som den inversa avbildningsmatrisen för avbildningen $f'(y)$.

Det lönar sig att komma ihåg att det i $\mathbb{R}^n$, $n \geq 2$, inte finns någon enkel produkt, som skulle garantera att avbildningen $f : D \to \mathbb{R}^n$ är en injektion. Av att f är en injektion i något klot $B(x, r_x) \subset D$ för varje $x \in D$ följer inte, att f skulle vara en injektion i hela D.

3.3.3. Exempel. Låt $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (x^2 - y^2, 2xy)$. Uppenbart gäller att $f \in \mathcal{C}^1(\mathbb{R}^2)$ och f:s derivata $f'(x,y)$ har avbildningsmatrisen

$$f'(x,y) = \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix}$$

i punkten $(x,y) \in D$. Nu är

$$\tau((x,y), f) = \det f'(x,y) = \begin{vmatrix} 2x & -2y \\ 2y & 2x \end{vmatrix} = 4x^2 + 4y^2 \neq 0,$$

då $(x,y) \neq (0,0)$.

Av Sats 3.3.1 följer att det för varje $(x,y) \neq (0,0)$ finns en omgivning $U = B((x,y), r)$ vars restriktion f är en injektion, att $f : U \to f(U)$ är en bijektion samt att $f^{-1} : f(U) \to U$ är en $\mathcal{C}^1(f(U))$-avbildning.

Å andra sidan märker vi att $f(-x,-y) = (x^2 - y^2, 2xy) = f(x,y)$, alltså är f inte en injektion i någon omgivning till $(0,0)$. Speciellt är f inte en injektion i mängden $R^2 \backslash \{(0,0)\}$, fastän $\tau \neq 0$ i denna mängd.

Funktionen f ovan är avbildningen $z \mapsto z^2$, om vi i $\mathbb{R}^2$ använder komplex beteckning $(x,y) = x + iy = z$.

3.4. Lagranges multiplikatorer

Maximi- och minimipunkter till realvärda funktioner med flera variabler undersöks ofta i mängden $A_0 \subset \mathbb{R}^n$, med andra ord undersöks funktionen $f|_{A_0}$ d.v.s. *restriktionen* av funktionen f till mängden A_0. Vanligtvis är det fråga om att mängden A_0 är en "kurva" γ eller en "yta" S. Vi granskar situationen, då $n = 2$ och mängden A_0 bestäms av stigen γ. Låt $\Delta \subset \mathbb{R}$ vara ett intervall, $A \subset \mathbb{R}^2$ en mängd, $\gamma : \Delta \to \mathbb{R}^2$ en stig och $f : A \to \mathbb{R}$ en avbildning. Antag dessutom, att $A_0 = \gamma(\Delta) \subset A$.

Vi har som uppgift att bestämma extrempunkterna till den sammansatta funktionen $f \circ \gamma : \Delta \to \mathbb{R}$. Antag att $t \in \Delta$ och $\gamma(t) =$

$(x(t), y(t))$. Då är

$$(f \circ \gamma)(t) = f(x(t), y(t)).$$

Uppgiften är i princip enkel: Om f är en $\mathcal{C}^1$-avbildning och γ är deriverbar, så finns extremvärdena bland de eventuella ändpunkterna i Δ eller bland de punkter, i vilka

$$(f \circ \gamma)'(t) = 0.$$

Tidigare har vi konstaterat, se Sats 2.6.6, att

$$(f \circ \gamma)'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t) = \partial_1 f(\gamma(t))\, x'(t) + \partial_2 f(\gamma(t))\, y'(t) \cdot \gamma(t) = 0,$$

så vi behöver endast finna denna ekvations lösningar. Tyvärr är situationen vanligen inte given i den här formen.

En stig ges vanligtvis som en nivåyta $g(x, y) = 0$, där $D \subset \mathbb{R}^2$ är öppen och $g : D \to \mathbb{R}$. Vi undersöker fallet, då $f : D \to \mathbb{R}$ och $g : D \to \mathbb{R}$ är $\mathcal{C}^1(D)$-funktioner och vi vill bestämma extrempunkterna till f i mängden

$$A_0 = \{(x, y) \in D : g(x, y) = 0\}.$$

Vi antar, att $(x_0, y_0) \in A_0$ och $\nabla g(x_0, y_0) \neq 0$. Då kan funktionen $f|_{A_0}$ inte ha en extrempunkt i punkten (x_0, y_0), ifall $\nabla f(x_0, y_0)$ och $\nabla g(x_0, y_0)$ inte är likriktade. Med andra ord, om (x_0, y_0) är en extrempunkt till $f|_{A_0}$, så är

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0) \tag{3.4.1}$$

för något $\lambda \in \mathbb{R}$. Observera, att värdet $\lambda = 0$ är tillåtet.

Att bevisa detta är besvärligt, men den geometriska idén bakom resultatet är rätt så tydlig. Vi skall alltså visa, att om (3.4.1) inte gäller, så har $f|_{A_0}$ inte någon extrempunkt i punkten (x_0, y_0). Ifall (3.4.1) inte gäller, så är $\nabla f(x_0, y_0) \neq 0$ (om $\nabla f(x_0, y_0) = 0$, kan vi välja $\lambda = 0$). Å andra sidan är $\nabla g(x_0, y_0) \neq 0$, varvid vi kan använda den så kallade *implicita funktionssatsen*, som säger att mängden A_0 i en omgivning till (x_0, y_0) kan uttryckas som grafen av funktionen $y = y(x)$ (om $\partial_2 g(x_0, y_0) \neq 0$) eller funktionen $x = x(y)$ (om $\partial_1 g(x_0, y_0) \neq 0$), d.v.s. som en $\mathcal{C}^1$-stig $\gamma : [a, b] \to A_0$, där $\gamma(t_0) = (x_0, y_0)$, $t_0 \in (a, b)$ och

$\gamma'(t_0) \neq 0$. Nu är $\gamma'(t_0) \cdot \nabla g(x_0, y_0) = 0$, se Sats 2.7.5, varvid vektorerna $\gamma'(t_0)$ och $\nabla g(x_0, y_0)$ bildar basen för planet i punkten (x_0, y_0). Linjen

$$t \mapsto (x_0, y_0) + t\nabla g(x_0, y_0)$$

delar planet i två delar T_1 och T_2, och eftersom $\nabla f(x_0, y_0)$ inte är likriktad med vektorn $\nabla g(x_0, y_0)$, så är den riktad mot någon av delarna, t.ex. mot T_1. Nu växer funktionen f kraftigast i riktningen $\nabla f(x_0, y_0)$ sett från punkten (x_0, y_0), och eftersom a_0 "nästan" sammanfaller med linjen

$$t \mapsto (x_0, y_0) + t\gamma'(t_0)$$

given av vektorn $\gamma'(t_0)$ nära punkten (x_0, y_0), så är $f(x, y) > f(x_0, y_0)$, där $(x, y) \in T_1 \cap A_0$ och (x, y) är nära punkten (x_0, y_0). På motsvarande sätt är $f(x, y) < f(x_0, y_0)$ då $(x, y) \in T_2 \cap A_0$ och (x, y) är tillräckligt nära punkten (x_0, y_0). Således kan funktionen $f|_{A_0}$ inte ha en extrempunkt i punkten (x_0, y_0).

Då den ovannämnda härledningen görs exakt, får vi följande sats:

3.4.2. Sats. *Låt $D \subset \mathbb{R}^2$ vara öppen, $f, g : D \to \mathbb{R}$, f och g $\mathcal{C}^1(D)$-funktioner samt*

$$A_0 = \{(x, y) \in D : g(x, y) = 0\}.$$

Om funktionen $f|_{A_0}$ har en lokal extrempunkt i punkten $(x_0, y_0) \in A_0$ och

$$\nabla g(x_0, y_0) \neq 0,$$

så är

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$$

för något $\lambda \in \mathbb{R}$. Dylika λ kallas Lagranges multiplikatorer för funktionen f.

3.4.3. Exempel. En likbent triangel, vars bas är parallell med x-axeln, är inskriven i enhetscirkeln. Vi skall bestämma den största möjliga arean för triangeln. Vi betecknar halva basens längd med x och antar att punkten $(0, y)$ är den punkt, där basen skär y-axeln. Då är triangelns area $f(x, y)$,

$$f(x, y) = \frac{1}{2} 2x(1 - y) = x(1 - y),$$

där $x \geq 0$. Punkten (x, y) uppfyller dessutom ekvationen $x^2 + y^2 = 1$.

Vi skall bestämma maximivärdet av f i mängden

$$A_0 = \{(x, y) \in \mathbb{R}^2 : x > 0,\ x^2 + y^2 = 1\}.$$

Vi definierar $g(x, y) = x^2 + y^2 - 1$, varvid

$$A_0 = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0,\ x > 0\}.$$

Nu kan vi använda Sats 3.4.2. Som mängd D kan vi ta den öppna mängden

$$D = \{(x, y) : x > 0\}$$

d.v.s. det högra öppna halvplanet. Funktionerna f och g är uppenbart $\mathcal{C}^1(D)$-funktioner. Om f har ett lokalt extremvärde i punkten (x, y) i A_0, så gäller alltså antingen att

$$(1 - y, -x) = \nabla f(x, y) = \lambda \nabla g(x, y)$$

eller att $\nabla g(x, y) = 0$. Det senare villkoret kommer inte ifråga i mängden D och således inte heller i mängden A_0. Av ekvationen ovan får vi ekvationssystemet

$$\begin{cases} (1 - y, -x) = \lambda(2x, 2y) = (\lambda 2x, \lambda 2y), \\ g(x, y) = 0, \end{cases}$$

som kan skrivas i formen

$$\begin{cases} 1 - y = \lambda 2x \\ -x = \lambda 2y \\ x^2 + y^2 = 1 \end{cases}$$

Av den första ekvationen får vi att

$$\lambda = \frac{1 - y}{2x},$$

och genom att substituera detta i den andra ekvationen har vi kvar att lösa ekvationssystemet

$$\begin{cases} -x^2 + y^2 - y = 0, \\ x^2 + y^2 = 1, \end{cases}$$

vars lösningar är $y = 1$ och $y = -\dfrac{1}{2}$. Värdet $y = 1$ duger inte, ty då måste x vara 0. Därmed är punkten $(\sqrt{3}/2, -1/2)$ den enda eventuella lokala extrempunkten. En kontroll visar, att denna faktiskt uppfyller den ursprungliga ekvationen.

Vi utreder, ifall det är fråga om ett absolut maximum. Mängden $x^2 + y^2 = 1$, $x \geq 0$ är kompakt (sluten och uppenbart begränsad). Därmed får den kontinuerliga funktionen $f(x, y) = x(1 - y)$ sitt maximivärde i mängden. I punkterna $(0, 1)$ och $(0, -1)$ antar funktionen värdet 0. Således uppnår funktionen sitt absoluta maximum i den enda lokala extrempunkten $(\sqrt{3}/2, -1/2)$, ty $f > 0$ i denna punkt. Arean är således störst då triangeln är liksidig.

3.4.4. Exempel. Vi söker det kortaste avståndet från origo till mängden

$$A_0 = \{(x, y) \in \mathbb{R}^2 : x^2 y = 16\}.$$

Avståndet från en punkt till origo är $\sqrt{x^2 + y^2}$, men i fråga om minimering kan vi lika väl undersöka funktionen $f(x, y) = x^2 + y^2$. Vi betecknar $g(x, y) = x^2 y - 16$. Då är

$$A_0 = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}.$$

Vi räknar de partiella derivatorna

$$\frac{\partial f}{\partial x} = 2x, \ \frac{\partial f}{\partial y} = 2y, \ \frac{\partial g}{\partial x} = 2xy, \ \frac{\partial g}{\partial y} = x^2,$$

och tillämpar Sats 3.4.2. Vi får att

$$\begin{cases} 0 = 2x + 2\lambda xy \\ 0 = 2y + \lambda x^2 \\ 0 = x^2 y - 16 \end{cases}$$

Av den första ekvationen får vi att $x = 0$ eller $\lambda y = -1$, av vilka den första inte duger. Med substitution i de övriga ekvationerna får vi att $x = \pm\sqrt{2}y$, vilket ger att $y = 2$. De eventuella extrempunkterna är således $(\pm 2\sqrt{2}, 2)$. Dessutom är $\nabla g(x, y) \neq 0$ i mängden A_0. Det kortasete avståndet är således $\sqrt{8 + 4} = 2\sqrt{3}$. Geometriskt sett är det klart, att det kortaste avståndet existerar; att bevisa detta analytiskt lämnas som övningsuppgift.

Genom att kombinera tidigare resultat om eventuella extrempunkter i en öppen mängd med Sats 3.4.2, erhåller vi en metod, som används för att finna extrempunkter till funktionen $f \in \mathcal{C}^1(D)$ i den kompakta mängden $\overline{A} = A \cup \partial A \subset D$ under följande förutsättningar:

(i) Mängden A är öppen och $\overline{A} = A \cup \partial A$ är komapkt.

(ii) Randen för mängden A är
$$\partial A = \bigcup_{i=1}^{k} C_i,$$
där
$$C_i \subset \{(x, y) : g_i(x, y) = 0\}$$
och g_i är en $\mathcal{C}^1(D_i)$-funktion i någon omgivning $D_i,\ i = 1, \ldots, k$ till mängden C_i.

(iii) Mängderna C_i är *disjunkta*, med andra ord är $C_i \cap C_j = \emptyset$ för varje $i \neq j$, förutom i en ändlig mängd "hörnpunkter".

De eventuella lokala extrempunkterna till funktionen f i mängden $\overline{A}$ är då

a) punkterna $(x, y) \in A$, för vilka $\nabla f(x, y) = 0$, med andra ord de normala kritiska punkterna för f i A,

b) punkterna $(x, y) \in C_i$, för vilka $\nabla g_i(x, y) = 0,\ i = 1, \ldots, k$,

c) punkterna $(x, y) \in C_i$, för vilka $\nabla f(x, y) = \lambda \nabla g_i(x, y),\ i = 1, \ldots, k,\ \lambda \in \mathbb{R}$ (Sats 3.4.2) och

d) eventuella "hörnpunkter" i ∂A.

Eftersom $\overline{A}$ är kompakt, så finns absoluta minima och maxima av den kontinuerliga funktionen f bland dessa punkter.

Metoden ovan följer av att de eventuella extrempunkterna i en öppen mängd finns bland de kritiska punkterna, samt av att om punkten $(x, y) \in \partial A$ är en extrempunkt till funktionen f i $\overline{A}$, så är den ovillkorligen även en extrempunkt till funktionen $f|_{\partial A}$.

Ofta innehåller $\overline{A}$ punkter, i vilka f eller g inte är deriverbara. Dessa punkter måste undersökas skilt.

3.4.5. Anmärkning. Teorin kan omedelbart utvidgas till fallet $n \geq 3$. Då $n \geq 3$, är nivåytan
$$A_0 = \{x \in D : g(x) = 0\}$$

vanligtvis ytlik.

3.4.6. Exempel. Bestäm de punkter, i vilka funktionen

$$f(x,y) = 8x^2 - 12xy + 17y^2$$

antar sitt största och sitt minsta värde i den slutna enhetskulan $\overline{B}(0,1)$.

Funktionen f är uppenbart en $\mathcal{C}^1(\mathbb{R}^2)$-funktion; den är i själva verket en så kallad andragradsform. Eftersom $\overline{B}(0,1)$ är kompakt, erhåller f sitt största och minsta värde i $\overline{B}(0,1)$. Nu är

$$\overline{B}(0,1) = B(0,1) \cup \partial B(0,1),$$

där

$$\partial B(0,1) = \{(x,y) : x^2 + y^2 - 1 = 0\}.$$

Vi betecknar $g(x,y) = x^2 + y^2 - 1$. Nu har vi situationen, som beskrivs ovan. Vi bestämmer först de kritiska punkterna till f i $B(0,1)$. Vi får att

$$\partial_1 f(x,y) = 16x - 12y = 0$$
$$\partial_2 f(x,y) = -12x + 34y = 0.$$

Detta är ett lineärt ekvationssystem, vars determinant uppenbart är olik 0, d.v.s. det har endast den triviala lösningen $(x,y) = (0,0)$. Därmed är $(0,0)$ den enda kritiska punkten till funktionen f i kulan $B(0,1)$. Eftersom varje extrempunkt i $B(0,1)$ också är en kritisk punkt, så är origo den enda eventuella extrempunkten för funktionen f i den öppna enhetskulan.

Vi undersöker ännu eventuella extrempunkter i mängden $\partial B(0,1)$. Nu har vi endast en funktion g_i, som vi betecknar $g_i = g$, och

$$\nabla g(x,y) = (2x, 2y) \neq (0,0),$$

då (x,y) är på enhetscirkelns periferi. Det räcker alltså att undersöka fall c).

Ekvationen $\nabla f(x,y) = \lambda \nabla g(x,y)$ består av två “komponentekvationer”

$$\begin{cases} 16x - 12y - 2\lambda x &= 0 \\ -12x + 34y - 2\lambda y &= 0. \end{cases}$$

Genom att eliminera λ får vi att

$$\frac{x}{y} - \frac{y}{x} = \frac{3}{2},$$

då $x \neq 0$ och $y \neq 0$. Detta ger att

$$x^2 - \frac{3}{2}xy - y^2 = 0$$

alltså att

$$(x - 2y)(x + \frac{y}{2}) = 0,$$

av vilket vi ser, att de eventuella extrempunkterna i $\partial B(0,1)$ är skärningspunkterna mellan linjerna $x = 2y$ och $2x = -y$ samt cirkeln $x^2 + y^2 = 1$. Punkterna $x = 0$ eller $y = 0$ duger inte. Vi får att skärningspunkterna är

$$\pm(2/\sqrt{5}, 1/\sqrt{5}), \ \pm(1/\sqrt{5}, -2/\sqrt{5}).$$

Nu behöver vi endast räkna värdet av f i dessa punkter och i punkten $(0,0)$ samt välja det största och minsta värdet av dessa. På grund av att funktionen f är symmetrisk, så inverkar inte $\pm$ på den ifrågavarande punktens värde i f. I punkten $(0,0)$ antar funktionen f värdet 0, i punkten $(2/\sqrt{5}, 1/\sqrt{5})$ värdet 5 och i punkten $(1/\sqrt{5}, -2/\sqrt{5})$ värdet 20.

Genom att kombinera resultaten ovan observerar vi, att funktionen f antar sitt minimivärde i punkten $(0,0)$ och sitt maximivärde i punkterna

$$\pm(1/\sqrt{5}, -2/\sqrt{5}).$$

En liten tilläggsanalys visar, att extremvärdena i dessa punkter är stränga. För punkten $(0,0)$ framgår detta med hjälp av analys hos kritiska punkter, men för punkterna $\pm(1/\sqrt{5}, -2/\sqrt{5})$ krävs lite extra härledning. En tilläggsanalys visar, att den kvadratiska formen definierad av f är positivt definit, jämför med kapitel 2.9.

4. Integrering i planet

4.1. Integralens definition över en rektangel

I det euklidiska rummet $\mathbb{R}^n$ motsvaras itegrationsintervallet $[a, b]$ av en “låda”, som är en kartesisk produkt av n intervall. Riemann-integralen över lådan skiljer sig inte väsentligt från Riemann-integralen över ett endimensionellt intervall $[a, b]$. Svårigheter med integrering i $\mathbb{R}^n$ uppstår då man integrerar över mera komplicerade mängder än lådor. I praktiska tillämpningar hamnar vi ofta i dessa situationer. I detta kapitel behandlar vi fallet $n = 2$. Generaliseringar till högre dimensioner är enkla. Vi utgår från en sluten rektangel i $\mathbb{R}^2$,

$$D = \{(x_1, x_2) \in \mathbb{R}^2 : a \leq x_1 \leq b, c \leq x_2 \leq d\} = [a, b] \times [c, d],$$

där $a < b$ och $c < d$.

Delningen $\{R_{ij}\}$ av rektangeln består av delningen av intervallen $[a, b]$ och $[c, d]$: $a = x_0 < x_1 < \ldots < x_m = b$, $c = y_0 < y_1 < \ldots < y_n = d$, där rektanglarna $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$, $1 \leq i \leq m,\ 1 \leq j \leq n$, bildar delningen $\{R_{ij}\}$ av mängden D.

Arean $\Delta(R_{ij})$ av rektangeln R_{ij} definieras på ett naturligt sätt som produkten av sidornas längder, d.v.s.

$$\Delta(R_{ij}) = \underbrace{(x_i - x_{i-1})}_{\Delta x_i} \underbrace{(y_j - y_{j-1})}_{\Delta y_j}.$$

Här använder vi beteckningen $\Delta(R_{ij})$; serare övergår vi dock till beteckningen $\mathrm{area}(R_{ij})$. Med diametern av rektangeln R_{ij} avser vi längden på dess diagonal

$$d(R_{ij}) = (\Delta x_i^2 + \Delta y_j^2)^{1/2},$$

och med normen $\|R_{ij}\|$ av delningen $\{R_{ij}\}$ avser vi den största diagonalen av rektanglarna R_{ij}, alltså

$$\|R_{ij}\| = \max\{d(R_{ij}) : 1 \leq i \leq m,\ 1 \leq j \leq m\}.$$

Låt $f : D \to \mathbb{R}$ vara en begränsad funktion. *Riemannsumman i delningen* $\{R_{ij}\}$ av funktionen f är

$$\begin{aligned} R(f,\ R_{ij}) &= \sum_{i,j} f(\xi_{ij})\,\Delta(R_{ij}) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} f(\xi_{ij})\,\Delta(R_{ij}), \end{aligned}$$

där $\xi_{ij} \in R_{ij}$. Märk, att $R(f,\ R_{ij})$ även beror på valet av punkterna ξ_{ij}, fastän vi inte skilt betecknar detta beroende.

