Stockholms matematiska cirkel

Variationskalkyl

Linus Bergqvist
Lukas Gustafsson

Institutionen för matematik, KTH och
Matematiska institutionen, Stockholms universitet
2022–2023

Stockholms matematiska cirkel genom tiderna
(tidigare KTH:s matematiska cirkel)

2022-2023	Variationskalkyl
2021-2022	Matematik och AI
2020-2021	Musik och matematik
2019-2020	Datorernas matematik
2018-2019	Grafteori med inriktning på färgläggning
2017-2018	Geometriska konstruktioner
2016-2017	Vad är ett tal?
2015-2016	Fraktaler
2014-2015	Polytoper
2013-2014	Grupper, mönster och symmetrier
2012-2013	Den matematiska analysens grunder
2011-2012	Diofantiska ekvationer
2010-2011	Polynom
2009-2010	Hyperbolisk geometri
2008-2009	Talteori
2007-2008	Sannolikhetsteori
2006-2007	Gruppteori
2005-2006	Vad är ett tal?
2004-2005	Integraler
2003-2004	Linjär algebra och bioinformatik
2002-2003	Algebra och kryptografi
2001-2002	Analysens grunder
2000-2001	Talföljder, rekursioner och iterationer
1999-2000	Linjära avbildningar

Innehåll

Lista över grekiska alfabetet

A	α	alfa
B	β	beta
Γ	γ	gamma
Δ	δ	delta
E	ε	epsilon
Z	ζ	zeta
H	η	eta
Θ	θ	theta
I	ι	iota
K	κ	kappa
Λ	λ	lambda
M	μ	my
N	ν	ny
Ξ	ξ	xi
O	o	omikron
Π	π	pi
P	ρ	rho
Σ	σ	sigma
T	τ	tau
Υ	υ	ypsilon
Φ	ϕ	fi
X	χ	chi
Ψ	ψ	psi
Ω	ω	omega

Några ord på vägen

Detta kompendium är skrivet för att användas som kurslitteratur till Stockholms matematiska cirkel under läsåret 2022–2023 och består av åtta kapitel.

Kompendiet är inte tänkt att läsas enbart på egen hand, utan ska ses som ett skriftligt komplement till undervisningen. Alla elever rekommenderas att läsa igenom varje kapitel själv innan föreläsningen. Det är inte nödvändigt att förstå alla detaljer vid den första genomläsningen.

Som de flesta matematiska skrifter på högre nivå är kompendiet kompakt skrivet. Detta innebär att man i allmänhet inte kan läsa det som en vanlig bok. Istället bör man pröva nya satser och definitioner genom att på egen hand exemplifiera. Därmed uppnår man oftast en mycket bättre förståelse av vad dessa satser och deras bevis går ut på.

Till varje kapitel finns ett antal övningsuppgifter. De udda övningarna har lösningar längst bak i kompendiet. Syftet med dessa är att eleverna ska kunna lösa dem och på egen hand kontrollera att de förstått materialet. Övningar med jämna nummer saknar facit och kan användas som examination. Det rekommenderas dock att man försöker lösa dessa uppgifter även om man inte examineras på dem.

Om man kör fast kan man alltid fråga en kompis, en lärare på sin skola eller någon av författarna. Under årets gång kommer det att finnas övningstillfällen där eleverna kan jobba med uppgifterna, själva eller i grupp, och få hjälp av oss.

De övningsuppgifter som är något svårare markeras med en stjärna ($\star$). Uppgifter som är extra utmanande markeras med två stjärnor ($\star\star$).

Vissa övningar kan ha flera lösningar och det som står i facit bör i detta fall endast ses som ett förslag.

Målet med årets kurs är att introducera teorin som står till grund för det matematiska verktyget *variationskalkyl*. Innan vi gör det så studerar och repeterar vi de grundläggande principerna i matematik i Kapitel 1. Sedan studerar vi Euklidiska rum i Kapitel 2 så att vi bättre förstår de underliggande mängderna i matematisk analys. Kapitel 3 behandlar grunderna för analys i en variabel och hur man löser en viss typ av differentialekvationer, vilket kommer att behövas för att lösa Brachistochronproblemet och andra problem inom variationskalkyl. Kapitel 4 och 5 behandlar derivering och optimiering i flera variabler. Detta är viktiga ämnen inom matematiken och det är dessa koncept vi måste generalisera för att nå fram till och definera variationskalkyl. I Kapitel 6 studerar vi funktionaler, en typ av exotiska funktioner som vi behöver för att kunna formulera Brachistochronproblemet matematiskt. Slutligen så formulerar och bevisar vi Euler–Lagranges ekvation som man skulle kunna kalla för *variationskalkylens fundamentalsats*. Slutligen i Kapitel 8 använder vi all den teori vi byggt upp och variationskalkyl för att lösa Brachistochronproblemet.

Några ord om cirkeln

Stockholms matematiska cirkel är en kurs för matematikintresserade gymnasieelever, som arrangeras av Kungliga Tekniska högskolan och Stockholms universitet. Cirkeln startade 1999. Vid starten hade den namnet KTH:s matematiska cirkel och hölls i KTH:s ensamma regi. Ambitionen med cirkeln är att sprida kunskap om matematiken och dess användningsområden utöver vad eleverna får genom gymnasiekurser, och att etablera ett närmare samarbete mellan gymnasieskolan och högskolan. Cirkeln ska särskilt stimulera elevernas matematikintresse och inspirera dem till fortsatta naturvetenskapliga och matematiska studier.

Till varje kurs skrivs ett kompendium som distribueras gratis till eleverna. Detta material, föreläsningsschema och övrig information om Stockholms matematiska cirkel finns tillgängligt på

`https://www.math-stockholm.se/cirkel`

Cirkeln godkänns ofta som en gymnasiekurs eller som matematisk breddning på gymnasieskolorna. Det är upp till varje skola att godkänna cirkeln som en kurs och det är lärarna från varje skola som sätter betyg på kursen. Lärarna är självklart också välkomna till cirkeln och många har kommit överens med sin egen skola om att få cirkeln godkänd som fortbildning eller som undervisning.

Vi vill gärna understryka att föreläsningarna är öppna för alla gymnasieelever, lärare eller andra matematikintresserade.

Vi har avsiktligt valt materialet för att ge eleverna en inblick i matematisk teori och tankesätt och presenterar därför både några huvudsatser inom varje område och bevisen för dessa resultat. Vi har också som målsättning att bevisa alla satser som används om de inte kan förutsättas bekanta av elever från gymnasiet.

Författarna, sommaren 2022

1 Matematik

Temat för årets matematiska cirkel är variationskalkyl. Detta kapitel är en introduktion i den matematiska metoden och ett antal grundbegrepp som vi kommer använda oss av i kursen.

1.1 Definition, axiom, sats och bevis

I detta avsnitt ska vi beskriva den matematiska metoden utifrån fyra begrepp: *definition*, *sats*, *bevis* och *axiom*.

En *definition* bestämmer vad en term betyder så att man kan arbeta matematiskt med den. Till exempel kan vi definiera udda och jämna tal på följande sätt.

Definition 1.1.1. Ett heltal n är *udda* om det finns ett heltal k som uppfyller att $n = 2k + 1$. △

Definition 1.1.2. Ett heltal n är *jämnt* om det finns ett heltal k som uppfyller att $n = 2k$. △

Ofta har man en intuition om vad en term betyder redan innan man definierar den. Läsaren hade till exempel säkert en uppfattning om vad udda och jämna tal är innan vi definierade dem. Syftet med en definition är att precisera detta.

När definitionen är gjord, så överger man sina tidigare uppfattningar om vad termen betyder och utgår endast ifrån definitionen. Man säger att definitionen är *stipulativ*. En definition är alltså inte rätt eller fel, utan bara mer eller mindre användbar och intuitiv.

Definitioner bygger ofta på begrepp som läsaren är bekant med. Till exempel utgår Definition 1.1.1 och 1.1.2 från att läsaren redan vet vad ett heltal är.

En *sats* är ett påstående som bevisats vara sant. Varje sats hör samman med ett *bevis*: ett argument för att påståendet är sant.

Sats 1.1.3. *Om n är udda, så är $n + 1$ jämnt.*

Bevis. Om n är udda så finns det ett heltal k så att $n = 2k + 1$. Då gäller att

$$n + 1 = 2k + 1 + 1 = 2k + 2 = 2(k + 1).$$

Eftersom $k + 1$ är ett heltal, så är $n + 1$ ett jämnt tal. □

Bevisen kombinerar definitioner och olika logiska slutledningsregler för att nå den önskade slutsatsen. Sats 1.1.3 har en syskonsats. Beviset är mer eller mindre identiskt, och lämnas som övning.

Sats 1.1.4. *Om n är jämnt, så är $n + 1$ udda.*

En sats vars främsta syfte är att användas i beviset av en annan sats kallas för en *hjälpsats* eller ett *lemma*. En sats som följer omedelbart ur en annan sats, till exempel som ett specialfall, kallas för en *följdsats* eller ett *korollarium*.

Ett påstående måste vara bevisat för att få kallas för en sats. Om man har goda skäl att tro att ett påstående är sant men inte formellt bevisat det kallas påståendet för en *förmodan*, eller *hypotes*. Två exempel är *Riemannhypotesen* och *primtalstvillingsförmodan*.

En förmodan kan förbli obevisad i hundratals år. Ett berömt exempel är *Fermats stora sats*, som formulerade av Pierre de Fermat (1607–1665) år 1637 men bevisades först av Andrew Wiles år 1995. Riemannhypotesen, som ännu är obevisad, formulerades 1859 av Bernard Riemann (1826–1866).

Eftersom bevisen utgår ifrån definitionen, och inte vår intuition, så behöver man ibland bevisa saker som känns uppenbara. Läsaren vet till exempel att

(i) alla tal antingen är udda eller jämna, och

(ii) ett tal kan inte vara udda och jämnt samtidigt.