4.1.1. Definition. En begränsad funktion f är *Riemann-integrerbar* över mängden D, om det existerar ett tal $I \in \mathbb{R}$, för vilket följande villkor gäller: För varje $\varepsilon > 0$ existerar det ett sådant $\delta > 0$, att

$$|R(f, R_{ij}) - I| < \varepsilon$$

då delningen $\{R_{ij}\}$ uppfyller villkoret $\|R_{ij}\| < \delta$. Vi betecknar

$$\begin{aligned} I &= \int_D f\ \mathrm{d}x\,\mathrm{d}y = \int_D f\ \mathrm{d}A = \iint_D f\ \mathrm{d}x\,\mathrm{d}y \\ &= \iint_D f(x,y)\ \mathrm{d}x\,\mathrm{d}y = \int_a^b \int_c^d f(x,y)\ \mathrm{d}y\ \mathrm{d}x. \end{aligned}$$

Såsom med funktioner med en variabel, är det möjligt att bilda över- och undersummor som hänför sig till funktionen f,

$$\begin{aligned} S(f, R_{ij}) &= \sum_{i,j} \sup\{R(f, R_{ij}) : \xi_{ij} \in R_{ij}\}\ \text{(översumma)}, \\ s(f, R_{ij}) &= \sum_{i,j} \inf\{R(f, R_{ij}) : \xi_{ij} \in R_{ij}\}\ \text{(undersumma)}. \end{aligned}$$

Båda dessa existerar, ty f är begränsad.

Bevisen för följande satser avviker inte väsentligt från det endimensionella fallet.

4.1.2. Sats. *Låt $f : D \to \mathbb{R}$ vara begränsad. Då är f Riemann-integrerbar över D om och endast om det för varje $\varepsilon > 0$ existerar en sådan delning $\{R_{ij}\}$, att $S(f, R_{ij}) - s(f, R_{ij}) < \varepsilon$.*

4.1.3. Sats. *Antag att $f : D \to \mathbb{R}$ är kontinuerlig. Då är f Riemann-integrerbar.*

4.1.4. Anmärkning. Beviset för Sats 4.1.3 motsvarar fullt beviset för den endimensionella motsvarigheten. Beviset grundar sig på likformig kontinuitet hos f, ty i en kompakt mängd, såsom mängden D, är en kontinuerlig funktion likformigt kontinuerlig.

4.1.5. Exempel. Integralens definition 4.1.1 är svår att använda för att bestämma värdet av integralen. Ibland lyckas det ändå. Antag att $D = [a, b] \times [c, d]$ och att $f : D \to \mathbb{R}$, $f(x, y) = \alpha$ är en konstant funktion. Nu är

$$\int_D f \,\mathrm{d}x\,\mathrm{d}y = \alpha \underbrace{\operatorname{area}(D)}_{(b-a)(d-c)},$$

eftersom varje Riemannsumma har detta samma konstanta värde $\alpha \operatorname{area}(D)$.

Beräkning av integralen som en *itererad integral* är en av de viktigaste metoderna för att bestämma flerdimensionella integraler. Antag att $D = [a, b] \times [c, d]$ och att $f : D \to \mathbb{R}$ är kontinuerlig. För ett fixerat $x \in [a, b]$ är då funktionen $y \mapsto f(x, y)$ kontinuerlig i intervallet $[c, d]$ och integralen

$$\int_c^d f(x, y) \,\mathrm{d}y$$

existerar. På motsvarande sätt, för ett fixerat $y \in [c, d]$, existerar integralen

$$\int_a^b f(x, y) \,\mathrm{d}x.$$

4.1.6. Sats. *Låt $f : D \to \mathbb{R}$ vara kontinuerlig. Då är*

$$\int_D f \,\mathrm{d}x\,\mathrm{d}y = \int_a^b \left(\int_c^d f(x, y) \,\mathrm{d}y \right) \mathrm{d}x = \int_c^d \left(\int_a^b f(x, y) \,\mathrm{d}x \right) \mathrm{d}y.$$

Bevis. Vi skisserar upp idén för beviset. Vi väljer $\xi_{ij} = (x_i, y_j) \in \mathbb{R}_{ij}$. Då är

$$\begin{aligned}
\int_D f(x, y) \,\mathrm{d}x\,\mathrm{d}y \approx R(f, R_{ij}) &= \sum_{i=1}^m \sum_{j=1}^n f(\xi_{ij}) \underbrace{\Delta(R_{ij})}_{\Delta x_i \Delta y_j} \\
&= \sum_{i=1}^m \Delta x_i \sum_{j=1}^n f(x_i, y_j) \Delta y_j \\
&\approx \sum_{i=1}^m \Delta x_i \int_c^d f(x_i, y) \,\mathrm{d}y \approx \int_a^b \int_c^d f(x, y) \,\mathrm{d}y \,\mathrm{d}x.
\end{aligned}$$

I beviset använder vi oss av kunskapen om att funktionen

$$x \mapsto \int_c^d f(x, y)\ \mathrm{d}y$$

är kontinuerlig och därmed integrerbar över intervallet $[a, b]$. Kontinuiteten hos funktionen är relativt enkel att se. Detta lämnas som övningsuppgift. Beviset för den andra ekvationen går på samma sätt. □

Märk, att det inte allmänt gäller att

$$\int_D f(x, y)\ \mathrm{d}x\,\mathrm{d}y = \int_a^b \int_c^d f(x, y)\ \mathrm{d}y\ \mathrm{d}x,$$

då f är Riemann-integrerbar. Orsaken till detta är att itererade integraler inte nödvändigtvis existerar. Formeln gäller, då f är kontinuerlig. Det föregående beviset använder sig av detta. Formeln i Sats 4.1.6 gäller även i allmännare fall.

4.1.7. Sats. *Antag att $f : D \to \mathbb{R}$ är begränsad och att integralen*

$$\int_D f\ \mathrm{d}x\,\mathrm{d}y$$

existerar, med andra ord att f är Riemann-integrerbar. Om integralen

$$\int_a^b f(x, y)\ \mathrm{d}x$$

existerar för varje $y \in [c, d]$ och funktionen

$$y \mapsto \int_a^b f(x, y)\ \mathrm{d}x$$

är Riemann-integrerbar i intervallet $[c, d]$, så är

$$\int_D f\ \mathrm{d}x\,\mathrm{d}y = \int_c^d \int_a^b f(x, y)\ \mathrm{d}x\,\mathrm{d}y.$$

Detsamma gäller, om vi i antagandena byter roller för x och y.

Bevis. Förbigås, fastän det inte är svårt. □

För integraler gäller samma egenskaper som i det endimensionella fallet.

4.1.8. Sats. *Antag att $D = [a, b] \times [c, d]$, $\lambda \in \mathbb{R}$ och att funktionerna $f, g : D \to \mathbb{R}$ är Riemann-integrerbara. Då gäller att*

a)
$$\int_D (f+g)\,\mathrm{d}x\,\mathrm{d}y = \int_D f\,\mathrm{d}x\,\mathrm{d}y + \int_D g\,\mathrm{d}x\,\mathrm{d}y,$$
b)
$$\int_D \lambda f\,\mathrm{d}x\,\mathrm{d}y = \lambda \int_D f\,\mathrm{d}x\,\mathrm{d}y,$$
c) *om* $f(x,y) \le g(x,y)$ *för varje* $(x,y) \in D$, *så är*
$$\int_D f\,\mathrm{d}x\,\mathrm{d}y \le \int_D g\,\mathrm{d}x\,\mathrm{d}y,$$
d)
$$\int_D f\,\mathrm{d}x\,\mathrm{d}y \le \int_D |f|\,\mathrm{d}x\,\mathrm{d}y,$$
e) *om* $D_1, \ldots, D_n \subset D$ *är väsentligt disjunkta rektanglar och* $\bigcup_{i=1}^n D_i = D$, *så är*
$$\sum_{i=1}^n \int_{D_i} f\,\mathrm{d}x\,\mathrm{d}y = \int_D f\,\mathrm{d}x\,\mathrm{d}y.$$

Bevis. Beviset följer till stor del beviset för fallet $n = 1$. □

Med *väsentligt disjunkta* rektanglar menas, att de har gemensamma punkter endast på sina kanter. Märk, att alla rektanglar är slutna.

4.1.9. Exempel. Låt $D = [1,3] \times [0,1]$. Vi räknar
$$\iint_D \frac{x}{(1+xy)^2}\,\mathrm{d}x\,\mathrm{d}y.$$
Nu är funktionen
$$f(x,y) = \frac{x}{(1+xy)^2}$$
kontinuerlig i mängden D. Detta ser vi av att nämnaren är ett polynom med två variabler, och att denna inte får värdet noll i mängden D. Därmed kan vi räkna integralen som en itererad integral:
$$\begin{aligned}\iint_D f\,\mathrm{d}x\,\mathrm{d}y &= \int_1^3 \int_0^1 \frac{x}{(1+xy)^2}\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_1^3 \left(\Big/_{0}^{1} -\frac{1}{1+xy} \right) \mathrm{d}x = \int_1^3 \left(1 - \frac{1}{1+x}\right) \mathrm{d}x \\ &= \Big/_{1}^{3} (x - \log(1+x)) = 2 - \log 2.\end{aligned}$$

4.1.10. Anmärkning. Teorin är likadan i fallet $n \geq 3$. Då är D "lådan"

$$D = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n].$$

4.2. Integrering över godtyckliga mängder

Följande teori gäller även då $n = 1$, men den behandlas sällan skilt. Antag att $A \subset \mathbb{R}^2$ är en begränsad mängd och att $f : A \to \mathbb{R}$ är en begränsad funktion. Märk, att f inte behöver vara kontinuerlig. Eftersom A är begränsad, existerar det en sådan rektangel D, att $D \supset A$.

4.2.1. Definition. Vi betecknar

$$\int_A f \,\mathrm{d}x\,\mathrm{d}y = \int_D \tilde{f} \,\mathrm{d}x\,\mathrm{d}y,$$

där

$$\tilde{f}(x) = \begin{cases} f(x) & , \text{då } x \in A \\ 0 & , \text{då } x \in D \backslash A, \end{cases}$$

om den senare integralen existerar.

Det är relativt enkelt att se, att integralens definition är oberoende av den valda rektangeln D.

Märk, att $\tilde{f}$ sällan är kontinuerlig, fastän f vore kontinuerlig i A. När existerar då integralen

$$\int_D \tilde{f} \,\mathrm{d}x\,\mathrm{d}y \ ?$$

För detta räcker det inte att f är kontinuerlig i A, utan vi måste ta komplexiteten hos ∂A i beaktande.

Vi behandlar inte detta problem mera ingående, utan vi ger ett svar som täcker våra praktiska behov. Detta är i princip inte svårt att bevisa. Till svaret behöver vi begreppet *Lebesgues nollmängd.* Med nollmängd avser vi en mängd, vars mått är noll med avseende på *Lebesgues mått.*

Mängden $N \subset \mathbb{R}^2$ är en nollmängd om det för varje $\varepsilon > 0$ existerar en sådan, möjligen ändlig, följd rektanglar $R_i = [a_i, b_i] \times [c_i, d_i]$, att

$$\sum_{i=1}^{\infty} \underbrace{(b_i - a_i)(d_i - c_i)}_{\text{area}(R_i)} \leq \varepsilon$$

och

$$N \subset \bigcup_i R_i,$$

med andra ord *täcker* rektanglarna R_i mängden N.

4.2.2. Exempel. Sträckan $I = \{(t, 0) : 0 \leq t \leq 1\}$ är en nollmängd i planet $\mathbb{R}^2$, ty rektangeln

$$R = [0, 1] \times [-\varepsilon/2,\ \varepsilon/2],\ \varepsilon > 0$$

täcker I, m.a.o. $I \subset R$ och

$$\text{area}(R) = 1 \cdot (\varepsilon/2 + \varepsilon/2) = \varepsilon.$$

En delmängd av en nollmängd är uppenbart en nollmängd.

4.2.3. Sats (Lebesgue). a) *Antag att $A \subset \mathbb{R}^2$ är kompakt, att ∂A är en nollmängd och att $f : A \to \mathbb{R}$ är kontinuerlig. Då existerar integralen*

$$\int_A f \;\mathrm{d}x\,\mathrm{d}y.$$

b) *Låt $f : A \to \mathbb{R}$ vara begränsad och $A \subset \mathbb{R}^2$ en kompakt nollmängd. Då är*

$$\int_A f \;\mathrm{d}x\,\mathrm{d}y = 0.$$

Bevis. Förbigås. □

4.2.4. Anmärkning. I punkt a) i Sats 4.2.3 behöver vi inte anta att f är begränsad, eftersom en kontinuerlig funktion i en kompakt mängd alltid är begränsad. I punkt b) behövs antagandet: Riemann-integralen kan endast anpassas på begränsade funktioner: Den är inte definierad för obegränsade funktioner, t.ex.

$$\int_0^1 \frac{1}{\sqrt{x}} \;\mathrm{d}x$$

är inte Riemann-integrerbar. Den är en så kallad oegentlig integral, som beräknas med gränsvärdet

$$\lim_{\varepsilon\to 0+}\int_\varepsilon^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x.$$

Vi återkommer till oegentliga integraler i kapitel 4.7.

Grundegenskaperna givna av Sats 4.1.8 generaliseras från rektanglar till godtyckliga mängder med hjälp av följande sats.

4.2.5. Sats. a) *Antag att funktionerna $f, g : A \to \mathbb{R}$ är integrerbara över den begränsade mängden $A \subset \mathbb{R}^2$ samt att $\lambda \in \mathbb{R}$. Då gäller*

a$_1$)

$$\int_A (f+g)\,\mathrm{d}x\,\mathrm{d}y = \int_A f\,\mathrm{d}x\,\mathrm{d}y + \int_A g\,\mathrm{d}x\,\mathrm{d}y,$$

a$_2$)

$$\int_A \lambda f\,\mathrm{d}x\,\mathrm{d}y = \lambda\int_A f\,\mathrm{d}x\,\mathrm{d}y,$$

a$_3$) *om $f(x,y) \le g(x,y)$ för varje $(x,y)\in D$, så är*

$$\int_A f\,\mathrm{d}x\,\mathrm{d}y \le \int_A g\,\mathrm{d}x\,\mathrm{d}y.$$

b) *Antag att funktionen $f : A \to \mathbb{R}$ är integrerbar över mängden A,*

$$A_1,\dots,A_k \subset A,$$

och att ∂A_i är en nollmängd för varje $i = 1,\dots,k$. Antag dessutom att $A_i \cap A_j \subset \partial A_i \cap \partial A_j$, då $i \neq j$, samt att

$$\bigcup_{i=1}^k A_i = A.$$

Då är

$$\sum_{i=1}^k \int_{A_i} f\,\mathrm{d}x\,\mathrm{d}y = \int_A f\,\mathrm{d}x\,\mathrm{d}y$$

Bevis. Förbigås. □

4.2.6. Anmärkning. Formeln i punkt b) i föregående sats är inte möjligast allmän, men räcker i de flesta tillämpningar.

Riemann-integralen gör det möjligt att bestämma arean av mängden $A \subset \mathbb{R}^2$. Låt $A \subset \mathbb{R}^2$ vara en begränsad mängd. Om integralen

$$\int_A \mathrm{d}x\,\mathrm{d}y \tag{4.2.7}$$

existerar, så kallas denna *Jordans mått* av A, eller i det tvådimensionella fallet *arean* av A, och betecknas $\mathrm{area}(A)$. Med integralen (4.2.7) menas alltså integralen

$$\int_D \chi_A \;\mathrm{d}x\,\mathrm{d}y,$$

där D är en kvardat som innehåller mängden A, och χ_A är den *karakteristiska funktionen* till mängden A, definierad genom

$$\chi_A(x) = \begin{cases} 1, & \text{då } x \in A, \\ 0, & \text{då } x \notin A, \end{cases}$$

Följande sats ger ett tillräckligt och nödvändigt villkor för $\mathrm{area}(A)$:s existens.

4.2.8. Sats (Lebesgue)**.** *En begränsad mängd $A \subset \mathbb{R}^2$ har en area exakt då, när ∂A är en nollmängd. Då är*

$$\mathrm{area}(A) = \int_{\overline{A}} 1 \;\mathrm{d}x\,\mathrm{d}y, \tag{4.2.9}$$

där $\overline{A} = A \cup \partial A$ är det slutna höljet till A.

Bevis. Förbigås. $\square$

Märk, att eftersom A är begränsad, så är även $\overline{A}$ begränsad. Eftersom $\overline{A}$ dessutom är sluten, så är den kompakt.

4.2.10. Anmärkning. Funktionen $f : \overline{A} \to \mathbb{R}$, $f(x) \equiv 1$, är kontinuerlig i den kompakta mängden $\overline{A}$ och enligt Sats 4.2.3 a) existerar integralen, eftersom vi har antagit, att ∂A är en nollmängd. Av Sats 4.2.3 b) och 4.2.5 b) följer att

$$\int_{\overline{A}} 1 \;\mathrm{d}x\,\mathrm{d}y = \int_{\partial A \setminus A} 1 \;\mathrm{d}x\,\mathrm{d}y + \int_A f \;\mathrm{d}x\,\mathrm{d}y = \int_A 1 \;\mathrm{d}x\,\mathrm{d}y.$$

Om $A = [a,b] \times [c,d]$, så har vi definierat att arean av A är $\mathrm{area}(A) = (b-a)(d-c)$. Eftersom en sträcka i $\mathbb{R}^2$ är en nollmängd, så följer det enkelt att ∂A, d.v.s. rektangelns rand, också är en nollmängd och med stöd av Sats 4.2.8 har rektangeln den ovannämnda arean. Av Sats 4.2.8 följer också, att mängden av innerpunkter till A, $A' = (a,b) \times (c,d)$ uppfyller ekvationen $\text{area}(A') = \text{area}(A)$, ty $\overline{A} = A$.

Ovan har vi behandlat integrering i planet, d.v.s. i $\mathbb{R}^2$. Det är skäl att noggrannare jämföra situationen med det bekanta endimensionella fallet. En begränsad mängd $A \subset \mathbb{R}$ har en Jordan-längd, med andra ord *Jordans endimensionella mått* $l(A)$, om Riemann-integralen

$$l(A) = \int_a^b \chi_A(t)\,\mathrm{d}t$$

existerar. Här är $[a,b]$ ett slutet intervall, som innehåller mängden A. Lebesgues sats 4.2.8 gäller alltså även i det endimensionella fallet, och säger, att en begränsad mängd $A \subset \mathbb{R}$ har en längd om och endast om randen ∂A till A är en endimensionell nollmängd. Detta betyder, att det för varje $\varepsilon > 0$ finns sådana intervall $[a_i, b_i]$, $i = 1, 2, \ldots$, att $\partial A \subset \bigcup_i [a_i, b_i]$ och

$$\sum_i (b_i - a_i) < \varepsilon.$$

Låt $\mathbb{Q}$ vara mängden av rationella tal och $A = \mathbb{Q} \cap [0,1]$. Då är A en begränsad mängd, men χ_A är inte Riemann-integrerbar i intervallet $[0,1]$. Detta är enkelt att se med hjälp av endimensionella över- och undersummor. Det följer naturligtvis också av Lebesgues sats 4.2.8, eftersom $\partial A = [0,1]$ och således är ∂A inte en endimensionell nollmängd. Att $\partial A = [0,1]$ är enkelt att se: Varje omgivning till $x \in [0,1]$ innehåller både punkter i mängden $A = \mathbb{Q} \cap [0,1]$ och mängden $[0,1] \setminus \mathbb{Q}$, varvid $x \in \partial A$ enligt definitionen på randen av mängden A. Vi lämnar beviset för att $[0,1]$ inte är en endimensionell nollmängd som övningsuppgift. Eftersom $\mathbb{Q}$ är en *numrerbar* mängd, inser vi lätt, att A själv är en endimensionell nollmängd. Mängden A är ändå inte kompakt, så vi kan inte använda den endimensionella versionen av Sats 4.2.3 b).

Problemet med begreppen längd, area och volym, som är definierade med Riemann-integraler, är att alltför få mängder i $\mathbb{R}$, $\mathbb{R}^2$ och $\mathbb{R}^3$ har ett Jordan-mått, med andra ord en längd, area eller volym. Den mått- och integralteori, som i huvudsak utvecklades av H. Lebesgue i början av 1900-talet, kan svara på dessa frågor bättre än den ovan presenterade teorin baserad på Riemann-integralen. Till exempel ger Lebesgues teori att längden av mängden $A = \mathbb{Q} \cap [0,1]$ är 0, fastän mängden A inte har någon längd enligt teorin ovan.

4.3. RIEMANNSUMMOR I EN GODTYCKLIG MÄNGD

Antag att D är en kompakt mängd i $\mathbb{R}^2$ och att ∂D är en nollmängd. Med delningen $\{D_i\}$ av mängden D i delarna $D_i,\ i = 1, \ldots, k$, avser vi det följande:

a) Mängderna D_i är kompakta,
b) Ränderna ∂D_i av mängderna D_i är nollmängder,
c) $D_i \cap D_j \subset \partial D_i \cap \partial D_j,\ i \neq j$, med andra ord är mängderna D_i och D_j väsentligt disjunkta och
d) $\bigcup D_i = D$.

Normen av delningen $\{D_i\}$ är $||\{D_i\}|| = \max\{d(D_i)\}$, där $d(D_i)$ är diametern för D, med andra ord är $d(D_i) = \sup\{|x-y| : x, y \in D_i\}$.

4.3.1. Sats. *Om $f : D \to \mathbb{R}$ är kontinuerlig, så är*

$$\sum_{i=1}^{k} f(x_i, y_i)\text{area}(D_i) \to \int_D f \, \mathrm{d}x \, \mathrm{d}y \tag{4.3.2}$$

då $||\{D_i\}|| \to 0$ och $(x_i, y_i) \in D_i,\ i = 1, 2, \ldots, k$.

Bevis. Beviset grundar sig på likformig kontinuitet hos f i en kompakt mängd. Vi förbigår detaljerna. □

4.3.3. Anmärkning. Sats 4.3.1 är i praktiken viktig, då vi försöker approximera värdet av integralen. Gränsvärdet (4.3.2) betyder det följande: För varje $\varepsilon > 0$ existerar det ett sådant $\delta > 0$, att då $\|\{D_i\}\| < \delta$, är

$$\left| \sum_{i=1}^{k} f(x_i, y_i)\,\text{area}(D_i) - \int_D f \, \mathrm{d}x \, \mathrm{d}y \right| < \varepsilon$$

där (x_i, y_i) är vilka som punkter som helst i D_i.

4.4. Beräkning av integraler i praktiken

Ofta har man som uppgift att räkna integralen av en kontinuerlig funktion $f : A \to \mathbb{R}$ över mängden

$$A = \{(x, y) \in \mathbb{R}^2 : g_1(x) \leq y \leq g_2(x),\ x \in [a, b]\},$$

där funktionerna $g_i : [a, b] \to \mathbb{R}$ är kontinuerliga, $i = 1, 2$, och $g_1(x) \leq g_2(x)$ för alla $x \in [a, b]$.

Observation 1. Mängden A är kompakt. Detta följer av att funktionerna är kontinuerliga. Vi lämnar beviset som övningsuppgift.

Observation 2. Randen ∂A av mängden A är en nollmängd. Detta följer av att:

a) A:s ändor är sträckor, och därmed är de nollmängder.
b) Grafen av funktionerna $g_i,\ i = 1, 2, \ldots$, är nollmängder. Detta grundar sig på en kontinuerlig funktions likformiga kontinuitet i ett slutet intervall. Vi lämnar detta som övningsuppgift.
c) Ett ändligt snitt av nollmängder är en nollmängd.

Genom att kombinera observationerna 1 och 2 med Sats 4.2.3, ser vi att integralen

$$\int_A f \; \mathrm{d}x \, \mathrm{d}y$$

existerar. Hur räknar man integraler i praktiken? Det är naturligt att använda itererade integraler. Vi väljer rektangeln

$$D = [a, b] \times [c, d]$$

där exempelvis

$$d = \max\{g_2(x) : x \in [a, b]\},\ c = \min\{g_1(x) : x \in [a, b]\}.$$

Då är $D \supset A$ och

$$\int_A f \; \mathrm{d}x \, \mathrm{d}y = \int_D \tilde{f} \; \mathrm{d}x \, \mathrm{d}y,$$

där

$$\tilde{f}(x) = \begin{cases} f(x) & , \text{då } x \in A \\ 0 & , \text{då } x \in D \backslash A. \end{cases}$$

Av Sats 4.1.7 följer nu att

$$\int_D \tilde{f}\ \mathrm{d}x\,\mathrm{d}y = \int_a^b \int_c^d \tilde{f}\ \mathrm{d}y\,\mathrm{d}x,$$

ifall de itererade integralerna existerar. En itererad integral räknas på följande sätt. Vi fixerar $x \in [a,b]$. Vi skall räkna

$$\int_c^d \tilde{f}(x,y)\ \mathrm{d}y.$$

Nu har funktionen $y \mapsto \tilde{f}(x,y)$ högst två diskontinuitetsställen i intervallet $[c,d]$, nämligen $g_1(x)$ och $g_2(x)$, och integralkalkyl med en variabel ger att

$$\int_c^d \tilde{f}(x,y)\ \mathrm{d}y = \int_{g_1(x)}^{g_2(x)} f(x,y)\ \mathrm{d}y,$$

ty $\tilde{f}(x,y) = 0$, då $y \in [c,d]\setminus[g_1(x), g_2(x)]$.

Av detta följer att

$$\int_A f\ \mathrm{d}x\,\mathrm{d}y = \int_D \tilde{f}\ \mathrm{d}x\,\mathrm{d}y = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\ \mathrm{d}y\,\mathrm{d}x,$$

eftersom det inte är svårt att se, att funktionen

$$x \mapsto \int_{g_1(x)}^{g_2(x)} f(x,y)\ \mathrm{d}y$$

är kontinuerlig i intervallet $[a,b]$.

4.4.1. Exempel. Vi skall räkna

$$I = \int_D (x^2 + y^2)\ \mathrm{d}x\,\mathrm{d}y,$$

där D är en triangel, vars hörn är punkterna $(0,0)$, $(1,0)$ och $(1,1)$. Vi betecknar $g_1(x) = 0$ och $g_2(x) = x$. Då kan vi skriva D i formen

$$\begin{aligned} D &= \{(x,y) \in \mathbb{R}^2 : g_1(x) \leq y \leq g_2(x),\ x \in [0,1]\} \\ &= \{(x,y) \in \mathbb{R}^2 : 0 \leq y \leq x,\ x \in [0,1]\}. \end{aligned}$$

Dessutom är $g_1,\ g_2$ och $f(x,y)=x^2+y^2$ kontinuerliga. Därmed gäller att

$$
\begin{aligned}
I &= \int_0^1 \left(\int_0^x (x^2+y^2)\ \mathrm{d}y \right)\ \mathrm{d}x = \int_0^1 \left(\Big/_{0}^{x} (x^2y + y^3/3) \right)\ \mathrm{d}x \\
&= \int_0^1 \underbrace{(x^3+x^3/3)}_{=4x^3/3}\ \mathrm{d}x = \Big/_{0}^{1} x^4/3 = 1/3.
\end{aligned}
$$

Det föregående kan vi även integrera i en annan ordning,

$$
I = \int_0^1 \left(\int_y^1 (x^2+y^2)\ \mathrm{d}x \right)\ \mathrm{d}y = \ldots = 1/3.
$$

I detta fall är uträkningarna aningen mera komplicerade.

4.4.2. Exempel. Vi bestämmer massan av planskivan $D \subset \mathbb{R}^2$. Vi antar, att skivans densitet i punkten (x,y) är $\rho(x,y)$ (kg/m^2). Märk, att densiteten inte nödvändigtvis är konstant. Nu kan vi tänka oss att vi approximerar massan av D med Riemann-summorna

$$
\sum \rho(x_i,y_i)\mathrm{area}(D_i),
$$

som går mot integralen

$$
m = \iint_D \rho(x,y)\ \mathrm{d}x\,\mathrm{d}y,
$$

då $\|\{D_i\}\| \to 0$, ifall den ifrågavarande integralen existerar. Denna definierar massan m. Om densiteten $\rho(x,y)=\rho_0$ är konstant, så är

$$
m = \iint_D \ \mathrm{d}x\,\mathrm{d}y = \rho_0\,\mathrm{area}(D).
$$

En planskiva D med konstant densitet har alltså en massa, om den har en area. Detta inträffar enligt Sats 4.2.8 exakt då, när D är begränsad och då ∂D är en nollmängd.

4.5. VARIABELBYTE I INTEGRALER

Den kanske mest användbara transformationsformeln för endimensionella Riemann-integraler är

$$
\int_a^b f(x)\ \mathrm{d}x = \int_\alpha^\beta f(g(t))g'(t)\,\mathrm{d}t, \tag{4.5.1}
$$

där $g(\alpha) = a,\ g(\beta) = b$ och alla ifrågavarande funktioner är kontinuerliga. Märk, att vi ovan inte kräver, att $g : [a, b] \to \mathbb{R}$ skulle definiera en bijektion till mängden $[a, b]$. Intervallet $[\alpha, \beta]$ kan dessutom vara "bakvänt". Beviset för formel (4.5.1) grundar sig på begreppet primitiv funktion och är enkelt. Antag att funktionen F är en primitiv funktion till funktionen f, alltså att $F'(x) = f(x)$. Med stöd av deriveringsregeln för sammansatta funktioner gäller då att

$$f(g(t))\, g'(t) = F'(g(t))\, g'(t) = (F \circ g)'(t),$$

och därmed är

$$\begin{aligned}\int_a^b f(x)\ \mathrm{d}x &= F(b) - F(a) = F(g(\beta)) - F(g(\alpha)) \\ &= \int_\alpha^\beta (F \circ g)'(t)\,\mathrm{d}t = \int_\alpha^\beta f(g(t))\, g'(t)\,\mathrm{d}t.\end{aligned}$$

Situationen är jobbigare, då $n \geq 2$. Teorin kan då inte grunda sig på begreppet primitiv funktion. Dessutom orsakar integreringsmängderna extra svårigheter.

Vi undersöker situationen, där $D,\ D' \subset \mathbb{R}^2$ är öppna, $\overline{D},\ \overline{D'}$ är kompakta och $\partial D,\ \partial D'$ är nollmängder. Vi antar att avbildningen $w : D \to D'$ är en bijektion (avbildningen w behöver inve vara en bijektion $\overline{D} \to \overline{D'}$) och att $w = (w_1, w_2)$ är en $\mathcal{C}^1(D)$-avbildning, med andra ord är koordinatfunktionerna $w_1, w_2 \in \mathcal{C}^1(D)$. Antag dessutom att $f : \overline{D'} \to \mathbb{R}$ är kontinuerlig, eller endast att $f : D' \to \mathbb{R}$ är Riemann-integrerbar.

4.5.2. Sats. *Med antagandena ovan gäller att*

$$\begin{aligned}&\iint_{D'} f(x, y)\ \mathrm{d}x\,\mathrm{d}y \\ &= \iint_D f(w_1(x, y), w_2(x, y))\, |\det w'(x, y)|\ \mathrm{d}x\,\mathrm{d}y \\ &= \int_D f(w)\, |\tau(w)|\ \mathrm{d}x\,\mathrm{d}y,\end{aligned}$$

där $w'(x,y) : \mathbb{R}^2 \to \mathbb{R}^2$ *är derivatan av avbildningen* w *i punkten* $(x,y) \in D$. *Avbildningen* $w'(x,y)$ *är en linjärfunktion, vars avbildningsmatris är*

$$\begin{bmatrix} \partial_1 w_1 & \partial_2 w_1 \\ \partial_1 w_2 & \partial_2 w_2 \end{bmatrix}$$

och

$$\tau(w)(x,y) = \det w'(x,y) = \begin{vmatrix} \partial_1 w_1 & \partial_2 w_1 \\ \partial_1 w_2 & \partial_2 w_2 \end{vmatrix}$$

är determinanten för avbildningsmatrisen av linjärfunktionen $w'(x,y)$, *d.v.s. Jacobis determinant till funktionen* w *i punkten* $(x,y) \in D$.

4.5.3. Anmärkningar. a) Sats 4.5.2 gäller inte helt i formen ovan, eftersom funktionen

$$(x,y) \mapsto f(w(x,y))\,|\tau(w)(x,y)| \tag{4.5.4}$$

inte behöver vara Riemann-integrerbar i mängden D. Orsaken till detta är att funktionen ifråga inte nödvändigtvis är begränsad. För varje D_i, där $D_i \subset D$, $\overline{D}_i$ är kompakt i D och ∂D_i är en nollmängd, gäller dock att funktionen (4.5.4) är Riemann-integrerbar i mängden D_i. Ifall

$$D_1 \subset D_2 \subset \ldots$$

och

$$\bigcup D_i = D,$$

så gäller dessutom att

$$\lim_{i\to\infty} \int_{D_i} f(w)\,|\tau(w)|\,\mathrm{d}x\,\mathrm{d}y$$

existerar och är oberoende av följden D_i. Detta gränsvärde betecknas

$$\int_D f(w)\,|\tau(w)|\,\mathrm{d}x\,\mathrm{d}y, \tag{4.5.5}$$

och det är detsamma som integralen i Sats 4.5.2. I själva verket är (4.5.5) ett exempel på en oegentlig integral, vilket vi återkommer till i kapitel 4.7.

b) Formeln i Sats 4.5.2 gäller också då vi istället för D har $\overline{D}$ och istället för D' har $\overline{D}'$. Orsaken till detta är att både ∂D och $\partial D'$ är nollmängder.

Bevis för Sats 4.5.2. Vi presenterar idén för beviset. Anmärkning 4.5.3 a) visar att detaljerna kräver ganska mycket arbete. Vi betecknar w:s inversa funktion $w^{-1} : D' \to D$. Låt $\{D_i\}$ vara delningen av mängden D, se kapitel 4.3. Nu gäller att

$$\int_{D'} f \, \mathrm{d}x \, \mathrm{d}y \approx \sum_{i=1}^{k} f(x_i', y_i') \operatorname{area}(D_i'),$$

där $D_i = w^{-1}(D_i')$, $w(D_i) = D_i'$ och $w(x_i, y_i) = (x_i', y_i')$. Ju mindre D_i är, desto bättre approximerar avbildningen $x \mapsto w(x_i, y_i) + w'(x_i, y_i)(x - (x_i, y_i))$ avbildningen w i mängden D_i. Detta beror på att w är deriverbar, och att det är fråga om lokal approximering av funktionen w med sin derivatafunktion.

Låt $L : \mathbb{R}^2 \to \mathbb{R}^2$ vara en lineärfunktion och $A = [a, b] \times [a, b] \subset \mathbb{R}^2$ en kvadrat. Då är $\operatorname{area}(A) = (b - a)^2$ arean av A, och med stöd av kunskaper i lineäralgebra uppfyller arean av kvadraten ekvationen

$$\operatorname{area}(L(A)) = |\det A| \operatorname{area}(A).$$

Detta gäller inte enbart för kvadrater, utan för alla mängder A, som har en area. Arean av bildmängden för lineärfunktionen erhåller vi från formeln ovan. Genom att tillämpa denna omvandlingsformel för areor på Riemannsumman av f och lineärfunktionen $L = w'(x_i, y_i)$, samt genom att ta i beaktande att

$$\operatorname{area}(D_i') = \operatorname{area}(w(D_i)) \approx \operatorname{area}(L(D_i)) = |\det w'(x_i, y_i)| \operatorname{area}(D_i)$$

får vi att

$$\begin{aligned}\sum_{i=1}^{k} f(x_i', y_i') \operatorname{area}(D_i') &\approx \sum_{i=1}^{k} f(w(x_i, y_i)) |\det w'(x_i, y_i)| \operatorname{area}(D_i) \\ &\approx \int_D f(w) |\det w'| \, \mathrm{d}x \, \mathrm{d}y.\end{aligned}$$

□

4.5.6. Exempel. Vi räknar

$$I = \iint_D \frac{x}{x + 2y} \, \mathrm{d}x \, \mathrm{d}y,$$

där D är rektangeln begränsad av linjerna $2x - y = 0$, $x + 2y = 2$, $x + 2y = 1$ och $2x - y = 2$. Avbildningen $w : \mathbb{R}^2 \to \mathbb{R}^2$, $w(x, y) =$

$(\underbrace{2x-y}_{u}, \underbrace{x+2y}_{v})$, är en lineärfunktion, och dess inversa funktion w^{-1} : $\mathbb{R}^2 \to \mathbb{R}^2$ har uttrycket

$$w^{-1}(u,v) = \left(\frac{1}{5}(2u+v), \frac{1}{5}(-u+2v)\right).$$

Detta får vi genom att lösa ekvationssystemet

$$\begin{cases} u = 2x - y \\ v = x + 2y \end{cases}$$

med avseende på x och y. Avbildningen w är vald så att den avbildar kvadraten D bijektivt till kvadraten D', vars hörnpunkter är $(0,1)$, $(0,2)$, $(2,1)$ och $(2,2)$. Om en lineärfunktion har en invers funktion, så är den inversa funktionen också en lineärfunktion. Avbildningsmatrisen av w^{-1} är

$$\begin{bmatrix} 2/5 & 1/5 \\ -1/5 & 2/5 \end{bmatrix},$$

varvid dess determinant är

$$\tau(w^{-1}) = \begin{vmatrix} 2/5 & 1/5 \\ -1/5 & 2/5 \end{vmatrix} = \frac{1}{5}.$$

Avbildningen w^{-1} är uppenbart en $\mathcal{C}^1$-avbildning i $\mathbb{R}^2$. Genom att definiera

$$f(x,y) = \frac{x}{x+2y}$$

och tillämpa Sats 4.5.2 på avbildningen w^{-1} får vi att

$$\begin{aligned} I &= \iint_D \frac{x}{x+2y}\,\mathrm{d}x\,\mathrm{d}y = \iint_{D'} f(w^{-1}(u,v))\,|\tau(w^{-1})(u,v)|\,\mathrm{d}u\,\mathrm{d}v \\ &= \iint_{D'} \frac{2u+v}{5v}\left|\frac{1}{5}\right|\mathrm{d}u\,\mathrm{d}v = \frac{1}{25}\iint_{D'}\left(\frac{2u}{v}+1\right)\mathrm{d}u\,\mathrm{d}v \\ &= \frac{1}{25}\int_0^2\left(\int_1^2\left(\frac{2u}{v}+1\right)\mathrm{d}v\right)\mathrm{d}u = \frac{1}{25}\int_0^2 (2u\ln 2+1)\,\mathrm{d}v \\ &= \frac{4\ln 2+2}{25}. \end{aligned}$$

I uppgiften hade vi nytta av att D avbildades till en enklare mängd D', i vilken det var enkelt att tillämpa itererade integraler.

4.5.7. Exempel. Vi skall beräkna integralen

$$I = \iint_D e^{-x^2-y^2} \, \mathrm{d}x \, \mathrm{d}y,$$

där $D = \{(x,y) \in \mathbb{R}^2 : a^2 \leq x^2 + y^2 \leq b^2\}$, $0 \leq a < b < \infty$, är avsnittet mellan två koncentriska cirklar. Mängden D är kompakt; då $a = 0$ är den en sluten skiva $\overline{B}(0,b)$. Vi granskar avbildningen $w : [a,b] \times [0,2\pi] \to \mathbb{R}^2$,

$$w(r,\varphi) = (x,y) = (r\cos\varphi, r\sin\varphi).$$

Det är alltså fråga om den *polära framställningen* av punkten (x,y). Avbildningen w avbildar den öppna rektangeln $(a,b) \times (0,2\pi)$ bijektivt till mängden

$$D' = \mathrm{int} D \backslash \{(x,y) \in \mathbb{R}^2 : y = 0, x > 0\}.$$

Avbildningen w är uppenbart en $\mathcal{C}^1$-avbildning i mängden $(a,b)\times(0,2\pi)$ och Jacobis determinant för uttrycket är

$$\tau(w)(r,\varphi) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r\cos^2\varphi + r\sin^2\varphi = r.$$

Eftersom funktionen som skall integreras är kontinuerlig i en kompakt mängd D, och den del som utelämnas från mängden D är en nollmängd, får vi att

$$\begin{aligned} I &= \iint_D e^{-x^2-y^2} \, \mathrm{d}x \, \mathrm{d}y = \iint_{D'} e^{-x^2-y^2} \, \mathrm{d}x \, \mathrm{d}y \\ &= \iint_{[a,b]\times[0,2\pi]} e^{-r^2} r \, \mathrm{d}r \, \mathrm{d}\varphi = \left(\int_0^{2\pi} \mathrm{d}\varphi \right) \left(\int_a^b e^{-r^2} r \, \mathrm{d}r \right) \\ &= 2\pi \Big/_{a}^{b} -\frac{1}{2} e^{-r^2} = \pi \left(e^{-a^2} - e^{-b^2} \right). \end{aligned}$$

Integralen ovan används inom statistik.