Men om man läser Definition 1.1.1 och 1.1.2 så är detta inte självklart. Kan man inte tänka sig tal som varken är udda eller jämnt? Eller tal som är både och?

Bevis bygger på antaganden. Dessa antaganden måste dock bevisas innan de kan anses giltiga. Men dessa bevis måste också bygga på antaganden, som också måste bevisas, och så vidare.

För att undvika en oändlig kedja av bevis, eller ett cirkulärt bevis (ett bevis som använder sig av det man försöker bevisa) så måste man göra grundantaganden som inte behöver bevisa. Dessa kallas för *axiom*.

1.2 Mängder

En *mängd* är en samling objekt. Man kan samla nästan vad man vill i en mängd: tal, katter, och andra mängder.[1] Det viktiga är att man alltid kan avgöra ifall ett objekt tillhör mängden eller inte. De objekt som ligger i mängden kallas för *element*.

Det lättaste sättet att beskriva en mängd är att räkna upp element som ingår i den. För att markera att objekten ligger i en mängd, så omger man listan med *mängdklamrar* $\{$ och $\}$. Mängden som innehåller 1, 2 och 3 skrivs alltså som

$$\{1, 2, 3\}.$$

Två mängder är lika om de innehåller samma element, vilket skrivs $A = B$. Det spelar ingen roll i vilken ordning man skriver elementen eller hur många gånger de listas. Därför gäller att

$$\{1, 1, 2, 3\} = \{1, 2, 3\} = \{2, 3, 1\}.$$

Om ett element x tillhör en mängd A kan man skriva $x \in A$, vilket uttalas som x tillhör A. Om x inte tillhör A skriver man $x \notin A$. Mängden som inte

[1] Vi skriver *nästan* av en anledning. Det finns samlingar av objekt som kan beskrivas men som inte utgör en mängd. Detta kallas *Russells paradox*, efter Bertrand Russell (1872–1970). Russells exempel är samlingen av alla mängder som inte innehåller sig själva.

innehåller några element alls kallas den *tomma mängden*, och betecknas med $\emptyset$.

En mängd kan innehålla andra mängder som element. Mängden

$$A = \{\{1, 2\}, 3\}$$

har två element: mängden $\{1, 2\}$ och talet 3. Mängden $\{1, 2\}$ innehåller i sin tur elementen 1 och 2. Däremot innehåller A varken 1 eller 2.

Att mängder kan innehålla andra mängder kan ha paradoxala konsekvenser. Till exempel kan vi lägga den tomma mängden i en mängd, och bilda mängden av den tomma mängden.

$$A = \{\emptyset\} = \{\{\}\}$$

Mängden A innehåller ett element, den tomma mängden, och är därför inte tom. Mängden av den tomma mängden är alltså inte lika med den tomma mängden.

Detta verkar motsägelsefullt. Den tomma mängden är ju tom, så mängden av den tomma mängden borde ju också vara tom? Tricket är att skilja på mängden och elementen i mängden. Den tomma mängden är ju ett element i sig, även om den inte innehåller några element, precis som att 0 är ett tal, trots att representerar ett antal som inte finns.

Det finns ingen begränsning på hur stor en mängd kan vara, och de flesta mängder man studerar innehåller oändligt många element. Dessa mängder kan naturligtvis inte skrivas ut som en lista. Istället beskriver man dem med *mängdbyggaren*, som har följande allmänna form.

$$\{x \mid \text{villkor på } x\}$$

Den här mängden består av alla element som uppfyller villkoret. Ett exempel är mängden

$$\{n \mid n \text{ är jämnt}\} = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$$

som innehåller alla jämna tal.

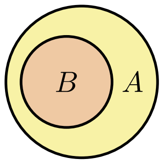
En mängd B är en *delmängd* av en mängd A om alla element i B är ett element i A. Man skriver detta som $B \subset A$. Till exempel så är $\{1, 2\}$ en delmängd av $\{1, 2, 3\}$, eftersom 1 och 2 är element i båda mängderna. Om två mängder är delmängder av varandra så är de lika.

En mängd har alltid minst två delmängder: sig själv och den tomma mängden. En delmängd B av A är *äkta* om $B \neq A$.

Det är lätt att blanda ihop element och delmängder. Det beror på att mängder kan innehålla andra mängder, så att en delmängd av en mängd kan vara ett element i mängden. Mängden $A = \{\emptyset\}$ är ett bra exempel. Den tomma mängden är både ett element i och en delmängd av A.

I mängden $A = \{1, 2, \{1, 2\}\}$ är $\{1, 2\}$ både en delmängd och ett element. Däremot så är $\{1\}$ enbart en delmängd av A, medan 1 enbart är ett element.

För att illustrera mängder använder man *Venndiagram*, efter matematikern John Venn (1834–1923). Där representeras mängder som enkla former, oftast cirklar, och formernas förhållanden till varandra motsvarar mängdernas. Till exempel kan man illustrera att B är en delmängd av A genom att rita dem som två cirklar, där B ligger inuti A.