4.6. Beräkning av integraler med hjälp av nivåytor

Metoden lämpar sig speciellt för att beräkna integraler av typen

$$\int_D h(g(x,y)) \, \mathrm{d}x \, \mathrm{d}y.$$

Antag att $\Omega \subset \mathbb{R}^2$ är öppen, att $D \subset \Omega$ är kompakt, $g : \Omega \to \mathbb{R}$, $g \in \mathcal{C}^1(\Omega)$ är sådan, att $g(D) \subset [a, b]$ samt att $h : [a, b] \to \mathbb{R}$ är kontinuerlig. Vi betecknar

$$A(t) = \text{area}(G_t),\ t \in \mathbb{R},$$

där $G_t = \{(x, y) \in D : g(x, y) \leq t\}$. Vi antar att $A \in \mathcal{C}^1([a, b])$; detta förutsätter naturligtvis att $\text{area}(G_t)$ existerar.

4.6.1. Sats. *Antag att funktionerna f, g och A är som ovan. Då är*

$$\iint_D h(g(x, y))\ \mathrm{d}x\,\mathrm{d}y = \int_a^b h(t)A'(t)\,\mathrm{d}t.$$

Bevis. Vi skisserar upp beviset. Låt

$$a = t_0 < t_1 < \ldots < t_n = b$$

vara en delning av intervallet $[a, b]$. Då gäller att

$$h(g(x, y)) \approx h(t_i),$$

då $(x, y) \in G_{t_i} \backslash G_{t_{i-1}}$. Ju mindre differensen $t_i - t_{i-1}$ är, desto noggrannare är estimatet. Nu är

$$\begin{aligned}\iint_D h(g(x, y))\ \mathrm{d}x\,\mathrm{d}y &\approx \sum_i h(t_i)(A(t_i) - A(t_{i+1})) \\ &\approx \sum_i h(t_i)A'(t_i)(t_i - t_{i-1}) \approx \int_a^b h(t)A'(t)\,\mathrm{d}t.\end{aligned}$$

□

Satsen gäller även med svagare villkor, med då är den svår att formulera.

4.6.2. Exempel. Låt $D = \overline{B}(0, 1)$ och

$$I = \iint_D \sin(x^2 + y^2)\ \mathrm{d}x\,\mathrm{d}y.$$

Nu är $h(t) = \sin t,\ g(x, y) = x^2 + y^2$ och

$$A(t) = \text{area}(G_t) = \text{area}(\{(x, y) : x^2 + y^2 \leq t\}) = \pi t.$$

Vi har att

$$I = \int_0^1 h(t)A'(t)\,\mathrm{d}t = \pi \int_0^1 \sin t\,\mathrm{d}t = \pi \Big/_{0}^{1}(-\cos t) = \pi(1 - \cos 1).$$

Istället för intervallet $[0,1]$ kan vi använda ett bredare intervall $[a,b] \supset [0,1]$. Då är dock $A(t)$ konstant när $t \in [a,b] \backslash [0,1]$. Observera, att A inte är en $\mathcal{C}^1$-funktion i intervallet $[0,b]$, om $b > 1$. Sats 4.6.1 gäller även i detta specialfall.

Märk, att vi även kan räkna I med polära koordinater.

4.7. OEGENTLIGA INTEGRALER

I fallet $n = 1$ representerar t.ex.

$$\int_0^\infty f(x)\,\mathrm{d}x = \lim_{M\to\infty} \int_0^M f(x)\,\mathrm{d}x.$$

en så kallad *oegentlig integral*. Den existerar, ifall det både gäller att f är Riemann-integrerbar i varje intervall $[0,M]$ och att gränsvärdet ovan existerar. Situationen försvåras då vi övergår till flera dimensioner, ty de oegentliga integralerna i $\mathbb{R}$ grundar sig på de reella talens naturliga ordning. Något dylikt finns inte till förfogande i $\mathbb{R}^n$, då $n \geq 2$.

Antag att $A \subset \mathbb{R}^2$ är sluten eller öppen, att ∂A är en nollmängd, samt att $f : A \to \mathbb{R}$, $f \geq 0$ och att f är kontinuerlig. Vi antar inte att $\overline{A}$ är kompakt; A kan till exempel vara obegränsad. Dessutom är det ofta möjligt att ersätta kontinuiteten för f med Riemann-integrerbarhet.

Låt $(x_0, y_0) \in \partial A$, $(x,y) \in A \backslash \{(x_0,y_0)\}$ och $f(x,y) \to \infty$ då $(x,y) \to (x_0,y_0)$. Vi antar att Riemann-integralen

$$\int_D f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

existerar i de mängder D, som *uttömmer mängden* $A \backslash \{(x_0,y_0)\}$. Vi betecknar

$$I = \sup_D \int_D f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

där vi tar supremum över de ovannämnda mängderna D. Märk, att vi antog att $f \geq 0$, varvid också alla integraler antingen är positiva eller lika med noll. Med rätt tolkning kan (x_0,y_0) vara en "oändlighetspunkt".

Om $I = \infty$, d.v.s. mängden av integraler är inte begränsad ovanifrån, så säger vi att den *oegentliga integralen*

$$\int_A f(x,y)\ \mathrm{d}x\,\mathrm{d}y$$

divergerar i punkten (x_0, y_0). Om $I < \infty$, så säger vi att (den oegentliga) integralen

$$\int_A f(x,y)\ \mathrm{d}x\,\mathrm{d}y$$

konvergerar i punkten (x_0, y_0) och dess värde är I.

Följande härledning gäller: Antag att $D_1 \subset D_2 \subset \ldots$, att f är Riemann-integrerbar över varje mängd D_i, $D_i \subset A$ samt att

$$\bigcup D_i = A \backslash \{(x_0, y_0)\}.$$

Då är

$$I = \lim_{i\to\infty} \int_{D_i} f\ \mathrm{d}x\,\mathrm{d}y.$$

Detta är inte svårt att bevisa, ty $f \geq 0$. Vi går till väga på motsvarande sätt, om $f \leq 0$. Om f antar både positiva och negativa värden, så kan den uppdelas i formen

$$f = f^+ - f^-,$$

där $f^+ = \max\{f, 0\}$ och $f^- = -\min\{f, 0\}$. Vi definierar

$$I = \int_A f\ \mathrm{d}x\,\mathrm{d}y = \int_A f^+\ \mathrm{d}x\,\mathrm{d}y - \int_A f^-\ \mathrm{d}x\,\mathrm{d}y,$$

om båda oegentliga integralerna konvergerar.

I det föregående bestod "singulariteten" av en enda punkt (x_0, y_0). Singulariteten kan dock även vara en större mängd. Då går vi till väga på motsvarande sätt.

4.7.1. Exempel. Låt $A = (0,1) \times (0,1)$ vara en kvadrat, $(x,y) \in A$ och $f(x,y) = 1/\sqrt{x}$. Då är integralen

$$\int_A f(x,y)\ \mathrm{d}x\,\mathrm{d}y$$

inte definierad som en vanlig Riemann-integral, eftersom en Riemann-integrerbar funktion måste vara kontinuerlig. Vi uppfattar integralen som en oegentlig integral, varvid vi får att

$$\int_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \lim_{t\to 0+} \iint_{[t,1]\times[0,1]} \frac{1}{\sqrt{x}}\,\mathrm{d}x\,\mathrm{d}y$$
$$= \lim_{t\to 0+} \int_0^1 \int_t^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x\,\mathrm{d}y = \lim_{t\to 0+} \int_0^1 \left(\Big/_{t}^{1} 2\sqrt{x} \right)\,\mathrm{d}y$$
$$= \lim_{t\to 0+} \int_0^1 2 - 2\sqrt{t}\,\mathrm{d}y = \lim_{t\to 0+} 2(1-\sqrt{t}) = 2.$$

4.7.2. Exempel. Antag att

$$A = \{(x,y) \in \mathbb{R}^2 : x \geq 1,\ 0 \leq y \leq 1/x\}.$$

Vi skall bestämma

$$\iint_A \frac{\mathrm{d}x\,\mathrm{d}y}{x+y}.$$

Nu är funktionen

$$f(x,y) = \frac{1}{x+y}$$

kontinuerlig i A och $f(x,y) \geq 0$. Mängden A är obegränsad, och således kan integralen inte existera i enlighet med den vanliga Riemann-integralen. Singulariteten (x_0, y_0) måste tolkas befinna sig i "oändligheten".

Vi betecknar $A_t = A \cap ([1,t] \times [0,1])$ och definierar

$$\iint_A \frac{\mathrm{d}x\,\mathrm{d}y}{x+y} = \lim_{t\to\infty} \iint_{A_t} \frac{\mathrm{d}x\,\mathrm{d}y}{x+y}$$

om gränsvärdet existerar. Vi observerar, att gränsvärdet säkert existerar, eftersom funktionen

$$t \mapsto \iint_{A_t} \frac{\mathrm{d}x\,\mathrm{d}y}{x+y}$$

är definierad och växande. Det enda problemet som kan uppstå är att gränsvärdet är oändligt.

$$\iint_{A_t} \frac{\mathrm{d}x\,\mathrm{d}y}{x+y} = \int_1^t \left(\int_0^{1/x} \frac{\mathrm{d}y}{x+y} \right) \mathrm{d}x = \int_1^t \left(\Big/_{0}^{1/x} \ln(x+y) \right) \mathrm{d}x$$
$$= \int_1^t \ln\left(\frac{x+1/x}{x} \right) \mathrm{d}x = \int_1^t \ln\left(1 + \frac{1}{x^2} \right) \mathrm{d}x$$

Vi observerar att $\ln(1+s) \leq s$, varvid vi får estimatet

$$\int_1^t \ln\left(1+\frac{1}{x^2}\right)\,\mathrm{d}x \leq \int_1^t \frac{1}{x^2}\,\mathrm{d}x = \Big/_{1}^{t} -\frac{1}{x} = -\frac{1}{t}+1 \leq 1.$$

Således existerar gränsvärdet

$$\lim_{t\to\infty}\iint_{A_t}\frac{\mathrm{d}x\,\mathrm{d}y}{x+y}$$

och är ändligt.

Ett noggrannt värde för uttrycket

$$\lim_{t\to\infty}\int_1^t \ln\left(1+\frac{1}{x^2}\right)\,\mathrm{d}x$$

kan räknas med partiell integrering, ty den integrerade funktionens derivata är ett rationellt uttryck.

5. Integrering i flerdimensionella rum

5.1. Grundläggande definitioner

De grundläggande definitionerna är likadana som i föregående kapitel, då $n = 2$. Riemann-integralen av en begränsad funktion $f : D \to \mathbb{R}$ över "lådan"

$$D = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$$

kan definieras som i $\mathbb{R}^2$; även över- och undersummor kan användas på motsvarande sätt som i det tvådimensionella fallet. Speciellt gäller, att då de itererade integralerna existerar, så är

$$\int_D f \,\mathrm{d}x_1 \ldots \mathrm{d}x_n = \underbrace{\int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) \,\mathrm{d}x_n \ldots \mathrm{d}x_1}_{\text{itererade integraler}}$$

och om $f : D \to \mathbb{R}$ är kontinuerlig, så existerar både integralen

$$\int_D f \,\mathrm{d}x_1 \ldots \mathrm{d}x_n$$

och de itererade integralerna.

Antag sedan att $A \subset \mathbb{R}^n$ är en begränsad mängd och att $f : A \to \mathbb{R}$ är en begränsad funktion. Vi definierar

$$\tilde{f}(x) = \begin{cases} f(x), \text{ då } x \in A \\ 0, \text{ då } x \in \mathbb{R}^n \setminus A. \end{cases}$$

Vi väljer lådan D så att $D \supset A$, och definierar sedan

$$\int_A f \,\mathrm{d}x_1 \ldots \mathrm{d}x_n = \int_D \tilde{f} \,\mathrm{d}x_1 \ldots \mathrm{d}x_n,$$

ifall den senare integralen existerar.

Mängden $N \subset \mathbb{R}^n$ är en *Lebesgue nollmängd* i $\mathbb{R}^n$, med andra ord är dess mått noll med avseende på det *n-dimensionella Lebesgue måttet*, om det för varje $\varepsilon > 0$ existerar en sådan, möjligen ändlig, följd lådor $D_i,\ i = 1, 2, \ldots$, att

$$\sum_i \mathrm{vol}(D_i) < \varepsilon$$

och

$$N \subset \bigcup_{i=1}^{\infty} D_i.$$

Ovan är

$$\operatorname{vol}(D) = (b_1 - a_1)(b_2 - a_2) \dots (b_n - a_n)$$

volymen av lådan

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

Såsom i fallet $n = 2$ (Sata 4.2.3 a)), gäller följande sats.

5.1.1. Sats (Lebesgue). *Låt $A \subset \mathbb{R}^n$ vara kompakt, ∂A en nollmängd och $f : A \to \mathbb{R}$ kontinuerlig. Då existerar integralen*

$$\int_A f \, \mathrm{d}x_1 \dots \mathrm{d}x_n.$$

5.1.2. Anmärkning. Ränderna ∂A av "regelbundna" mängder A är vanligtvis nollmängder, men det lönar sig att vara noggrann.

Volymen av mängden $A \subset \mathbb{R}^n$ definieras med formeln

$$\operatorname{vol}(A) = \int_A 1 \, \mathrm{d}x_1 \dots \mathrm{d}x_n,$$

om integralen existerar. Oegentliga integraler behandlas på samma sätt som i fallet $n = 2$.

5.2. Beräkning av integraler i praktiken

För att beräkna integraler kan vi använda samma metoder som i fallet $n = 2$. Vi behandlar dessa med hjälp av exempel.

5.2.1. Exempel. Vi beräknar integralen

$$I = \int_A x_3 \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3$$

som en itererad integral, där A är den mängd som definieras av olikheterna

$$\begin{cases} x_1^2 + x_2^2 \leq x_3^2, \\ x_1^2 + x_2^2 + x_3^2 \leq 1, \\ x_3 \geq 0. \end{cases}$$

Nu är funktionen $f(x_1, x_2, x_3) = x_3$ kontinuerlig, mängden A är kompakt och ∂A är en nollmängd, fastän detta inte är helt enkelt att se. Således existerar integralen enligt Sats 5.1.1, varvid vi får att

$$
\begin{aligned}
I &= \iint_{B(0,\frac{1}{\sqrt{2}})} \left(\int_{\sqrt{x_1^2+x_2^2}}^{\sqrt{1-x_1^2-x_2^2}} x_3 \, \mathrm{d}x_3 \right) \mathrm{d}x_1 \, \mathrm{d}x_2 \\
&= \frac{1}{2} \iint_{B(0,\frac{1}{\sqrt{2}})} (1 - 2x_1^2 - 2x_2^2) \, \mathrm{d}x_1 \, \mathrm{d}x_2.
\end{aligned}
$$

Genom att övergå till polära koordinater kan den sista integralen skrivas i formen

$$
\begin{aligned}
&\frac{1}{2} \int_0^{2\pi} \int_0^{\frac{1}{\sqrt{2}}} (1 - 2r^2) r \, \mathrm{d}r \, \mathrm{d}\varphi \\
&= \pi \int_0^{\frac{1}{\sqrt{2}}} (1 - 2r^2) r \, \mathrm{d}r = \frac{\pi}{8}.
\end{aligned}
$$

5.2.2. Exempel. Vi kan behandla integraler med variabelbyte såsom i fallet $n = 2$. Antag att $D, D' \subset \mathbb{R}^n$ är kompakta, att ∂D, $\partial D'$ är nollmängder samt att w är en bijektiv $\mathcal{C}^1$-avbildning, $w : D \to D'$. Då gäller att

$$
\int_{D'} f \, \mathrm{d}x_1 \ldots \mathrm{d}x_n = \int_D f(w) |\tau(w)| \, \mathrm{d}x_1 \ldots \mathrm{d}x_n,
$$

där $\tau(w)$ är Jacobianen av avbildningen w, d.v.s. Jacobis determinant

$$
\tau(w) = \det w' = \begin{vmatrix} \partial_1 w_1 & \cdots & \partial_n w_1 \\ \vdots & \ddots & \vdots \\ \partial_1 w_n & \cdots & \partial_n w_n \end{vmatrix},
$$

och $w_1, \ldots, w_n$ är koordinatfunktionerna till w. Vi undersöker som exempel på funktionen w den s.k. polära koordinattransformationen i rummet $\mathbb{R}^3$.

$$
\begin{cases}
x_1 = r \sin\theta \cos\varphi, \\
x_2 = r \sin\theta \sin\varphi, \\
x_3 = r \cos\theta.
\end{cases}
$$

Den geometriska tolkningen av detta är följande: $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$ representerar längden av vektorn $x = (x_1, x_2, x_3)$ och $\theta \in [0, \pi]$ är vinkeln mellan x och den positiva x_3-axeln. Vinkeln $\varphi \in [0, 2\pi)$ är vinkeln mellan vektorn $(x_1, x_2, 0)$ och den positiva x_1-axeln i x_1x_2-planet.

Antag att

$$w : [0, \infty) \times [0, 2\pi) \times [0, \pi] \to \mathbb{R}^3$$

är avbildningen definierad av formeln ovan, d.v.s. att $w(r, \varphi, \theta) = (x_1, x_2, x_3)$. Denna är inte en bijektion, men däremot är restriktionen av w,

$$w : (0, \infty) \times [0, 2\pi) \times [0, \pi] \to \mathbb{R}^3 \backslash \{0\},$$

en bijektion. Mängden $(0, \infty) \times [0, 2\pi) \times [0, \pi]$ är varken öppen eller sluten, men dess rand och bilden av randen är nollmängder. Enkel derivering ger att Jacobis determinant av funktionen w är

$$\begin{aligned}\tau(w)(r, \varphi, \theta) &= \det \begin{bmatrix} \sin\theta\cos\varphi & -r\sin\theta\sin\varphi & r\cos\theta\sin\varphi \\ \sin\theta\sin\varphi & r\sin\theta\cos\varphi & r\cos\theta\sin\varphi \\ \cos\theta & 0 & -r\sin\theta \end{bmatrix} \\ &= r^2\sin\theta.\end{aligned}$$

Att utveckla denna determinant lämnas som övningsuppgift.

5.2.3. Anmärkning. En motsvarande transformation förekommer även i $\mathbb{R}^n$. Då $n = 2$, är denna transformation den polära framställningen, vars Jacobi-determinant är f.

Vi räknar volymen av klotet $\overline{B}(0, R)$ genom att använda transformationen ovan. Enligt definitionen är

$$\begin{aligned}\mathrm{vol}(\overline{B}(0, R)) &= \int_{\overline{B}(0,R)} 1 \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3 \\ &= \int_0^R \int_0^\pi \int_0^{2\pi} 1 \cdot r^2 \sin\theta \, \mathrm{d}\varphi \, \mathrm{d}\theta \, \mathrm{d}r \\ &= 2\pi \left(\Big/_{0}^{\pi} - \cos\theta \right) \left(\Big/_{0}^{R} \frac{r^3}{3} \right) = \frac{4}{3}\pi R^3.\end{aligned}$$

5.2.4. Exempel. Vi beräknar integralen

$$I = \int_{B(0,R)} \frac{1}{1 + x_1^2 + x_2^2 + x_3^2} \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3.$$

som ett annat exempel på variabelbyte. Mängden $\partial B(0, R)$ är en nollmängd, så vi kan lika väl räkna integralen över mängden $\overline{B}(0, R)$. Funktionen

$$f(x_1, x_2, x_3) = \frac{1}{1 + |x|^2}$$

är kontinuerlig i $\mathbb{R}^3$, och därmed får vi att

$$\begin{aligned}
I &= \int_0^R \int_0^\pi \int_0^{2\pi} \underbrace{f(w(r,\varphi,\theta))}_{=\frac{1}{1+r^2}} \underbrace{|\tau(w)(r,\varphi,\theta)|}_{=r^2\sin\theta} \,\mathrm{d}\varphi \,\mathrm{d}\theta \,\mathrm{d}r \\
&= 2\pi \int_0^R \frac{r^2}{1+r^2} \left(\int_0^\pi \sin\theta \,\mathrm{d}\theta \right) \mathrm{d}r \\
&= 4\pi \int_0^R \underbrace{\frac{r^2}{1+r^2}}_{=1-\frac{1}{1+r^2}} \,\mathrm{d}r = 4\pi(R - \arctan R).
\end{aligned}$$

5.2.5. Exempel. Såsom i fallet $n = 2$, kan vi också räkna integraler med hjälp av nivåytor. I $\mathbb{R}^n$ antar formeln formen

$$\int_A h(f(x_1,\ldots,x_n)) \,\mathrm{d}x_1 \ldots \mathrm{d}x_n = \int_a^b h(t)V'(t) \,\mathrm{d}t,$$

där $V(t) = \mathrm{vol}\{(x_1,\ldots,x_n) \in A : f(x_1,\ldots,x_n) \le t\}$.

Vi räknar det föregående exemplet med denna teknik. Nu är $h(t) = 1/(1+t)$ och $f(x_1,x_2,x_3) = x_1^2 + x_2^2 + x_3^2$. Funktionen $h \circ f$ är alltså

$$(h \circ f)(x_1,x_2,x_3) = \frac{1}{1 + x_1^2 + x_2^2 + x_3^2}.$$

Om $f \le t$, så är $x_1^2 + x_2^2 + x_3^2 \le (\sqrt{t})^2$, och då $0 \le t \le R^2$, är

$$V(t) = \frac{4\pi}{3} t^{3/2},$$

ty volymen av ett klot med radien r är $4\pi r^3/3$. Genom att derivera får vi att $V'(t) = 2\pi\, t^{1/2}$ och därmed är

$$I = \int_0^{R^2} \frac{1}{1+t} V'(t) \,\mathrm{d}t = 2\pi \int_0^{R^2} \frac{t^{1/2}}{1+t} \,\mathrm{d}t.$$

Med substitutionen $t = u^2$ omvandlas integralen till formen

$$\begin{aligned}
I &= 4\pi \int_0^R \left(1 - \frac{1}{1+u^2}\right) \mathrm{d}u^2 \\
&= 4\pi \int_0^R \frac{u^2}{1+u^2} \,\mathrm{d}u = 4\pi(R - \arctan R).
\end{aligned}$$

5.2.6. Exempel. Tolkningen av oegentliga intergaler är också likadan som i planet. Vi bestämmer den oegentliga integralen

$$I = \int_{B(0,1)} \frac{1}{|x|^\alpha} \,\mathrm{d}x_1 \,\mathrm{d}x_2 \,\mathrm{d}x_3,$$

där $\alpha > 0$. Nu är funktionen

$$f(x) = \frac{1}{|x|^\alpha}$$

inte definierad i klotet $B(0,1)$ och inte heller är den begränsad i mängden $B(0,1)\setminus\{0\}$. En punkt, i detta fall origo 0, är en tredimensionell nollmängd. Därmed är

$$I = \int_{B(0,1)\setminus\{0\}} f \; \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3,$$

ifall den oegentliga integralen

$$\int_{B(0,1)\setminus\{0\}} f \; \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3$$

existerar. Vi observerar att $f \geq 0$. Den oegentliga integralen I uppfattas som gränsvärdet

$$I = \lim_{r\to 0+} \int_{B(0,1)\setminus B(0,r)} \frac{1}{|x|^\alpha} \; \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3 =: \lim_{r\to 0+} I_r.$$

Då $0 < r < 1$, är integralen I_r en vanlig Riemann-integral av en kontinuerlig funktion, och det lönar sig att räkna med polära koordinater:

$$\begin{aligned} I_r &= \int_r^1 \int_0^{2\pi} \int_{-\pi}^{\pi} \frac{1}{r^\alpha} r^2 \sin\theta \; \mathrm{d}\theta \, \mathrm{d}\varphi \, \mathrm{d}r \\ &= 4\pi \int_r^1 r^{2-\alpha} \, \mathrm{d}r. \end{aligned}$$

Antag att $\alpha \neq 3$. Då $r \to 0$ gäller det att

$$\frac{4\pi}{3-\alpha} \Big/_{\!r}^{\,1} r^{3-\alpha} = \frac{4\pi}{3-\alpha}(1 - r^{3-\alpha}) \to \begin{cases} \dfrac{4\pi}{3-\alpha}, \text{ då } \alpha < 3, \\ \infty, \text{ då } \alpha > 3. \end{cases}$$

Om $\alpha = 3$, så är

$$I_r = 4\pi \int_r^1 r^{-1} \, \mathrm{d}r = 4\pi \Big/_{\!r}^{\,1} \log r = -4\pi \log r \to \infty,$$

när $r \to 0$. Med andra ord konvergerar integralen I endast då $\alpha < 3$.

6. Integralformler

För den reella axeln $\mathbb{R}$ gäller följande sats.

6.0.7. Sats (Analysens grundsats)**.** *Antag att f är en $\mathcal{C}^1$-funktion i intervallet $[a, b]$. Då är*

$$f(b) - f(a) = \int_a^b f'(t)\,\mathrm{d}t.$$

Denna formel har flera motsvarigheter i $\mathbb{R}^n$. Många av dessa innehåller rätt så djupgående information.

6.1. Om kurvintegraler

Dessa förekommer i praktiken av två typer,

$$\int_\gamma f\ \mathrm{d}s$$

och

$$\int_\gamma F \cdot\ \mathrm{d}\overline{s};$$

i specialfall använder vi andra beteckingar för den senare typen. Inledningsvis behandlar vi integralen

$$\int_\gamma f\ \mathrm{d}s.$$

Antag att $\gamma : [a, b] \to \mathbb{R}^n$ är en (styckvis) deriverbar stig

$$\gamma(t) = (\gamma_1(t), \ldots, \gamma_n(t)),$$

där $\gamma_i \in \mathcal{C}^1([a, b])$, $\gamma([a, b]) \subset A \subset \mathbb{R}^n$ och $f : A \to \mathbb{R}$ är kontinuerlig.

6.1.1. Definition. Integralen av funktionen F längs stigen γ är

$$\int_\gamma f\ \mathrm{d}s = \int_a^b f(\gamma(t))\sqrt{\gamma_1'(t)^2 + \ldots + \gamma_n'(t)^2}\,\mathrm{d}t.$$

Integralen kallas en *kurvintegral.*

Kurvintegralen är en helt vanlig endimensionell Riemann-integral. Om $f = 1$, så är

$$\int_\gamma\ \mathrm{d}s = \int_a^b \sqrt{\gamma_1'(t)^2 + \ldots + \gamma_n'(t)^2}\,\mathrm{d}t$$

längden av stigen γ.

6.1.2. Anmärkning. Kurvintegralen är oberoende av parametriseringen av stigen γ i följande avseende: Låt $\varphi : [c,d] \to [a,b]$ vara en växande $\mathcal{C}^1$-funktion, $\varphi(c) = a$ och $\varphi(d) = b$. Vi betecknar $\tilde{\gamma}(t) = \gamma(\varphi(t))$, $t \in [c,d]$. Nu gäller att

$$\begin{aligned}\int_{\tilde{\gamma}} f \ \mathrm{d}s &= \int_c^d f(\tilde{\gamma}(t))\sqrt{\tilde{\gamma}_1'(t)^2 + \ldots + \tilde{\gamma}_n'(t)^2}\,\mathrm{d}t \\ &= \int_c^d f(\gamma(\varphi(t)))|\varphi'(t)|\sqrt{\gamma_1'(\varphi(t))^2 + \ldots + \gamma_n'(\varphi(t))^2}\,\mathrm{d}t \\ &= \int_a^b f(\gamma(t))\sqrt{\gamma_1'(t)^2 + \ldots + \gamma_n'(t)^2}\,\mathrm{d}t = \int_\gamma f \ \mathrm{d}s,\end{aligned}$$

där integralen över intervallet $[a,b]$ fås med endimensionell integraltransformation. Om φ är avtagande, så är resultatet detsamma, med andra ord ändras inte tecknet.

Integreringen är utförd i förhållande till stigens (kurvans) längd. Den geometriska tolkningen av kurvintegralen är arean av den "inhägnad", vars höjd i punkten $\gamma(t)$ är $f(\gamma(t))$.

6.1.3. Exempel. Låt $\gamma(t) = (\cos t, \sin t)$, $t \in [0,2\pi]$ och $f(x,y) \equiv 1$. Då är

$$\int_\gamma \mathrm{d}s = \int_0^{2\pi} \sqrt{\sin^2 t + \cos^2 t}\,\mathrm{d}t = \Big/_{0}^{2\pi} t = 2\pi$$

längden av enhetscirkelns båge, vilket alltså är detsamma som längden av stigen γ.

6.1.4. Exempel. Antag att $\gamma(t) = (\cos t, \sin t)$, $t \in [0,4\pi]$. Då är

$$\int_\gamma \mathrm{d}s = 4\pi,$$

vilket är två gånger längden av enhetscirkelns båge. Detta är naturligt med tanke på den geometriska tolkningen: Bågen har genomlöpts två gånger och "inhägnaden" är dubbel.

6.1.5. Exempel. Låt $\gamma(t) = (\cos t, -\sin t)$, $t \in [0,2\pi]$. Då är

$$\int_\gamma \mathrm{d}s = 2\pi,$$

eftersom riktningen, i vilken vi "färdas på bågen", inte inverkar på kurvintegralen.

Härnäst behandlar vi integralen

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s} = \int_\gamma F \cdot \, \mathrm{d}s.$$

Antag att γ är såsom ovan, $\gamma([a, b]) \subset A \subset \mathbb{R}^n$ och $F : A \to \mathbb{R}^n$. Vi antar (vanligtvis) att avbildningen F är kontinuerlig och den kallas ofta ett *vektorfält*. Eftersom $F : A \to \mathbb{R}^n$, så definierar F n stycken koordinatfunktioner, med andra ord är $F = (F_1, \ldots, F_n)$, där $F_i : A \to \mathbb{R}$.

6.1.6. Definition.

$$\begin{aligned}
\int_\gamma F \cdot \, \mathrm{d}\overline{s} &= \int_a^b F(\gamma(t)) \cdot \gamma'(t) \, \mathrm{d}t \\
&= \int_a^b \left(\sum_{i=1}^n F_i(\gamma(t)) \gamma_i'(t) \right) \mathrm{d}t.
\end{aligned}$$

I fysiken är vektorfältet F en kraft, och integralen är arbetet utfört av kraften F, då "den rörliga partikeln" γ rör sig från $\gamma(a)$ till $\gamma(b)$. Integralen ovan är fortfarande vanlig endimensionell Riemann-integral.

Integralen

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s}$$

är liksom integralen

$$\int_\gamma f \, \mathrm{d}s$$

också oberoende av parametriseringen för γ. Det som skiljer dessa åt är att den senare definierade byter tecken, om vi ändrar riktning för stigen γ. Med att ändra riktning för stigen γ menar vi det följande:

Antag att $\gamma : [a, b] \to \mathbb{R}^n$. Vi definierar en ny stig $-\gamma$ genom att beteckna $-\gamma(t) = \gamma(b - (t - a))$, $t \in [a, b]$. Nu färdas $-\gamma$ i motsatt riktning i förhållande till stigen γ, och en enkel uträkning visar att

$$\int_{-\gamma} F \cdot \, \mathrm{d}\overline{s} = -\int_\gamma F \cdot \, \mathrm{d}\overline{s}.$$

Detta fenomen är naturligt med tanke på den fysikaliska tolkningen av integralen

$$\int_\gamma F \cdot \,\mathrm{d}\overline{s}.$$

I själva verket är integralen rätt så oberoende av parametriseringen av stigen γ: Det räcker, att avbildningen $\varphi : [c,d] \to [a,b]$ är en $\mathcal{C}^1$-avbildning, för vilken det gäller att $\varphi(c) = a$ och $\varphi(d) = b$. Då ger stigen $\gamma \circ \varphi$ samma resultat som γ.

6.1.7. Exempel. Antag att $\alpha, \beta \in \mathbb{R}$, $F(x,y) = (\alpha, \beta)$ är en "konstant kraft" och $\gamma(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$. Nu är

$$\begin{aligned}\int_\gamma F \cdot \,\mathrm{d}\overline{s} &= \int_0^{2\pi} (\alpha,\beta)\cdot(-\sin t, \cos t)\,\mathrm{d}t \\ &= \int_0^{2\pi} (-\alpha \sin t + \beta \cos t)\,\mathrm{d}t = 0.\end{aligned}$$

6.1.8. Exempel. Låt F vara som ovan, men $\gamma(t) = (\cos t, \sin t)$, $t \in [0, \pi]$. Nu gäller att

$$\begin{aligned}\int_\gamma F \cdot \,\mathrm{d}\overline{s} &= \int_0^{\pi} (\alpha,\beta)\cdot(-\sin t, \cos t)\,\mathrm{d}t \\ &= \int_0^{\pi} (-\alpha \sin t + \beta \cos t)\,\mathrm{d}t \\ &= \Big/_{0}^{\pi} \alpha \cos t + \beta \sin t = \alpha(-1-1) + 0 = -2\alpha.\end{aligned}$$

Flera andra kurvintegraler förekommer i tillämpningar. Dessa kan dock reduceras till de föregående. Ofta är det endast fråga om beteckningsskiljaktigheter. En beteckning som ofta förekommer i litteraturen är

$$\int_\gamma F \cdot \overline{T}\,\mathrm{d}s = \int_\gamma F \cdot \,\mathrm{d}\overline{s},$$

där vi med $\overline{T}$ menar tangentvektorn till stigen γ, d.v.s. $\overline{T} = (\gamma_1', \ldots, \gamma_n')$.

Ofta används också beteckningen

$$\oint_\gamma F \cdot \,\mathrm{d}\overline{s}$$

och, då $n = 2$, beteckningen

$$\oint_{\partial D} F \cdot \,\mathrm{d}s.$$

Av dessa betyder den första endast att stigen $\gamma : [a,b] \to \mathbb{R}^n$, längs vilken vi integrerar, är *sluten*, alltså att $\gamma(a) = \gamma(b)$. Integralens definition är dock densamma, d.v.s.

$$\oint_\gamma F \cdot \, \mathrm{d}\overline{s} = \int_\gamma F \cdot \, \mathrm{d}\overline{s}.$$

Den andra integralen är aningen mera komplicerad. Låt $D \subset \mathbb{R}^2$ vara en sådan öppen mängd, att ∂D kan framställas (styckvis) med en regelbunden sluten $\mathcal{C}^1$-stig $\gamma : [a,b] \to \partial D$ så att

$$\partial D = \gamma([a,b]).$$

Att en stig är *regelbunden* betyder att $\gamma'(t) \neq 0$ för varje $t \in [a,b]$. Vi antar dessutom, att γ är *positivt orienterad*, d.v.s. att mängden D är på vänstra sidan om vektorn $\gamma'(t)$ i närheten av punkten $\gamma(t)$, och att γ genomlöper mängden D endast en gång. Nu betyder beteckningen

$$\oint_{\partial D} F \cdot \, \mathrm{d}\overline{s}$$

helt enkelt integralen

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s}.$$

Observera, att vi kan beräkna kurvintegralen

$$\oint_{\partial D} F \cdot \, \mathrm{d}\overline{s}$$

också i mera komplicerade fall. Randen ∂D till mängden D kan till exempel bestå av två disjunkta delar, som representeras av stigarna γ_1 och γ_2. Ett typiskt exempel av en sådan mängd är cirkelringen $B(0,2)\backslash\overline{B}(0,1)$. Då är

$$\oint_{\partial D} F \cdot \, \mathrm{d}\overline{s} = \int_{\gamma_1} F \cdot \, \mathrm{d}\overline{s} + \int_{\gamma_2} F \cdot \, \mathrm{d}\overline{s},$$

där vi måste vara noggranna i förhållande till orienteringen för stigarna γ_1 och γ_2.

6.1.9. Exempel. Låt $F(x,y) = (x,y)$ och $D = B(0,1)$. Vi skall bestämma

$$\oint_{\partial D} F \cdot \, \mathrm{d}\overline{s}.$$

Stigen $\gamma(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$ ger ∂D genomlöpt i positiv riktning. Dessutom genomlöper stigen mängden $B(0,1)$ endast en gång, varvid

$$\begin{aligned}\oint_{\partial D} F \cdot \, \mathrm{d}\bar{s} &= \int_0^{2\pi} F(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t \\ &= \int_0^{2\pi} (\cos t, \sin t) \cdot (-\sin t, \cos t)\,\mathrm{d}t \\ &= \int_0^{2\pi} (-\cos t \sin t + \sin t \cos t)\,\mathrm{d}t = 0.\end{aligned}$$

Geometriskt sett är integralens värde klart utan några uträkningar, ty vektorn $F(x,y)$ och tangenten av $\partial B(0,1)$, som representeras av vektorn $\gamma'(t)$, står vinkelrätt mot varandra i punkten $(x,y) \in \partial B(0,1)$, med andra ord är $F(\gamma(t)) \cdot \gamma'(t) = 0$ för varje $t \in [0, 2\pi]$.

6.1.10. Exempel. Låt $D = B(0,1) \backslash \overline{B}(0,1/2)$ och vektorfältet $F(x,y) = (x, 1)$. Vi skall bestämma

$$\oint_{\partial D} F \cdot \, \mathrm{d}\bar{s}.$$

Nu är D cirkelringen mellan skivorna $B(0,1)$ och $B(0,1/2)$. Dess rand består av två delar: Av cirkeln $\partial B(0,1)$ och av cirkeln $\partial B(0,1/2)$. Vi uttrycker cirkeln $\partial B(0,1)$ som en positivt orienterad sluten stig

$$\gamma_1(t) = (\cos t, \sin t),$$

$t \in [0, 2\pi]$ och cirkeln $\partial B(0,1/2)$ på motsvarande sätt som en positivt orienterad sluten stig

$$\gamma_2(t) = (1/2 \cos t, -1/2 \sin t) = 1/2\,(\cos t, -\sin t),$$

$t \in [0, 2\pi]$. Märk, att stigarna γ_1 och γ_2 färdas i motsatta riktningar, ty mängden D måste förbli på vänstra sidan om vektorn $\gamma_2'(t)$. Till exempel är $\gamma_2'(0) = 1/2(-\sin 0, -\cos 0) = (0, -1/2)$ riktad nedåt i punkten $\gamma_2(0) = (1/2, 0)$ och D är faktiskt på vänstra sidan om vektorn

$\gamma_2'(0)$. Vi tolkar alltså integralen i formen

$$\begin{aligned}
\oint_{\partial D} F\cdot\ \mathrm{d}\overline{s} &= \int_{\gamma_1} F\cdot\ \mathrm{d}\overline{s} + \int_{\gamma_2} F\cdot\ \mathrm{d}\overline{s}\\
&= \int_0^{2\pi} (\cos t, 1)\cdot(-\sin t, \cos t)\,\mathrm{d}t\\
&\qquad + \int_0^{2\pi} (1/2\cos t, 1)\cdot(-1/2\sin t, -1/2\cos t)\,\mathrm{d}t\\
&= \int_0^{2\pi} (-\cos t\sin t + \cos t)\,\mathrm{d}t\\
&\qquad + \int_0^{2\pi} (-1/4\cos t\sin t - 1/2\cos t)\,\mathrm{d}t\\
&= -5/4\int_0^{2\pi} \cos t\sin t\,\mathrm{d}t + 1/2\int_0^{2\pi}\cos t\,\mathrm{d}t\\
&= -5/8 \Big/_{0}^{2\pi} \sin^2 t - 1/2 \Big/_{0}^{2\pi} \sin t = 0 + 0 = 0.
\end{aligned}$$

Integralen

$$\oint_{\partial D} F\cdot\ \mathrm{d}s$$

är inte heller beroende av "parametriseringen" av ∂D. Situationen är densamma som i fallet med integralen

$$\int_\gamma F\cdot\ \mathrm{d}\overline{s}.$$

Analysens grundsats i en dimension,

$$f(b) - f(a) = \int_a^b f'(t)\,\mathrm{d}t,$$

har en motsvarighet med kurvintegraler i rummet $\mathbb{R}^n$. Låt $f : D \to \mathbb{R}$ vara en $\mathcal{C}^1(D)$-funktion och $\gamma : [a,b] \to D$ en $\mathcal{C}^1$-stig i mängden D. Då gäller att

$$f(\gamma(b)) - f(\gamma(a)) = \int_\gamma \nabla f\cdot\ \mathrm{d}\overline{s}. \tag{6.1.11}$$

Integralen på den högra sidan i ekvation (6.1.11) är i formen

$$\int_\gamma F\cdot\ \mathrm{d}\overline{s},$$

där vektorfältet $F = \nabla f$. Av formel (6.1.11) följer den endimensionella formeln, då vi väljer $\gamma : [a,b] \to [a,b]$, $\gamma(t) = t$, och märker, att det i fallet $n = 1$ gäller att $\nabla f = f'$.

För att bevisa formel (6.1.11) går vi till väga på följande sätt: Antag att $\varphi : [a,b] \to \mathbb{R}$ är en sammansatt funktion $\varphi = f \circ \gamma$. Nu är φ en $\mathcal{C}^1$-funktion i intervallet $[a,b]$ och det gäller att

$$\varphi'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t),$$

se Sats 2.5.11. Vi får att

$$\begin{aligned}\int_\gamma \nabla f \cdot \,\mathrm{d}s &= \int_a^b \nabla f(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t \\ &= \int_a^b \varphi'(t)\,\mathrm{d}t = \Big/_{a}^{b} \varphi(t) = \varphi(b) - \varphi(a) \\ &= f(\gamma(b)) - f(\gamma(a)),\end{aligned}$$

vilket är formel (6.1.11).