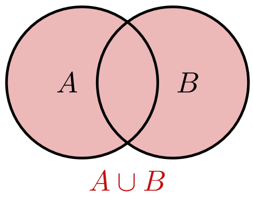
Figur 1.1: Venndiagram för $B \subset A$.

Ur gamla mängder kan vi skapa nya genom de så kallade *mängdoperationerna*.

(i) *Unionen* av mängderna A och B är mängden som består av alla element som ligger i A *eller* i B. Den betecknas med $A \cup B$ och definieras som

$$A \cup B = \{x \mid x \in A \text{ eller } x \in B\}.$$

Ett exempel är $\{1, 2, 3\} \cup \{2, 3, 4\} = \{1, 2, 3, 4\}$.

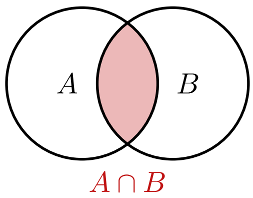


Figur 1.2: Venndiagram för $A \cup B$.

(ii) *Snittet* av mängderna A och B är mängden som består av alla element som ligger i A *och* i B. Den betecknas med $A \cap B$ och definieras som

$$A \cap B = \{x \mid x \in A \text{ och } x \in B\}.$$

Ett exempel är $\{1, 2, 3\} \cap \{2, 3, 4\} = \{2, 3\}$.

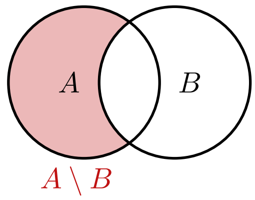


Figur 1.3: Venndiagram för $A \cap B$.

(iii) *Differensen* av en mängd A och B är mängden som består av alla element som ligger i A men inte i B. Den betecknas med $A \setminus B$ och definieras som

$$A \setminus B = \{x \mid x \in A \text{ och } x \notin B\}.$$

Ett exempel är $\{1, 2, 3\} \setminus \{2, 3, 4\} = \{1\}$.



Figur 1.4: Venndiagram för $A \setminus B$.

> Notera att $A \setminus B$ inte är lika med $B \setminus A$, exempelvis gäller $\{2,3,4\} \setminus \{1,2,3\} = \{4\}$.

Två mängder A och B är *disjunkta* om de inte har några gemensamma element, det vill säga om $A \cap B = \emptyset$.

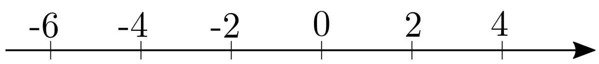
Om alla uppträdande mängder är delmängder av en viss mer eller mindre underförstådd grundmängd M talar man ofta om $M \setminus A$ som *komplementet* till A (med avseende på M). Vi kommer att beteckna komplementet till A med A^c.

Om vi till exempel pratar om mängder av heltal, och vi till exempel betraktar mängden $A = \{1, 2, 3\}$, så avser komplementet till A mängden av alla heltal *förutom* $1, 2$ och 3.

De olika talsystemen kan ses som mängder av tal, och har fått egna beteckningar. De *naturliga talen* betecknas med $\mathbb{N}$ och består av talen 0, 1, 2, 3, och så vidare.[2]

Naturliga tal kan adderas och multipliceras utan problem. Resultatet är alltid ett nytt naturligt tal. För att subtrahera behöver vi införa de negativa talen -1, -2, och så vidare. De naturliga talen tillsammans med de negativa talen kallas för *heltalen*, och betecknas med $\mathbb{Z}$ (av tyskans *Zahl* = tal).

Heltal kan adderas, subtraheras och multipliceras. Man kan däremot inte dividera dem med varandra. För detta krävs *rationella tal*. De definieras som alla kvoter a/b, där a och b är heltal och b är skilt från 0. Mängden av alla rationella tal betecknas med $\mathbb{Q}$.

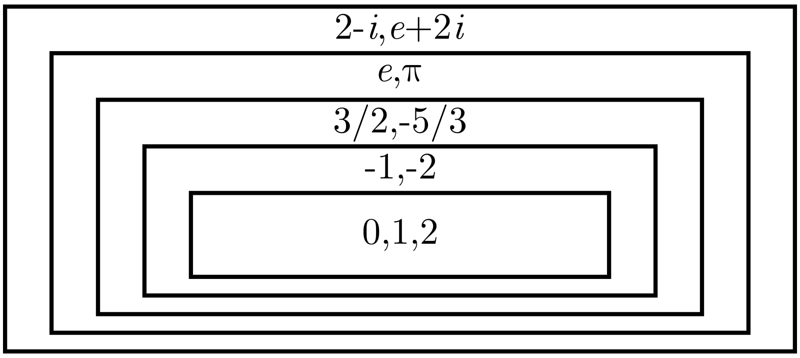


Figur 1.5: Tallinjen runt 0.