6.1.12. Exempel. Vi beräknar

$$\oint_{\partial D} F \cdot \,\mathrm{d}\overline{s} \tag{6.1.13}$$

då D är en kvadrat $(0,1) \times (0,1)$ och $F(x,y) = (0, y^2)$. Att beräkna integralen (6.1.13) verkar inte ha något samband med formel (6.1.11). Vi observerar dock, att $F(x,y) = \nabla f(x,y)$, där $f(x,y) = 1/3\, y^3$. Randen ∂D till kvadraten D representerar en styckvis regelbunden sluten $\mathcal{C}^1$-stig $\gamma : [0,4] \to \partial D$, som vi kan sätta ihop av fyra sträckor. En sådan stig kan användas istället för $\mathcal{C}^1$-stigen i teorin ovan. Därmed ger formel (6.1.11) att

$$\begin{aligned}\oint_{\partial D} F \cdot \,\mathrm{d}\overline{s} = \int_{\partial D} \nabla f \cdot \,\mathrm{d}\overline{s} = \int_\gamma \nabla f \cdot \,\mathrm{d}\overline{s} &= f(\gamma(4)) - f(\gamma(0)) \\ = f(\gamma(0)) - f(\gamma(0)) &= 0,\end{aligned}$$

eftersom γ är sluten, d.v.s. $\gamma(4) = \gamma(0)$. Vi upptäcker, att det essentiella i den föregående uträkningen var att $F = \nabla f$ för någon $\mathcal{C}^1$-funktion f. Som resultat får vi då alltid 0. Vi behandlar detta fall noggrannare i samband med potentialer, i kapitel 6.3.

6.2. Greens formel i planet

Detta är en annan motsvarighet till formeln

$$f(b) - f(a) = \int_a^b f'(t)\,\mathrm{d}t$$

i fallet $n = 2$. Formeln har också motsvarigheter då $n \geq 3$. Dessa formler är dock inte lika uppenbara som (6.1.11).

6.2.1. Sats (Greens formel)**.** *Antag att $\Omega \subset \mathbb{R}^2$ är öppen och låt $F = (F_1, F_2) : \Omega \to \mathbb{R}^2$ vara en $\mathcal{C}^1(\Omega)$-avbildning. Vi antar att D är öppen så, att $\overline{D} \subset \Omega$ och randen ∂D uttrycks av en (styckvis) regelbunden positivt orienterad stig. Då är*

$$\oint_{\partial D} F \cdot \,\mathrm{d}\overline{s} = \iint_D (\partial_1 F_2 - \partial_2 F_1)\,\mathrm{d}x\,\mathrm{d}y. \tag{6.2.2}$$

Skiss för beviset. Vi undersöker specialfallet, där randen ∂D till den öppna mängden D består av en "undre kurva", som framställs av grafen av funktionen $\varphi_1 : [a, b] \to \mathbb{R}$, och en "övre kurva", som framställs av grafen av funktionen $\varphi_2 : [a, b] \to \mathbb{R}$. Vi antar, att φ_1 och φ_2 är $\mathcal{C}^1$-funktioner och att $\varphi_1(t) \leq \varphi_2(t)$, $t \in [a, b]$. Nu uttrycks randen ∂D av en styckvis regelbunden stig γ, som består av två stigar

$$\gamma_1(t) = (t, \varphi_1(t)),\ \ \gamma_2(t) = (t, \varphi_2(t)),\ \ t \in [a, b].$$

Riktningen för stigen γ_2 är felaktig i förhållande till mängden D. Om den avbildas på en motsatt riktad stig $-\gamma_2$, får vi en stig, som har den öppna mängden D på sin vänstra sida. Genom att placera γ_1 och $-\gamma_2$ efter varandra, vilket förutsätter parameterbyte i stigen $-\gamma_2$, får vi en styckvis regelbunden $\mathcal{C}^1$-stig γ, som representerar mängden ∂D i rätt vridriktning. Det egentliga uttrycket för stigen γ är onödigt, vilket beror på att stigintegraler i formen

$$\int_\gamma F \cdot \,\mathrm{d}\overline{s}$$

är oberoende av parameterframställningen av γ.

Nu gäller att

$$\oint_{\partial D} F \cdot \, \mathrm{d}\overline{s} = \int_{\gamma} F \cdot \, \mathrm{d}\overline{s} = \int_{\gamma_1} F \cdot \, \mathrm{d}\overline{s} + \int_{-\gamma_2} F \cdot \, \mathrm{d}\overline{s}.$$

För beviset lönar det sig att uppdela vektorfältet F i komponenter

$$F = (F_1, F_2) = (F_1, 0) + (0, F_2),$$

varvid även integralen uppdelas i formen

$$\oint_{\partial D} F \cdot \, \mathrm{d}\overline{s} = \oint_{\partial D} (F_1, 0) \cdot \, \mathrm{d}\overline{s} + \oint_{\partial D} (0, F_2) \cdot \, \mathrm{d}\overline{s}$$

Vi behandlar först termen som innehåller F_1:

$$\begin{aligned}
\oint_{\partial D} (F_1, 0) \cdot \, \mathrm{d}\overline{s} &= \int_{\gamma_1} (F_1, 0) \cdot \, \mathrm{d}\overline{s} + \int_{-\gamma_2} (F_1, 0) \cdot \, \mathrm{d}\overline{s} \\
&= \int_{\gamma_1} (F_1, 0) \cdot \, \mathrm{d}\overline{s} - \int_{\gamma_2} (F_1, 0) \cdot \, \mathrm{d}\overline{s} \\
&= \int_a^b (F_1(\varphi_1), 0) \cdot (1, \varphi_1') \, \mathrm{d}t - \int_a^b (F_1(\varphi_2), 0) \cdot (1, \varphi_2') \, \mathrm{d}t \\
&= \int_a^b F_1(t, \varphi_1(t)) \, \mathrm{d}t - \int_a^b F_1(t, \varphi_2(t)) \, \mathrm{d}t \\
&= \int_a^b (F_1(t, \varphi_1(t)) - F_1(t, \varphi_2(t))) \, \mathrm{d}t.
\end{aligned}$$

Å andra sidan är

$$\begin{aligned}
\iint_D (-\partial_2 F_1) \, \mathrm{d}x \, \mathrm{d}y &= \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} (-\partial_2 F_1(x, y)) \, \mathrm{d}y \, \mathrm{d}x \\
&= \int_a^b \left(\Big/_{\varphi_1(x)}^{\varphi_2(x)} -F_1(x, y) \right) \, \mathrm{d}x \\
&= \int_a^b (F_1(x, \varphi_1(x)) - F_1(x, \varphi_2))) \, \mathrm{d}x,
\end{aligned}$$

och således är

$$\oint_{\partial D} (F_1, 0) \cdot \, \mathrm{d}\overline{s} = \iint_D (-\partial_2 F_1) \, \mathrm{d}x \, \mathrm{d}y.$$

På motsvarande sätt, genom att anta att ∂D i x-riktningen kan uttryckas med grafer av funktioner av typen $y \mapsto \psi(y)$, får vi att

$$\oint_{\partial D} (0, F_2) \cdot \, \mathrm{d}\overline{s} = \iint_D \partial_1 F_2 \, \mathrm{d}x \, \mathrm{d}y.$$

Påståendet är således bevisat, om D är av typen ovan. □

Formel (6.2.2) gäller också i betydligt mera komplicerade fall. Beviset kan lätt generaliseras till fallet, där det är möjligt att indela D i delar, i vilka vi kan använda metoden ovan. Kurvintegralerna över delarnas gemensamma ränder tar ut varandra, eftersom de färdas i motsatta riktningar.

Kom också ihåg, att formel (6.2.2) kan användas, fastän ∂D endast skulle vara styckvis regelbunden, t.ex. en rektangel. Dessutom gäller (6.2.2) fastän ∂D skulle bestå av flera slutna stigar. Däremot är regelbundenheten för F ett nödvändigt villkor.

6.2.3. Exempel. Vi bestämmer volymen av en cirkulär cylinder med hjälp av kurvintegraler. Låt $F : \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (0, hx)$, $h > 0$ och $D = B(0, r) \subset \mathbb{R}^2$. Nu gäller att

$$\begin{aligned}\iint_D (\partial_1 F_2 - \partial_2 F_1)\ \mathrm{d}x\,\mathrm{d}y &= \iint_{B(0,r)} (h - 0)\ \mathrm{d}x\,\mathrm{d}y \\ &= \iint_{B(0,r)} h\ \mathrm{d}x\,\mathrm{d}y = \pi r^2 h,\end{aligned}$$

vilket är volymen av cylindern $B(0, r) \times (0, h)$.

Med stöd av Greens formel är

$$\iint_D (\partial_1 F_2 - \partial_2 F_1)\ \mathrm{d}x\,\mathrm{d}y = \oint_{\partial D} F \cdot\ \mathrm{d}\overline{s}.$$

Randen ∂D representeras av stigen $\gamma(t) = (r\cos t, r\sin t)$, $t \in [0, 2\pi]$, vars färdriktning är åt höger, och dess derivata är $\gamma'(t) = (-r\sin t, r\cos t)$. Nu är

$$\begin{aligned}\oint_{\partial D} F \cdot\ \mathrm{d}\overline{s} &= \int_0^{2\pi} F(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t \\ &= \int_0^{2\pi} (0, hr\cos t) \cdot (-r\sin t, r\cos t)\,\mathrm{d}t \\ &= \int_0^{2\pi} hr^2 \cos^2 t\,\mathrm{d}t = hr^2 \int_0^{2\pi} \cos^2 t\,\mathrm{d}t \\ &= hr^2 \cdot 1/2 \int_0^{2\pi} (\cos^2 t + \sin^2 t)\,\mathrm{d}t \\ &= 1/2\, hr^2 \cdot 2\pi = hr^2\pi,\end{aligned}$$

vilket är volymen av cylindern, såsom det enligt Greens formel skall vara.

6.3. Exakta vektorfält och potentialer

Antag att $D \subset \mathbb{R}^2$ är en öppen mängd och att $F : D \to \mathbb{R}^2$ är en avbildning, d.v.s. ett vektorfält, $F(x) = (F_1(x), F_2(x))$, där $F_i : D \to \mathbb{R}$, $i = 1, 2$. Vi antar dessutom att F är kontinuerlig, d.v.s. att koordinatfunktionerna F_i är kontinuerliga.

6.3.1. Definition. Vektorfältet F är *exakt* om det existerar en sådan funktion $u : D \to \mathbb{R}$, $u \in \mathcal{C}^1(D)$, att $\nabla u = F$ i mängden D. Funktionen u kallas en *potential* till vektorfältet F.

I fysiken tolkas F som ett kraftfält och kallas *konservativt*, om det är exakt.

6.3.2. Lemma. a) *Om funktionen u är en potential till vektorfältet F, så är också funktionen $u + c$, $c \in \mathbb{R}$, en potential till vektorfältet F.*

b) *Om funktionen u är en potential till vektorfältet F och mängden D är sammanhängande, så är varje potential till vektorfältet F av formen $u + c$, $c \in \mathbb{R}$.*

Bevis. a) Detta är uppenbart, ty $\nabla(u + c) = \nabla u$.

b) Om v är en potiential till F i mängden D, så är

$$\nabla v = F = \nabla u,$$

av vilket det följer att

$$\nabla(v - u) = F - F = 0.$$

Därmed är $v - u$ en $\mathcal{C}^1$-funktion, vars gradient är 0, i mängden D. I en sammanhängande öppen mängd D följer av detta, att $v - u$ är konstant. Detta grundar sig på, att en sammanhängande öppen mängd i $\mathbb{R}^n$ även är *stigsammanhängande*, vilket i detta fall betyder, att om $x, y \in D$, så existerar det en sådan $\mathcal{C}^1$-stig $\gamma : [a, b] \to D$, att $\gamma(a) = x$ och $\gamma(b) = y$. Att mängden D är *sammanhängande* betyder att mängden D inte kan uttryckas som en union $D_1 \cup D_2$, där $D_1 \neq \emptyset \neq D_2$, $D_1 \cap D_2 = \emptyset$ och där D_1 och D_2 är öppna. Beviset för att en sammanhängande öppen delmängd i $\mathbb{R}^n$ är stigsammanhängande hör till kursen i topologi. Beviset är inte väldigt enkelt, speciellt då vi kräver att en $\mathcal{C}^1$-stig sammanbinder punkterna i D. Saken är dock intuitivt uppenbar.

Om det nu för varje $x, y \in D$ finns en $\mathcal{C}^1$-stig γ som sammanbinder dem i mängden D, så följer av formel (6.1.11) att $v-u$ är konstant i D:

$$(v-u)(y)-(v-u)(x)=\int_\gamma \nabla(v-u)\cdot\ \mathrm{d}\overline{s}=\int_\gamma 0\cdot\ \mathrm{d}\overline{s}=0.$$

Således antar $v-u$ samma värde i varje punkt i mängden D, d.v.s. den är konstant. □

6.3.3. Exempel. Låt $D=\mathbb{R}^2$ och $F(x,y)=(x,y)$. Vi undersöker huruvida F är exakt. Till detta måste vi finna en sådan $\mathcal{C}^1$-funktion u, att $\nabla u = F$, alltså att

$$\nabla u(x,y)=(x,y).$$

Detta är ekvivalent med att ekvationerna $\partial_1 u(x,y)=x$ och $\partial_2 u(x,y)=y$ håller. Om $\partial_1 u(x,y)=x$, så är $u(x,y)=x^2/2+\varphi(y)$, där $\varphi:\mathbb{R}\to\mathbb{R}$. Orsaken till detta är följande: För ett fixerat $y\in\mathbb{R}$ gäller att $\partial_1 u(x,y)=x$, av vilket det följer att $u(x,y)=x^2/2+C_y$ för någon konstant $C_y\in\mathbb{R}$ som beror av y. För varje $y\in\mathbb{R}$ är således en funktion $\varphi(y)=C_y$, $\varphi:\mathbb{R}\to\mathbb{R}$ definierad.

Å andra sidan är $\partial_2 u(x,y)=y$, och av samma orsak som ovan är $u(x,y)=y^2/2+\psi(x)$, där $\psi:\mathbb{R}\to\mathbb{R}$. Genom att välja $\varphi(y)=y^2/2$ och $\psi(x)=x^2/2$ får vi att $u(x,y)=x^2/2+y^2/2$ och vi observerar att $\nabla u(x,y)=(x,y)$. Därmed är vektorfältet F exakt.

Av Lemma 6.3.2 b) följer att varje potential för vektorfältet F kan skrivas i formen $x^2/2+y^2/2+c$.

6.3.4. Exempel. Antag att $D=\mathbb{R}^2$ och $F(x,y)=(x^2,xy)$. Vi undersöker om vektorfältet F har en potential. Om en sådan existerar, så är $\nabla u(x,y)=(x^2,xy)$ d.v.s. $\partial_1 u(x,y)=x^2$ och $\partial_2 u(x,y)=xy$.

Eftersom $\partial_1 u$ och $\partial_2 u$ tillhör klassen $\mathcal{C}^\infty(\mathbb{R}^2)$, är funktionen u i samma klass, och således är $\partial_1\partial_2 u(x,y)=\partial_2\partial_1 u(x,y)$. Det gäller dock att $\partial_1\partial_2 u(x,y)=y$ och $\partial_2\partial_1 u(x,y)=0$, varvid F inte kan ha en potential.

Härnäst använder vi ekvationen (6.1.11), då det i integralen på högra sidan förekommer ett vektorfält F istället för gradienten ∇f.

6.3.5. Sats. *Antag att $F : D \to \mathbb{R}^2$ är ett exakt vektorfält och låt $\gamma_i : [a_i, b_i] \to D$, $i = 1, 2$, vara (styckvis) $\mathcal{C}^1$-stigar i mängden D, med samma start- och ändpunkter, d.v.s. $\gamma_1(a_1) = \gamma_2(a_2)$ och $\gamma_1(b_1) = \gamma_2(b_2)$. Då är*

$$\int_{\gamma_1} F \cdot \, \mathrm{d}\overline{s} = \int_{\gamma_2} F \cdot \, \mathrm{d}\overline{s} = u(\gamma_1(b_1)) - u(\gamma_1(a_1)).$$

Bevis. Genom att ta isär den första integralen får vi med stöd av definitionen att

$$\int_{\gamma_1} F \cdot \, \mathrm{d}\overline{s} = \int_{a_1}^{b_1} \nabla u(\gamma_1(t)) \cdot \gamma_1'(t) \, \mathrm{d}t.$$

Vi utför variabelbytet $t \overset{\varphi}{\mapsto} u(\gamma_{1,1}(t), \gamma_{1,2}(t))$, varvid deriveringsregeln för sammansatta funktioner ger att

$$\varphi'(t) = \nabla u(\gamma_1(t)) \cdot \gamma'(t).$$

Således är

$$\begin{aligned}\oint_{\gamma_1} F \cdot \, \mathrm{d}\overline{s} &= \int_{a_1}^{b_1} \varphi'(t) \, \mathrm{d}t = \Big/_{a_1}^{b_1} \varphi(t) = \varphi(b_1) - \varphi(a_1) \\ &= u(\gamma_1(b_1)) - u(\gamma_1(a_1)) \\ &= u(\gamma_2(b_1)) - u(\gamma_2(a_1)) = \int_{\gamma_2} F \cdot \, \mathrm{d}\overline{s},\end{aligned}$$

där vi har använt antagandet. Den sista ekvationen får vi med samma uträkning som i fallet med stigen γ_1. $\square$

6.3.6. Korollarium. *Antag att $F : D \to \mathbb{R}^2$ är ett exakt vektorfält och att $\gamma : [a, b] \to D$ är en sluten (styckvis) $\mathcal{C}^1$-stig i D. Då är*

$$\oint_{\gamma} F \cdot \, \mathrm{d}\overline{s} = 0.$$

Bevis. Detta följer direkt av formeln i Sats 6.3.5,

$$\oint_{\gamma} F \cdot \, \mathrm{d}\overline{s} = u(\gamma(b)) - u(\gamma(a)) = u(\gamma(a)) - u(\gamma(a)) = 0.$$

$\square$

Den fysikaliska tolkningen av Sats 6.3.5 är följande. Då vi förflyttar oss i ett exakt (konservativt) vektorfält från punkten x_0 till punkten y_0, så är det utförda arbetet oberoende av den valda stigen.

6.3.7. Anmärkningar. a) Sats 6.3.5 har också en annan sida: Om integralen

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s}$$

endast beror på γ:s ändpunkter, så är F exakt, med andra ord har den en potential. Vi förbigår beviset.

b) Om F är ett $\mathcal{C}^1(D)$-vektorfält och F har en potential, så är

$$\partial_2 F_1 = \partial_1 F_2. \tag{6.3.8}$$

Detta följer av villkoret $\partial_1\partial_2 u(x, y) = \partial_2\partial_1 u(x, y)$. Ifall formel (6.3.8) inte gäller, så kan F inte ha en potential. Om (6.3.8) gäller, har F då en potential? Svaret till detta är komplicerat. Det är "ja", om D är en öppen mängd, som inte har "hål", med andra ord, ifall D är *enkelt sammanhängande*. Till exempel är skivan $B(x, r)$ enkelt sammanhängande, men den så kallade punkterade skivan $B(x, r)\backslash\{x\}$ är det inte. Vi hoppar över dessa granskningar.

6.3.9. Exempel. Låt $F : \mathbb{R}^2 \to \mathbb{R}^2$ vara ett vektorfält, $F(x, y) = (2x \cos y, -x^2 \sin y)$. Mängden $\mathbb{R}^2$ är uppenbart enkelt sammanhängande. Desutom är $F_1(x, y) = 2x \cos y$ och $F_2(x, y) = -x^2 \sin y$, varvid

$$\partial_2 F_1(x, y) = -2x \sin y = \partial_1 F_2(x, y).$$

Med stöd av anmärkning 6.3.7 b) har således vektorfältet F en potential. Vi bestämmer denna med metoden vi använde tidigare: $\partial_1 u = F_1 = 2x \cos y$, av vilket det följer att $u = x^2 \cos y + \varphi(y)$. Å andra sidan är $-x^2 \sin y + \varphi'(y) = \partial_2 u = F_2 = -x^2 \sin y$, av vilket det följer att $\varphi'(y) = 0$, d.v.s. $\varphi(y) = c$ är konstant. Som potential u får vi alltså $u = x^2 \cos y + c$. Att konstatera att korsderivatorna var lika stora var alltså onödigt, eftersom man på detta sätt kunde ha funnit potentialen med direkt räkning. Efteråt är det lätt att med derivering verifiera, att u verkligen är en potential för F.

6.3.10. Exempel. Antag att F är som i det föregående exemplet och att $\gamma(t) = (t, t^3)$, $t \in [0, 1]$. Vi bestämmer

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s}.$$

Detta är svårt att räkna direkt utgående från integralens definition. Det gäller dock att $\gamma(0) = (0, 0)$, $\gamma(1) = (1, 1)$ och att F har en potential u. Av Sats 6.3.5 följer att

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s} = u(1, 1) - u(0, 0) = 1 \cos 1 - 0 \cos 0 = \cos 1.$$

Motsvarande teori om potentialer för vektorfält förekommer även i fallet $n \geq 3$.

6.4. Ytintegraler

En stig är en kontinuerlig avbilning γ från intervallet $[a, b]$ till rummet $\mathbb{R}^n$. Stigens bild $\gamma([a, b]) \subset \mathbb{R}^n$ kallas en *kurva*. Av detta får vi lätt en sådan uppfattning, att kurvan vore någonting “endimensionellt”. Detta stämmer inte. Kurvan kan till exempel vara den slutna enhetskulan $\overline{B}(0, 1) \subset \mathbb{R}^n$. Vi behandlar dock inte teorin för kurvor vidare, utan vi övergår till att behandla *ytor* och begränsar oss till $\mathbb{R}^3$. För ytor gäller samma fenomen som för stigar. En parametriserad yta i $\mathbb{R}^3$ är också såsom stigen en avbildning, men definitionsmängden är nu en mängd i $\mathbb{R}^2$ istället för ett intervall och avbildningens värden är i $\mathbb{R}^3$.

Låt $D \subset \mathbb{R}^2$ vara en öppen mängd och $r : D \to \mathbb{R}^3$. Vi säger att r är en *parametriserad yta*, om r är kontinuerlig. Eftersom

$$r(x, y) = (r_1(x, y), r_2(x, y), r_3(x, y)) \in \mathbb{R}^3,$$

betyder kontinuiteten för avbildningen r, att alla avbildningar $r_i : D \to \mathbb{R}$, $i = 1, 2, 3$, är kontinuerliga.

Vi antar nu att r är en $\mathcal{C}^1(D)$-funktion, vilket betyder, att de realvärda funktionerna r_i, $i = 1, 2, 3$, är $\mathcal{C}^1(D)$-funktioner. Nu är det möjligt att definiera de partiella derivatorna till avbildningen r i punkten

$(x, y) \in D$ med formeln

$$\partial_i r(x, y) = (\partial_i r_1(x, y), \partial_i r_2(x, y), \partial_i r_3(x, y)),$$

där $i = 1, 2$. Märk, att de partiella derivatorna $\partial_1 r$ och $\partial_2 r$ är vektorer i $\mathbb{R}^3$.

Vi återkallar i minnet begreppet om derivatan av avbildningen r i punkten $(x_0, y_0) \in D$. Derivatan av avbildningen r i punkten $(x_0, y_0) \in D$ är en lineärfunktion $r'(x_0, y_0) : \mathbb{R}^2 \to \mathbb{R}^3$. Avbildningen r har en derivata i varje punkt $(x_0, y_0) \in D$, ty r är en $\mathcal{C}^1$-avbildning, se kapitel 2.6. Avbildningsmatrisen för lineärfunktionen $r'(x_0, y_0)$ är en 3×2-matris

$$\begin{bmatrix} \partial_1 r_1(x_0, y_0) & \partial_2 r_1(x_0, y_0) \\ \partial_1 r_2(x_0, y_0) & \partial_2 r_2(x_0, y_0) \\ \partial_1 r_3(x_0, y_0) & \partial_2 r_3(x_0, y_0) \end{bmatrix}.$$

I teorin om ytor grundar sig derivatans användning på att avbildningen

$$(x, y) \mapsto r(x_0, y_0) + r'(x_0, y_0)(x - x_0, y - y_0) \tag{6.4.1}$$

approximerar avbilningen $(x, y) \mapsto r(x, y)$ "noggrant" i närheten av punkten (x_0, y_0).

De partiella derivatorna $\partial_1 r(x_0, y_0)$ och $\partial_2 r(x_0, y_0)$ är vektorer i $\mathbb{R}^3$, och således spänner de upp ett parallellogram, vars area är

$$|\partial_1 r(x_0, y_0) \times \partial_2 r(x_0, y_0)|,$$

där vi använt oss av kryssprodukten med vektorer i $\mathbb{R}^3$. Med andra ord, om $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3) \in \mathbb{R}^3$, så är

$$a \times b = \begin{vmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix},$$

och genom att formellt utveckla determinanten får vi en framställning av vektorn $a \times b$ uttryckt med standardenhetsvektorerna e_1, e_2 och e_3. Arean av bildparallellogrammet för rektangeln i punkten (x_0, y_0), sidorna Δx och Δy, i den affina avbildningen (6.4.1), är

$$|\partial_1 r(x_0, y_0) \times \partial_2 r(x_0, y_0)| \Delta x \Delta y.$$

Eftersom derivatafunktionen (6.4.1) approximerar avbildningen r i närheten av punkten (x_0, y_0), är det naturligt att definiera arean av den parametriserade ytan r som

$$\text{area}(r(D)) = \iint_D |\partial_1 r(x,y) \times \partial_2 r(x,y)| \,\mathrm{d}x\,\mathrm{d}y, \tag{6.4.2}$$

ifall denna integral existerar. Vi kallar mängden $r(D)$ i $\mathbb{R}^3$ den *parametriserade ytan definierad av avbildningen r.*

I formel (6.4.2) kan vi skriva kryssprodukten som

$$\begin{aligned} D(x,y) &= |\partial_1 r(x,y) \times \partial_2 r(x,y)| \\ &= \left[\left(\frac{\partial(r_1, r_2)}{\partial(x,y)} \right)^2 + \left(\frac{\partial(r_2, r_3)}{\partial(x,y)} \right)^2 + \left(\frac{\partial(r_3, r_1)}{\partial(x,y)} \right)^2 \right]^{1/2}, \end{aligned}$$

där

$$\frac{\partial(r_i, r_j)}{\partial(x,y)} = \begin{vmatrix} \partial_1 r_i & \partial_2 r_i \\ \partial_1 r_j & \partial_2 r_j \end{vmatrix}$$

är en 2×2-determinant. Detta är en relativt lång men enkel uträkning, som vi hoppar över. Således får formel (6.4.2) formen

$$\text{area}(r(D)) = \iint_D D(x,y) \,\mathrm{d}x\,\mathrm{d}y. \tag{6.4.3}$$

Ett ofta förekommande fall är en "yta" i $\mathbb{R}^3$, som framställs av grafen av funktionen $f : D \to \mathbb{R}$, där $D \subset \mathbb{R}^2$ är öppen. Grafen av funktionen f är en parametriserad yta, som definieras av avbildningen $r : D \to \mathbb{R}^3$,

$$r(x,y) = (x, y, f(x,y)),$$

ty

$$r(D) = \{(x, y, f(x,y)) \in \mathbb{R}^3 : (x,y) \in D\}$$

är grafen av f. Med en enkel uträkning får vi i detta fall att funktionen $D(x,y)$ är

$$D(x,y) = \sqrt{1 + \partial_1 f(x,y)^2 + \partial_2 f(x,y)^2},$$

och således är uttrycket för arean av grafen av f

$$\text{area}(r(D)) = \iint_D \sqrt{1 + \partial_1 f(x,y)^2 + \partial_2 f(x,y)^2} \,\mathrm{d}x\,\mathrm{d}y. \tag{6.4.4}$$

6.4.5. Exempel. Låt $D = (0,1) \times (0,1)$ vara en kvadrat och $f : D \to \mathbb{R}$, $f(x,y) = 1$. Det är uppenbart, att arean av grafen av f borde vara 1, med andra ord arean av kvadraten D. I detta fall är $\partial_1 f(x,y) = 0 = \partial_2 f(x,y)$ och formel (6.4.4) ger det önskade resultatet 1.

6.4.6. Exempel. Antag att D är som ovan och att $f : D \to \mathbb{R}$, $f(x,y) = x$, med andra ord har kvadraten D ett "snett tak", som bildas av grafen av f. Grafen av avbildningen f är en kvadrat i $\mathbb{R}^3$, och från bilden ser vi direkt, att dess sidor är 1 och $\sqrt{2}$. Därmed är dess area $\sqrt{2}$. I detta fall är $r(x,y) = (x,y,x)$ och

$$D(x,y) = (1 + 1^2 + 0^2)^{1/2} = \sqrt{2}$$

samt

$$\operatorname{area}(r(D)) = \iint_D \sqrt{2}\, \mathrm{d}x\, \mathrm{d}y = \sqrt{2},$$

som det också borde vara.

6.4.7. Exempel. Vi bestämmer arean av en rotationskropp. Antag att $f : [a,b] \to \mathbb{R}$ är en $\mathcal{C}^1$-funktion, $f \geq 0$. Då vi roterar grafen av funktionen f i $\mathbb{R}^2$ runt x_1-axeln, får vi en rotationskropp, vars graf vi vill bestämma.

Ytan av rotationskroppen kan på ett naturligt sätt uttryckas som en parametriserad yta $r : [a,b] \times [0, 2\pi) \to \mathbb{R}^3$, där koordinatfunktionerna till r är

$$\begin{aligned} r_1(x,\varphi) &= x, \\ r_2(x,\varphi) &= f(x)\cos\varphi, \\ r_3(x,\varphi) &= f(x)\sin\varphi. \end{aligned}$$

Rita en bild och skissera rotationsytan.

Mängden $D = [a,b] \times [0, 2\pi)$ är inte öppen, men den kan i dessa undersökningar ersättas med en öppen mängd $(a,b) \times (0, 2\pi)$ eller med en kompakt mängd $\overline{D} = [a, b] \times [0, 2\pi]$. En direkt uträkning ger att

$$\begin{aligned}
\frac{\partial(r_1, r_2)}{\partial(x, y)} &= -f(x)\sin\varphi, \\
\frac{\partial(r_2, r_3)}{\partial(x, y)} &= f(x)\, f'(x), \\
\frac{\partial(r_1, r_3)}{\partial(x, y)} &= f(x)\cos\varphi.
\end{aligned}$$

Antag att $S = r(D)$. Nu är

$$\begin{aligned}
\operatorname{area}(S) &= \iint_D D(x, y)\ \mathrm{d}x\,\mathrm{d}y \\
&= \iint_D \sqrt{f(x)^2\sin^2\varphi + f(x)^2 f'(x)^2 + f(x)^2\cos^2\varphi}\,\mathrm{d}x\,\mathrm{d}\varphi \\
&= \iint_D f(x)\sqrt{1 + f'(x)^2}\,\mathrm{d}x\,\mathrm{d}\varphi \\
&= 2\pi\int_a^b f(x)\sqrt{1 + f'(x)^2}\ \mathrm{d}x,
\end{aligned}$$

vilket är arean av rotationskroppen.

Kurvintegralen

$$\int_\gamma f\ \mathrm{d}s$$

motsvaras av en *ytintegral*

$$\iint_{r(D)} f\ \mathrm{d}S$$

över den parametriserade ytan r. Detta är naturligt att definiera med formeln

$$\iint_S f\ \mathrm{d}S = \iint_{r(D)} f\ \mathrm{d}S = \iint_D f(r(x, y))D(x, y)\ \mathrm{d}x\,\mathrm{d}y.$$

Då antar vi på ett naturligt sätt, att avbildningen

$$(x, y) \mapsto (f \circ r)(x, y)D(x, y)$$

är Riemann-integrerbar i mängden D. Funktionen f är en avbildning $f : A \to \mathbb{R}$, där $A \subset \mathbb{R}^3$ och $A \supset r(D)$, och således är den sammansatta avbildningen $f \circ r$ definierad i mängden D. I ytintegraler skriver vi ofta $\mathrm{d}A$ ($A =$ area) istället för $\mathrm{d}S$.

6.4.8. Exempel. Antag att $D \subset \mathbb{R}^2$,

$$D = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1,\ -1 \leq y \leq 1\}$$

och att $r(x, y) = (x, y, x^2 + y)$. Med andra ord är ytan $r(D)$ den graf, som bestäms av funktionen $(x, y) \mapsto x^2 + y$. Vi beräknar

$$\iint_{r(D)} x_1 \, \mathrm{d}S.$$

Vi har följande specialfall: $f(x_1, x_2, x_3) = x_1$, $f(r(x, y)) = x$ och $D(x, y) = \sqrt{1 + (2x)^2 + 1} = \sqrt{2}\,\sqrt{1 + 2x^2}$. Nu är

$$\begin{aligned}
\iint_{r(D)} x_1 \, \mathrm{d}S &= \iint_D x\,\sqrt{2}\,\sqrt{1 + 2x^2} \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_{-1}^{1} \left(\int_0^1 \sqrt{2}\,\sqrt{1 + 2x^2} \, \mathrm{d}x \right) \mathrm{d}y \\
&= \int_{-1}^{1} \left(\Big/_{\!0}^{1} \frac{2\sqrt{2}}{4 \cdot 3} \left(1 + 2x^2\right)^{3/2} \right) \mathrm{d}y \\
&= \int_{-1}^{1} \sqrt{2}/6\,(3^{3/2} - 1) \, \mathrm{d}y \\
&= \frac{\sqrt{2}}{3} \left(3^{3/2} - 1\right) = \sqrt{2}(\sqrt{3} - 1/3).
\end{aligned}$$

Egenskaperna hos ytintegraler är essentiellt likadana som hos kurvintegraler, fastän parameterbytet är en mera komplicerad process.

En annan viktig ytintegral är vektorfältets *flöde* genom den parametriserade ytan $S = r(D)$. Antag att $A \subset \mathbb{R}^3$, $F : A \to \mathbb{R}^3$ är en avbildning, d.v.s. ett vektorfält i $\mathbb{R}^3$ och att $A \supset r(D)$.

6.4.9. Definition. Flödet för vektorfältet F genom ytan $S = r(D)$ definieras med formeln

$$\begin{aligned}
\int_S F \cdot \, \mathrm{d}\overline{A} &= \iint_S F \cdot \, \mathrm{d}\overline{S} \\
&= \iint_D F(r(x, y)) \cdot (\partial_1 r(x, y) \times \partial_2 r(x, y)) \, \mathrm{d}x \, \mathrm{d}y, \qquad (6.4.10)
\end{aligned}$$

förutsatt att den senare Riemann-integralen existerar.

För att förstå integralen (6.4.10) undersöker vi situationen, där $\partial_1 r(x, y) \times \partial_2 r(x, y) \neq 0$ för varje $(x, y) \in D$. Då står vektorn $\partial_1 r(x, y) \times$

$\partial_2 r(x,y)$ vinkelrätt mot vektorerna $\partial_1 r(x,y)$ och $\partial_2 r(x,y)$. Vektorerna $\partial_1 r(x,y)$ och $\partial_2 r(x,y)$ spänner upp tangentplanet till ytan $S = r(D)$ i punkten $r(x,y)$ och vektorn $\partial_1 r(x,y) \times \partial_2 r(x,y)$ står vinkelrätt mot dessa vektorer. Denna är därmed normalen för ytan S i punkten $r(x,y)$. Enhetsnormalen $n(x,y)$ i denna punkt är

$$n(x,y) = \overline{n}(x,y) = \frac{\partial_1 r(x,y) \times \partial_2 r(x,y)}{|\partial_1 r(x,y) \times \partial_2 r(x,y)|}$$

Nu kan integralen

$$\iint_S F \cdot \, \mathrm{d}\overline{S}$$

reduceras till en ytintegral på följande sätt:

$$\begin{aligned}
\iint_S F \cdot \, \mathrm{d}\overline{S} &= \iint_D F(r(x,y)) \cdot (\partial_1 r \times \partial_2 r) \, \mathrm{d}x \, \mathrm{d}y \\
&= \iint_D F(r(x,y)) \cdot \underbrace{\frac{\partial_1 r \times \partial_2 r}{|\partial_1 r \times \partial_2 r|}}_{n(x,y)} \underbrace{|\partial_1 r \times \partial_2 r|}_{=D(x,y)} \, \mathrm{d}x \, \mathrm{d}y \\
&= \iint_D F(r(x,y)) \cdot n(x,y) \, D(x,y) \, \mathrm{d}x \, \mathrm{d}y \\
&= \iint_S F \cdot \overline{n} \, \mathrm{d}S.
\end{aligned}$$

Vi undersöker specialfallet, där den parametriserade ytan $S = r(D)$ är grafen av funktionen $(x,y) \mapsto f(x,y)$. Nu är $r(x,y) = (x, y, f(x,y))$ och

$$\begin{aligned}
\partial_1 r \times \partial_2 r &= \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & \partial_1 f \\ 0 & 1 & \partial_2 f \end{vmatrix} \\
&= -\partial_1 f \, e_1 - \partial_2 f \, e_2 + e_3.
\end{aligned}$$

Vi får att enhetsnormalen i punkten $r(x,y)$ är

$$n = \overline{n} = \frac{\partial_1 r \times \partial_2 r}{|\partial_1 r \times \partial_2 r|} = \frac{-\partial_1 f \, e_1 - \partial_2 f \, e_2 + e_3}{\sqrt{\partial_1 f^2 + \partial_2 f^2 + 1}}.$$

I detta fall är å andra sidan $D(x,y) = \sqrt{\partial_1 f^2 + \partial_2 f^2 + 1}$, vilket betyder att om $F = (F_1, F_2, F_3)$ är ett kontinuerligt vektorfält, så får integralen uttrycket

$$
\begin{aligned}
\iint_S F \cdot \, \mathrm{d}\overline{S} &= \iint_D F(r(x,y)) \cdot n(x,y) D(x,y) \, \mathrm{d}x \, \mathrm{d}y \\
&= \iint_D (F_1(x,y,f(x,y))(-\partial_1 f(x,y)) \\
&\quad + F_2(x,y,f(x,y))(-\partial_2 f(x,y)) \\
&\quad + F_3(x,y,f(x,y))) \, \mathrm{d}x \, \mathrm{d}y.
\end{aligned} \tag{6.4.11}
$$

6.4.12. Exempel. Antag att $F(x_1, x_2, x_3) = (x_1, x_2, x_3)$, d.v.s. $F(x) = x$. Vi skall beräkna

$$\iint_S F \cdot \, \mathrm{d}\overline{S}$$

då ytan S är cirkeln $x_1^2 + x_2^2 < 25$ på planet $x_3 = 12$. Ytan S kan framställas som grafen av funktionen $f : B(0,5) \to \mathbb{R}$, $f(x,y) = 12$. Normalen till ytan S är $e_3 = \overline{n}$, varvid $F \cdot \overline{n} = x_3$. Nu är

$$
\begin{aligned}
\iint_S F \cdot \, \mathrm{d}\overline{S} = \iint_S F \cdot \, \mathrm{d}S &= \iint_{B(0,5)} x_3 \, \mathrm{d}x \, \mathrm{d}y \\
&= 12 \iint_{B(0,5)} \mathrm{d}x \, \mathrm{d}y = 12\pi \cdot 5^2 = 300\pi.
\end{aligned}
$$

Observera, att det i detta fall gäller att $\partial_1 f = 0 = \partial_2 f$.

Integralen

$$\iint_S F \cdot \overline{n} \, \mathrm{d}S$$

är väsentligt oberoende av ytans parametrisering. Att tolka parametriseringen är däremot inte helt enkelt. Dessutom kan integralen tolkas på två sätt beroende på hur vi valt riktningen för normalen $\overline{n}$.

I fall av den parametriserade ytan definierar avbildningen r normalens riktning på ett naturligt sätt, förutsatt att $\partial_1 r(x,y) \times \partial_2 r(x,y) \neq 0$ i punkten (x,y). Om den parametriserade ytan är framställd från grafen av funktionen $f : D \to \mathbb{R}$ med formeln $r(x,y) = (x,y,f(x,y))$, så blir normalens riktning automatiskt bestämd såsom ovan. Detta ger att ytan bestämd av grafen av f får en så kallad positiv orientering, som har behandlats i samband med kurvintegraler i kapitel 6.1. Generellt är det problematiskt att orientera ytor i $\mathbb{R}^3$. Till exempel är det inte möjligt att hitta en vettig entydig riktning för enhetsnormaler till *Möbius band*, eftersom normalens rikting blir motsatt, då vi genomlöper bandet en gång.

6.5. Gauss sats

Gauss sats, m.a.o. divergenssatsen, ger totalflödet för vektorfältet F genom en sluten yta S i $\mathbb{R}^3$, d.v.s. integralen

$$\iint_S F \cdot \overline{n}\, \mathrm{d}S,$$

uttryckt som integralen av vektorfältets divergens $\nabla \cdot F$ över den av ytan begränsade öppna mängden. Vi gör följande antaganden:

(i) $\Omega \subset \mathbb{R}^3$ är öppen och $F : \Omega \to \mathbb{R}^3$, ett $\mathcal{C}^1$-vektorfält
(ii) $D \subset \mathbb{R}^3$ öppen, $\overline{D}$ kompakt, och $\overline{D} \subset \Omega$,
(iii) ∂D består av en ändlig mängd parametriserade $\mathcal{C}^1$-ytor.

Vektorfältets *divergens* $\nabla \cdot F$ är avbildningen $\nabla \cdot F : \Omega \to \mathbb{R}$,

$$(\nabla \cdot F)(x, y, z) = \partial_1 F_1(x, y, z) + \partial_2 F_2(x, y, z) + \partial_3 F_3(x, y, z),$$

där $F = (F_1, F_2, F_3)$.