De rationella talen ligger på den så kallade *tallinjen*, som går från negativa tal till vänster och till positiva tal till höger (se Figur 1.5). Det finns dock tal som inte är rationella, men som ändå ligger på tallinjen. Ett exempel är $\sqrt{2}$, som är längden på diagonalen i en kvadrat med sidan 1. Lägger man till dessa tal får

[2] Vissa exkluderar 0 från de naturliga talen. Att inkludera 0 har dock fördelar. Om man börjar räkna från 0 och går ett steg i taget kommer man ha gått n steg när man räknat till n. Exempel: om vi räknar till 3 från 0 så får vi $0 \to 1 \to 2 \to 3$, vilket är 3 steg. Om vi börjar från 1 får vi istället $1 \to 2 \to 3$, vilket är 2 steg.

de *reella talen*, som betecknas med $\mathbb{R}$.[3] Reella tal som inte är rationella kallas för *irrationella*.



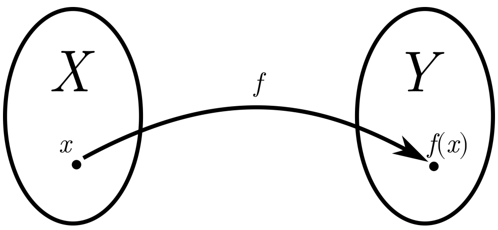
Figur 1.6: De olika talsystemen från $\mathbb{N}$ till $\mathbb{C}$.

De reella talen kan utvidgas ytterligare till de *komplexa talen*, som betecknas med $\mathbb{C}$, genom att lägga till ett tal i som uppfyller $i^2 = -1$.

1.3 Funktioner

En funktion $f : X \to Y$ parar ihop element i en mängd X med element i en mängd Y. Mängden X kallas för *definitionsmängd* och mängden Y kallas för *målmängd*. Man kan se f som en process som tar ett element i mängden X och avger ett element som ligger i mängden Y. När man tillämpar en funktion på ett element x i X så kallas x för funktionens *argument*. Mängden av alla värden en funktion i praktiken antar kallas för funktionens *värdemängd*, och denna betecknas ofta med V_f. Värdemängden är då alltså en delmängd av målmängden, och kan beskrivas som $V_f = \{f(x) : x \in X\}$. Till exempel är $\sin(x)$ en funktion från $\mathbb{R}$ till $\mathbb{R}$, så funktionens målmängd är alltså $\mathbb{R}$, men värdemängden är $[-1, 1]$.

Två funktioner är lika när de har samma definitionsmängd, samma målmängd och de är lika på alla element i definitionsmängden. Definitions- och målmängden är alltså en del av funktionen.



Figur 1.7: En funktion f från X till Y.

Vi säger att en funktion $f : X \to Y$ är *surjektiv* om alla element i Y är bilden

[3]Reella tal är mycket mystiska. Den matematiska cirkeln 2016–2017, *Vad är ett tal?*, handlade om hur man kan definiera dem i termer av rationella tal. Den intresserade läsaren uppmanas att söka upp kompendiet på Cirkelns hemsida: `https://www.math-stockholm.se/samverkan/cirkel/`

av något element i X, det vill säga om värdemängden överensstämmer med målmängden. En funktion sägs vara *injektiv* om varje element i värdemängden är bilden av exakt ett element i definitionsmängden. Detta är ekvivalent med att: $f(x) = f(y) \iff x = y$. Om en funktion är både surjektiv och injektiv säger vi att funktionen är *bijektiv*.

Notera att en bijektiv funktion är *inverterbar*: Det vill säga, det existerar en funktion $f^{-1} : Y \to X$ med egenskapen att $f^{-1}(f(x)) = x$ för alla $x \in X$ och $f(f^{-1}(y)) = y$ för alla $y \in Y$. Den inversa funktionen, om den finns, är alltså den funktion som för varje element $y \in Y$ ger det unika elementet $x \in X$ som har egenskapen att $f(x) = y$. I övning 3.7 ser vi exempelvis att en strikt växande funktion $f : \mathbb{R} \to \mathbb{R}$ är injektiv. Det är även värt att notera att en injektiv funktion alltid är inverterbar om vi begränsar målmängden till värdemängden.

Funktioner beskrivs ofta med formler. Exempelvis så kan funktionen $f : \mathbb{N} \to \mathbb{N}$ som tar ett naturligt tal och returnerar dess kvadrat beskrivas som $f(n) = n^2$. Alla polynom kan ses som en funktion från $\mathbb{R}$ till $\mathbb{R}$, som beräknas genom att man sätter in talet x i uttrycket.

Huruvida en funktion är inverterbar eller ej beror inte bara på regeln som beskriver funktionen, utan på definitionsmängd och målmängden.

Exempel 1.3.1. Betrakta funktionen $f : x \mapsto x^2$ som tar ett tal och kvadrerar det.

Betraktat som funktion från $\mathbb{R}$ till $\mathbb{R}$ är den inte inverterbar. Till exempel kan vi inte definiera $f^{-1}(-1)$ eftersom det inte finns något $x \in \mathbb{R}$ så att $x^2 = -1$. Problemet här är alltså att funktionens målmängd inte överensstämmer med dess värdemängd, det vill säga att funktionen inte är surjektiv.

Detta problem kan lösas genom att begränsa målmängden till värdemängden, men även om vi betraktar f som en funktion från $\mathbb{R}$ till dess värdemängd, det vill säga $[0, \infty)$, så är funktionen inte inverterbar. Vi har fortfarande problemet att funktionen inte är injektiv. Vi kan till exempel inte entydigt definiera $f^{-1}(1)$ eftersom det finns två element, 1 och -1, som har egenskapen att $f(x) = 1$.

Om vi däremot betraktar f som en funktion från $[0, \infty)$ till $[0, \infty)$ så är den både injektiv och surjektiv, det vill säga bijektiv, och således har funktionen en väldefinierad invers, nämligen $\sqrt{x}$.

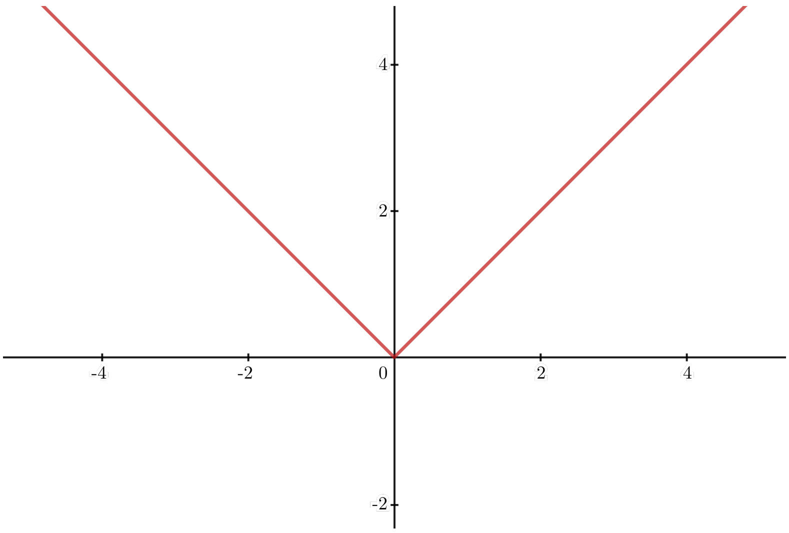
Vi har en liknande situation med $\sin(x)$ och $\cos(x)$. Betraktade som funktioner från $\mathbb{R}$ till $\mathbb{R}$ är dessa funktioner varken injektiva eller surjektiva, men genom att betrakta $\sin(x)$ som en funktion från $[-\pi/2, \pi/2]$ till $[-1, 1]$ och $\cos(x)$ som en funktion från $[0, \pi]$ till $[-1, 1]$ erhåller vi bijektiva, och således inverterbara funktioner. ▲

En funktion måste dock inte ges av en formel. Det enda som krävs är att funktionen är definierad för alla element i X, och att den alltid ger samma svar. Ett exempel är absolutbeloppet $|x|$ av ett reellt tal x, som definieras som avståndet på tallinjen från x till origo. Man kan beräkna det genom att man tar bort eventuella minustecken framför talet, det vill säga

$$|x| = \begin{cases} x & \text{om } x \geqslant 0 \\ -x & \text{om } x < 0. \end{cases}$$

Exempelvis så gäller $|-3| = -(-3) = 3$ och $|2| = 2$.

En funktion $f : \mathbb{R} \to \mathbb{R}$ kan beskrivas genom sin *graf*, som definieras som mängden av punkter i planet på formen $(x, f(x))$.



Figur 1.8: Grafen av funktionen $f(x) = |x|$.

Om x är ett reellt tal så är $\lfloor x \rfloor$ det största heltalet som är mindre än eller lika med x. Till exempel så är $\lfloor 5/2 \rfloor = 2$ och $\lfloor -\pi \rfloor = -4$. Man kan se det som att $\lfloor x \rfloor$ är avrundningen av x, förutsatt att man alltid avrundar neråt.

Definition 1.3.2. *Golvfunktionen* är funktionen som avbildar reella tal x på $\lfloor x \rfloor$. △