Vi antar dessutom, att ∂D är parametriserad så, att normalen $\overline{n}$ i varje punkt pekar utåt från mängden D, d.v.s. $\overline{n}$ är ytans *ytternormal*. I själva verket är de ovan beskrivna definitionerna inte tillräckligt exakta för att beskriva situationen. I praktiska situationer är det relativt klart, vad vi menar med att ∂D består av en ändlig mängd parametriserade ytor, vars normal $\overline{n}$ dessutom är väl definierad. Saken skulle dock kräva noggrannare fördjupade undersökningar. I praktiska situationer räcker det att $\mathcal{C}^1$-ytorna $r_i : D_i \to \partial D$, $i = 1, \ldots, k$, är sådana, att varje r_i är en injektiv avbildning, $r_i(D_i) \cap r_j(D_j) = \emptyset$, $i \neq j$, och att unionen av mängderna $r_i(D_i)$ täcker randen ∂D till den öppna mängden D med undantag av en mängd, som med avseende på ytintegreringen är en nollmängd. Injektiviteten för avbildningen r_i motsvaras hos kurvintegralen

$$\oint_{\partial D} F \cdot\, \mathrm{d}\overline{s},$$

av att stigen som framställer ∂D får genomlöpa mängden D endast en gång. Då kan integralen

$$\iint_{\partial D} F \cdot \overline{n} \, \mathrm{d}S$$

uppdelas i integraler över ytdelarna $S_i = r_i(D_i)$ på följande sätt:

$$\begin{aligned}\iint_{\partial D} F \cdot \overline{n} \, \mathrm{d}S &= \sum_{i=1}^{k} \iint_{S_i} F \cdot \overline{n} \, \mathrm{d}S \\ &= \sum_{i=1}^{k} \iint_{D_i} F(r_i(x,y)) \cdot \overline{n}(x,y) D_i(x,y) \, \mathrm{d}x \, \mathrm{d}y,\end{aligned}$$

där $D_i(x,y)$ är det $D(x,y)$ som bestäms av avbildningen r_i, se kapitel 6.4.

6.5.1. Sats (Gauss sats, d.v.s. divergenssatsen). *I den ovan beskrivna situationen gäller att*

$$\iint_{\partial D} F \cdot \overline{n} \, \mathrm{d}S = \iiint_{D} \nabla \cdot F \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z.$$

Beviset för Gauss sats grundar sig på samma ide som Greens sats. Vi förbigår detaljerna. Ofta betecknar vi

$$\iint_{\partial D} F \cdot \overline{n} \, \mathrm{d}S = \oiint_{\partial D} F \cdot \overline{n} \, \mathrm{d}S.$$

Gauss sats 6.5.1 har en motsvarighet i alla dimensioner $n \geq 1$. Dess motsvarighet i fallet $n = 1$ är den bekanta Analysens grundsats

$$f(b) - f(a) = \int_a^b f'(t) \, \mathrm{d}t.$$

På högra sidan förekommer divergensen för f, d.v.s. integralen av f', och den vänstra sidan $f(b) - f(a)$ kan också tolkas som integralen över randen till intervallet $[a, b]$, som endast består av punkterna a och b.

I fallet $n = 2$ är motsvarigheten för Gauss sats i själva verket Greens sats 6.2.1, som vi tidigare behandlade. Vi undersöker denna tolkning närmare. Antag att D är en öppen mängd i planet och att ∂D representeras av en positivt orienterad sluten $\mathcal{C}^1$-stig $\gamma : [a, b] \to \partial D$, $\gamma(t) = (\gamma_1(t), \gamma_2(t))$, som dessutom är regelbunden, d.v.s. $\gamma'(t) = (\gamma_1'(t), \gamma_2'(t)) \neq (0, 0)$ för varje $t \in [a, b]$. På naturligt sätt antar vi också, att γ genomlöper mängden D endast en gång. Vektorn $\gamma'(t)$

representerar tangentvektorn till ∂D i punkten $\gamma(t)$ och enhetsvektorn som står vinkelrätt mot denna, d.v.s. enhetsnormalen till ∂D, är

$$\overline{n}(t) = \frac{1}{|\gamma'(t)|}(\gamma_2'(t), -\gamma_1'(t)).$$

Eftersom γ är positivt orienterad i förhållande till mängden D, är $\overline{n}(t)$ riktad utåt från mängden D i punkten $\gamma(t)$. Den är alltså ytternormalen för ∂D i punkten $\gamma(t)$. Antag nu att $F : \Omega \to \mathbb{R}^2$ är ett tvådimensionellt $\mathcal{C}^1$-vektorfält i den öppna mängden Ω, som innehåller det slutna höljet $\overline{D}$. Situationen är nu likadan som i Gauss sats, förutom att $n = 2$. Tolkningen av integralen

$$\int_{\partial D} F \cdot \overline{n} \, \mathrm{d}s = \oint_{\partial D} F \cdot \overline{n} \, \mathrm{d}s$$

är en kurvintegral, se kapitel 6.1,

$$\int_{\partial D} F \cdot \overline{n} \, \mathrm{d}s = \int_{\gamma} F \cdot \overline{n} \, \mathrm{d}s$$

längs stigen γ.

Av Greens sats 6.2.1 följer nu Gauss sats 6.5.1, med andra ord divergenssatsen då $n = 2$, med följande resonemang. Vi definierar ett nytt vektorfält $\tilde{F} : \Omega \to \mathbb{R}^2$, $\tilde{F} = (-F_2, F_1)$. Då är

$$\begin{aligned}
\int_{\partial D} F \cdot \overline{n} \, \mathrm{d}s &= \int_{\gamma} F \cdot \overline{n} \, \mathrm{d}s \\
&= \int_a^b F(\gamma(t)) \cdot \overline{n}(t) |\gamma'(t)| \, \mathrm{d}t \\
&= \int_a^b \frac{F_1(\gamma(t))\gamma_2'(t) - F_2(\gamma(t))\gamma_1'(t)}{|\gamma'(t)|} |\gamma'(t)| \, \mathrm{d}t \\
&= \int_a^b \tilde{F}(\gamma(t)) \cdot \gamma'(t) \, \mathrm{d}t = \int_{\gamma} \tilde{F} \cdot \, \mathrm{d}\overline{S} \\
&= \oint_{\partial D} \tilde{F} \cdot \, \mathrm{d}\overline{S} = \iint_D \left(\partial_1 \tilde{F}_2 - \partial_2 \tilde{F}_1\right) \, \mathrm{d}x \, \mathrm{d}y \\
&= \iint_D (\partial_1 F_1 + \partial_2 F_2) \, \mathrm{d}x \, \mathrm{d}y = \iint_D \nabla \cdot F \, \mathrm{d}x \, \mathrm{d}y,
\end{aligned}$$

där vi använt Greens sats för vektorfältet $\tilde{F}$. Formeln som erhållits på detta sätt är enklare att komma ihåg än Greens sats. Man måste dock vara uppmärksam med normalens riktning. I själva verket är Greens sats och divergenssatsen i planet ekvivalenta.

6.5.2. Exempel. Vi härleder den så kallade *Arkhimedes lag*. Vi antar, att en vätskas yta är ett x_1x_2-plan och att den positiva x_3-axeln pekar nedåt, med andra ord fyller vätskan halvrummet $x_3 \geq 0$ och x_3 uppger vätskans djup. Trycket i punkten $(x_1, x_2, x_3) \in \mathbb{R}^3$ inne i vätskan är

$$p = p_o + g\rho\, x_3 = p(x_1, x_2, x_3),$$

där lufttrycket p_o, vätskans densitet ρ och gravitationskonstanten g är konstanter.

För enkelhetens skull antar vi, att kroppen D är helt nedsäkt i vätskan, varvid vi kan använda uttrycket ovan för trycket p. Låt A vara en "liten" del av ytan ∂D till kroppen D. Trycket i vätskan pressar då A ungefär med kraften

$$\overline{F} = -\text{area}(A) p(x_1, x_2, x_3) \overline{n}(x_1, x_2, x_3),$$

där $(x_1, x_2, x_3) \in A$ och $\overline{n}(x_1, x_2, x_3)$ är ytternormalen till ytan ∂D i punkten (x_1, x_2, x_3). Idén här är förstås, att A är så liten och ∂D tillräckligt regelbunden, att p och $\overline{n}$ inte nämnvärt förändras då vi förflyttar oss i A från en punkt till en annan.

Totalkraften $\overline{F}$, som orsakas av trycket på kroppen D, får vi genom att uppdela ∂D i små disjunkta delar A_i, $i = 1, \ldots, k$, och genom att addera de krafter som verkar på dessa. Vi får att

$$\overline{F} = \sum_{i=1}^{k} F_i = -\sum_{i=1}^{k} \text{area}(A_i) p(x_1^i, x_2^i, x_3^i) \overline{n}(x_1^i, x_2^i, x_3^i),$$

där punkten $x^i = (x_1^i, x_2^i, x_3^i) \in A_i$. Detta är ett uttryck som är likt en Riemannsumma, och då lämpliga villkor för regelbundenhet gäller, kan vi tolka totalkraften som integralen

$$\overline{F} = -\iint\!\!\!\!\!\!\bigcirc_{\partial D} p\overline{n}\, \mathrm{d}S, \tag{6.5.3}$$

eftersom det ovan är fråga om en Riemannsumma för denna integral.

Integralen (6.5.3) har inte förekommit tidigare. Det är nämligen en vektorvärd integral. Komponenterna för denna vektor är vanliga integraler

$$\begin{aligned}\overline{F} &= (e_1 \cdot \overline{F}, e_2 \cdot \overline{F}, e_3 \cdot \overline{F}) \\ &= -(\oiint_{\partial D} pe_1 \cdot \overline{n}\ \mathrm{d}S, \oiint_{\partial D} pe_2 \cdot \overline{n}\ \mathrm{d}S, \oiint_{\partial D} pe_3 \cdot \overline{n}\ \mathrm{d}S).\end{aligned}$$

Vi beräknar komponenternas integraler med hjälp av divergenssatsen. I den sista integralen undersöker vi vektorfältet $F^* = (0, 0, p)$, varvid

$$\oiint_{\partial D} pe_3 \cdot \overline{n}\ \mathrm{d}S = \oiint_{\partial D} F^* \cdot \overline{n}\ \mathrm{d}S,$$

och eftersom $\nabla \cdot F^* = \partial_3 p = g\rho$, får vi med stöd av Gauss sats, alltså divergenssatsen, att

$$\begin{aligned}-\oiint_{\partial D} pe_3 \cdot \overline{n}\ \mathrm{d}S &= -\iiint_D \nabla \cdot F^*\ \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3 \\ &= -g\rho \mathrm{vol}(D).\end{aligned}$$

Absolutbeloppet av detta är massan av den vätska som D undanträngt. Den tredje komponenten i kraften är negativ, eftersom kraften är riktad uppåt.

De två första komponenterna i $\overline{F}$ kan bestämmas ännu enklare med hjälp av divergenssatsen. För att bestämma den första komponenten $e_1 \cdot \overline{F}$ väljer vi vektorfältet $F^* = (p, 0, 0)$, varvid $\nabla \cdot F^* = 0$, och

$$\begin{aligned}-\oiint_{\partial D} pe_1 \cdot \overline{n}\ \mathrm{d}S &= \oiint_{\partial D} F^* \cdot \overline{n}\ \mathrm{d}S \\ &= \iiint_D \nabla \cdot F^*\ \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3 = 0.\end{aligned}$$

På motsvarande sätt får vi att

$$\oiint_{\partial D} pe_2 \cdot \overline{n}\ \mathrm{d}S = 0.$$

Därmed blir uttrycket för totalkraften $\overline{F}$

$$F = (0, 0, -g\rho \mathrm{vol}(D)),$$

vilket är Arkhimedes lag.

6.6. Stokes sats

Greens formel sammankopplar kurvintegralen längs randen till ett planområde, med integralen över ett planområde. Denna formel har en motsvarighet för krökta ytor i $\mathbb{R}^3$, där kurvintegralen är tagen längs stigen γ, som bildar randen till den tvådimensionella ytan S. Integralen över planområdet ersätts med integralen över ytan S. Formeln som erhållits på detta sätt kallas Stokes sats.

Vi behöver några definitioner. Antag att $\Omega \subset \mathbb{R}^2$ är öppen, $D \subset \mathbb{R}^2$ är öppen, $\overline{D}$ är kompakt och $\overline{D} \subset \Omega$, samt att $r : \Omega \to \mathbb{R}^3$ är en regelbunden $\mathcal{C}^1$-yta. Vi antar dessutom att avbildningen r är en injektion och vi undersöker ytan $S = r(D)$, som bestäms av restriktionen $r|_D$ av avbildningen r till mängden D. Normalvektorn $\overline{n}$ till ytan S är i denna situation väl definierad även på randen av ytan S, d.v.s. i mängden $r(\partial D)$, med hjälp av avbildningen r. Vi antar vidare, att randen av S representerar en (styckvis) regelbunden sluten stig $\gamma : [a, b] \to r(\partial D)$, som genomlöper randen endast en gång. Vi kräver ytterligare, att γ är positivt orienterad i förhållande till ytan S. Detta betyder, att vektorn

$$\overline{n}(\gamma(t)) \times \gamma'(t),$$

som står vinkelrätt mot vektorerna $\overline{n}(\gamma(t))$ och $\gamma'(t)$ i punkten $\gamma(t)$ på randen $r(\partial D)$, är riktad mot ytan S för varje $t \in [a, b]$. Detta är naturligtvis beroende av riktningen, som bestäms av avbildningen r, för vektorn $\overline{n}$.

6.6.1. Sats (Stokes)**.** *Antag att situationen är som ovan och att $F : \Omega \to \mathbb{R}^3$ är ett $\mathcal{C}^1$-vektorfält. Då gäller att*

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s} = \iint_S (\nabla \times F) \cdot \overline{n} \, \mathrm{d}S \tag{6.6.2}$$

Tolkningen av integralerna i formel (6.6.2) är klar bortsett från den vektorvärda avbildningen $\nabla \times F$. Vi skriver ut integralerna. Stigen $\gamma : [a, b] \to r(\partial D) \subset \mathbb{R}^3$ har tre koordinatfunktioner, med andra ord är $\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$. Dessutom är $\gamma'(t) = (\gamma_1'(t), \gamma_2'(t), \gamma_3'(t))$. Vektorfältet $F : \Omega \to \mathbb{R}^3$ har också tre koordinatfunktioner, med andra ord är $F = (F_1, F_2, F_3)$. Integralen

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s}$$

kan alltså skrivas i formen

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s} = \int_a^b (F_1(\gamma(t))\,\gamma_1'(t) + F_2(\gamma(t))\,\gamma_2'(t) + F_3(\gamma(t))\,\gamma_3'(t))\,\mathrm{d}t.$$

Integralen på högra sidan i formel (6.6.2) är aningen mera komplicerad. I den förekommer *rotationen för vektorfältet* F, $\nabla \times F$, vars tolkning är följande. Rotationen $\nabla \times F : \Omega \to \mathbb{R}^3$ är ett vektorfält, som erhålls ur vektorfältet F på följande sätt:

$$\begin{aligned}
\nabla \times F &= \begin{vmatrix} e_1 & e_2 & e_3 \\ \partial_1 & \partial_2 & \partial_3 \\ F_1 & F_2 & F_3 \end{vmatrix} \\
&= (\partial_2 F_3 - \partial_3 F_2)e_1 + (\partial_3 F_1 - \partial_1 F_3)e_2 + (\partial_1 F_2 - \partial_2 F_1)e_3 \\
&= (\partial_2 F_3 + \partial_3 F_2, \partial_3 F_1 - \partial_1 F_3, \partial_1 F_2 - \partial_2 F_1).
\end{aligned}$$

Det lönar sig att komma ihåg detta i den formella determinantformen ovan. Integralen

$$\iint_S (\nabla \times F) \cdot \overline{n} \, \mathrm{d}S$$

kan nu uppdelas på följande sätt:

$$\begin{aligned}
\iint_S (\nabla \times F) \cdot \overline{n} \, \mathrm{d}S &= \iint_S (\nabla \times F) \cdot \, \mathrm{d}S \\
&= \iint_D (\nabla \times F) \cdot (\partial_1 r \times \partial_2 r) \, \mathrm{d}x \, \mathrm{d}y,
\end{aligned}$$

där $\partial_1 r$ och $\partial_2 r$ är de partiella derivatavektorerna

$$\partial_i r = (\partial_i r_1, \partial_i r_2, \partial_i r_3), \ i = 1, 2,$$

till avbildningen $r : D \to \mathbb{R}^3$, se definition 6.4.9.

Beviset för Stokes sats är en rättfram, men inte helt enkel, tillämpning av Greens sats. Vi förbigår beviset.

Eftersom Greens och Stokes satser är nära relaterade till varandra, undersöker vi specialfallet, där $r(x, y) = (x, y, 0)$ uttrycker ett vanligt planområde D och den slutna stigen γ framställer randen ∂D av mängden D. Nu får vektorn $\partial_1 r \times \partial_2 r$ i punkten $(x, y) \in D$ formen

$$\partial_1 r(x,y) \times \partial_2 r(x,y) = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = e_3.$$

För att bestämma funktionen $\nabla \times F \cdot (\partial_1 r \times \partial_2 r)$ skall vi alltså endast räkna den tredje komponenten för $\nabla \times F$, eftersom effekten av de andra komponenterna försvinner då vi tar den skalära produkten med vektorn $\partial_1 r \times \partial_2 r = e_3$. Nu är $\partial_1 F_2 - \partial_2 F_1$ den tredje komponenten för $\nabla \times F$. Således är

$$\iint_S (\nabla \times F) \cdot \overline{n} \, \mathrm{d}S = \iint_D (\partial_1 F_2 - \partial_2 F_1) \, \mathrm{d}x \, \mathrm{d}y,$$

och å andra sidan är

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s} = \int_a^b (F_1(\gamma(t)) \, \gamma_1(t) + F_2(\gamma(t)) \, \gamma_2'(t)) \, \mathrm{d}t$$

en vanlig integral längs stigen γ i $x_1 x_2$-planet. Märk, att $\gamma_3'(t) = 0$, eftersom stigen γ är i formen $\gamma(t) = (\gamma_1(t), \gamma_2(t), 0)$. Stokes sats 6.6.1 reduceras därmed i detta specialfall till Greens formel, formel (6.2.2).

6.6.3. Exempel. Låt stigen γ representera snittkurvan av den oändliga cylindern, d.v.s. tuben $x_1^2 + x_2^2 = 1$ och planet $x_1 + x_2 + x_3 = 2$, roterad en gång medsols. Antag att vektorfältet F är

$$F(x_1, x_2, x_3) = (x_2, x_1^2, x_3).$$

Vi bestämmer

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s}. \tag{6.6.4}$$

Detta kan beräknas utgående från definitionen, ty vi kan välja

$$\gamma(t) = (\cos t, \sin t, 2 - \cos t - \sin t),$$

$t \in [0, 2\pi]$, varvid

$$\int_\gamma F \cdot \, \mathrm{d}\overline{s} = \int_a^b F(\gamma(t)) \cdot \gamma'(t) \, \mathrm{d}t.$$

Trots det beräknar vi integralen (6.6.4) med hjälp av Stokes sats. Stigen γ representerar randkurvan av planytan, som cylindern $x_1^2 + x_2^2 = 1$

avskiljer från planet $x_1+x_2+x_3=2$. Denna yta kan uttryckas i formen

$$r(x_1,x_2)=(x_1,x_2,2-x_1-x_2),$$

$(x_1,x_2)\in B(0,1)$, d.v.s. ytan är grafen av funktionen

$$f(x_1,x_2)=2-x_1-x_2.$$

Observera, att orienteringarna nu är korrekta (rita bild). Funktionen f får vi genom att lösa x_3 ur ekvationen för planet $x_1+x_2+x_3=2$. Med stöd av Stokes sats är

$$\int_\gamma F\cdot\ \mathrm{d}\overline{s}=\iint_S(\nabla\times F)\cdot\overline{n}\ \mathrm{d}S.$$

Vi beräknar den senare integralen.

Vi får att rotationen $\nabla\times F$ för vektorfältet F är

$$\nabla\times F=\begin{vmatrix}e_1 & e_2 & e_3\\ \partial_1 & \partial_2 & \partial_3\\ x_2 & x_1^2 & x_3\end{vmatrix}=(2x_1-1)e_3,$$

ty den första och andra komponenten i $\nabla\times F$ är 0. Då $S=r(D)$ är grafen av funktionen f, så ger formel (6.4.11) ett enkelt uttryck för integralerna

$$\iint_S G\cdot\overline{n}\ \mathrm{d}S=\iint_S G\cdot\ \mathrm{d}\overline{S},$$

i vektorfältet G. Vi tillämpar detta på fallet $G=\nabla\times F$, varvid vi får att

$$\iint_S(\nabla\times F)\cdot\overline{n}\ \mathrm{d}S=\iint_{B(0,1)}(2x_1-1)\ \mathrm{d}x_1\,\mathrm{d}x_2.$$

Den senare integralen lönar sig att beräkna med polära koordinater, vilket ger att

$$\begin{aligned}\iint_{B(0,1)}(2x_1-1)\ \mathrm{d}x_1\,\mathrm{d}x_2&=\int_0^1\int_0^{2\pi}(2r\cos\varphi-1)r\,\mathrm{d}\varphi\,\mathrm{d}r\\&=\int_0^1\Big/_0^{2\pi}(2r^2\sin\varphi-r\varphi)\,\mathrm{d}r\\&=-\int_0^1 2\pi r\,\mathrm{d}r=-\pi\Big/_0^1 r^2=-\pi.\end{aligned}$$

Märk, att vi ovan använde en transformation till polära koordinater, där $x_1 = r \cos \varphi$, $x_2 = r \sin \varphi$ och Jacobis determinant för transformationen är r.

Sakregister