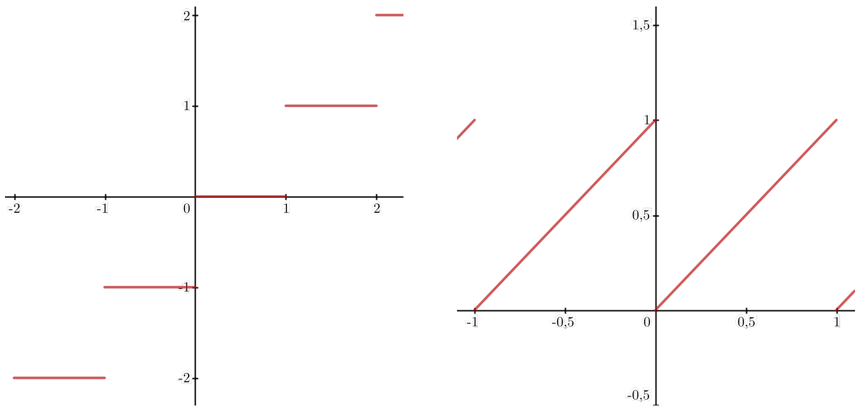
Ett annat sätt att se golvfunktionen är att den avbildar x på det största heltalet n som uppfyller olikheterna

$$n \leqslant x < n + 1.$$

Definition 1.3.3. *Fraktionsfunktionen* $\mathrm{frac} : \mathbb{R} \to \mathbb{R}$ är funktionen som avbildar reella tal x på $x - \lfloor x \rfloor$. △

Värdet $\mathrm{frac}(x)$ ligger alltid mellan 0 och 1, och $\lfloor x \rfloor + \mathrm{frac}(x) = x$. Det senare följer direkt av definitionen, eftersom

$$\lfloor x \rfloor + \mathrm{frac}(x) = \lfloor x \rfloor + x - \lfloor x \rfloor = x.$$



(a) Graf av $\lfloor x \rfloor$ **(b)** Graf av $\mathrm{frac}(x)$

Figur 1.9: Golv- och fraktionsfunktionen

Talet $\mathrm{frac}(x)$ kallas för *fraktionsdelen* av x. Obervera att $\mathrm{frac}(x) = 0$ när x är ett heltal. Två andra användbara samband är att $\mathrm{frac}(x + n) = \mathrm{frac}(x)$ och $\lfloor x + n \rfloor = \lfloor x \rfloor + n$ för alla heltal n.

1.4 Bevistekniker

Ett bevis för en sats är ett argument som förklarar varför satsen är sann. Vi har redan sett ett exempel när vi bevisade Sats 1.1.3. I detta avsnitt ska vi gå igenom tre tekniker för att bevisa matematiska satser: direkta bevis, motsägelsebevis och induktionsbevis.[4]

Ett *direkt bevis* utgår ifrån satsens antaganden och definitioner och bevisar satsen rakt på, så att säga. Beviset av Sats 1.1.3 är ett exempel på direkt bevis. Ett annat är följande sats.

Sats 1.4.1. *Antalet funktioner från en mängd A med n element till en mängd B med m element är m^n.*

Bevis. Varje funktion från A till B kan beskrivas som en tabell där varje element i A motsvaras av precis ett element i B. Listan innehåller totalt n platser, och på varje plats kan vi välja bland m element att välja bland. Alltså finns det totalt

$$\underbrace{m \cdot m \cdots m \cdot m}_{n \text{ stycken}} = m^n$$

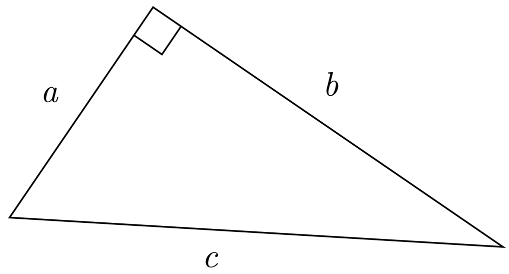
olika funktioner. □

Ibland går det inte att använda direkta bevis, till exempel när man ska bevisa att något inte är fallet. Då kan det vara enklare att anta att det man vill bevisa är falskt, och visa att detta leder till en motsägelse. Om alla steg i beviset är korrekta så måste det ursprungliga antagandet vara fel. Detta kallas för ett *motsägelsebevis*.

[4]Ibland förekommer termen *indirekt bevis*. Vissa använder det som synonym till motsägelsebevis, andra som en synonym till bevisregeln *modus tollens*. Vi undviker den helt.

Sats 1.4.2. *Längden av hypotenusan i en rätvinklig triangel är alltid mindre än summan av längderna av kateterna.*

Bevis. Låt a och b vara längderna av kateterna i triangeln och c längden av hypotenusan.



Figur 1.10: En rätvinklig triangel.

Enligt Pytagoras sats är $a^2 + b^2 = c^2$. Antag motsatsen till satsen, det vill säga att $a + b \leqslant c$. Då gäller att

$$(a+b)^2 \leqslant c^2 \implies a^2 + 2ab + b^2 \leqslant c^2 \implies 2ab \leqslant 0.$$

Eftersom a och b är längderna i en rätvinklig triangel så är a och b båda större än 0. Eftersom produkter av positiva tal är positiva, så är detta en motsägelse. Alltså måste motsatsen till $a + b \leqslant c$ gälla, det vill säga $a + b > c$. □

Sats 1.4.3. *Ett tal kan inte vara udda och jämnt samtidigt.*

Bevis. Antag att n är ett tal som är båda udda och jämnt. Då finns det två tal, k och l, så att $n = 2k$ och $n = 2l + 1$. Då gäller att

$$2k = n = 2l + 1 \implies 2k - 2l = 1 \implies 2(k - l) = 1.$$

Med andra ord finns det heltal $m = k - l$ så att $2m = 1$. Kan det finns ett sådant tal? Det finns två fall.

(i) Om $m \leqslant 0$, så är $1 = 2m \leqslant 0$. Motsägelse!

(ii) Om $m \geqslant 1$ så är $2m \geqslant 2 > 1$. Motsägelse! □

Ett berömt motsägelsebevis är följande.

Sats 1.4.4. *Talet $\sqrt{2}$ är irrationellt.*

Bevis. Antag motsatsen, det vill säga att $\sqrt{2} = a/b$ för några heltal a och b. Antag att a och b är förkortade så långt som möjligt. Då kan endast en av a eller b vara jämn, eftersom om båda är jämna kan vi skriva

$$\sqrt{2} = \frac{a}{b} = \frac{2c}{2d} = \frac{c}{d}$$

och då var inte a och b förkortade så långt som möjligt.

Av definitionen av $\sqrt{2}$ får vi att

$$\sqrt{2}^2 = 2 = \frac{a^2}{b^2} \implies 2b^2 = a^2.$$

Den sista ekvationen säger att a^2 är jämn. Eftersom kvadrater av udda tal är udda (se Övning 1.14), så måste a vara ett jämnt tal, det vill säga $a = 2k$ för något heltal k. Då får vi att

$$2b^2 = (2k)^2 = 4k^2 \implies b^2 = 2k^2.$$

Eftersom b^2 är jämnt, så måste b vara jämnt. Men nu har vi bevisat att både a och b är jämna, vilket var omöjligt eftersom vi hade förkortat bråket så långt som möjligt. Detta är en motsägelse. □

Bevisen av Sats 1.4.2 och 1.4.3 bygger båda på ett antagande som vi rent formellt inte bevisat (kan du se vilket?). Att bevis inkluderar sådana antaganden är snarare regel än undantag. Ifall man bevisade precis vartenda antagande utifrån axiomen skulle bevisen bli väldigt långa och komplicerade. Läsaren förväntas själv fylla i de luckor som uppstår.

Det händer dock att uppenbara antaganden är mycket svåra, till och med omöjliga, att bevisa utifrån definitionerna. Historien är fylld av matematiker som gjort till synes självklara antaganden som sedan visat sig vara svåra att bevisa.

Beviset av Sats 1.4.4 är ett exempel på det. Vi antar att ett bråk kan förkortas så långt som möjligt. Detta är inte självklart, utan bygger i själva verket på aritmetikens fundamentalsats, en sats som vi återkommer till senare i kursen.

Den tredje bevistekniken som vi kommer gå igenom kallas för *induktionsbevis*. Säg att du har en följd av påståenden P_0, P_1, P_2 och så vidare, ett för varje naturligt tal. För att bevisa att alla dessa påståenden gäller, så kan man göra det i två steg:

(i) **Basfall**: Bevisa att P_0 är sann.

(ii) **Induktionssteget:** Bevisa att om P_n är sant för något naturligt tal n så är P_{n+1} sant.

Idén är att om båda dessa gäller, så är P_m sant för alla naturliga tal m. Man kan motivera det på följande sätt. Säg att vi undrar ifall P_m är sant. Enligt basfallet är P_0 sant, och genom att tillämpa induktionssteget med $n = 0$ så kan vi dra slutsatsen att P_1 är sant. Vi kan nu sätta $n = 1$, och induktionssteget säger då att P_2 är sant. Genom att upprepa detta m gånger, kan vi bevisa att P_m är sant. Vi kan illustrera det hela som en följd av implikationer:

$$P_0 \implies P_1 \implies P_2 \implies \cdots \implies P_{m-1} \implies P_m \implies \cdots.$$

Ett exempel förtydligar hur det fungerar.

Sats 1.4.5. *Alla naturliga tal är udda eller jämna.*

Bevis. Låt P_n vara påståendet att n är antingen udda eller jämnt. Då är påståendet att alla naturliga tal är udda eller jämna samma som att P_n är sant för alla n.

(i) **Basfall**: När $n = 0$ säger satsen att 0 är udda eller jämnt, vilket är sant eftersom $0 = 2 \cdot 0$ är ett jämnt tal.

(ii) **Induktionssteg**: Antag att P_m är sant för något m, det vill säga m är antingen udda eller jämnt. Vi ska bevisa att $m + 1$ är udda eller jämnt. Det gör vi i två fall.

- Om m är udda så är $m + 1$ jämnt enligt Sats 1.1.3.
- Om m är jämnt så är $m + 1$ udda enligt Sats 1.1.4.

Detta bevisar att $m + 1$ antingen är udda eller jämnt, det vill säga att P_{m+1} är sant. □

Induktionsbevis reducerar satser som handlar om alla tal till satser som handlar om specifika tal. Dessa kan vara enklare att bevisa, eftersom du istället får ett konkret tal att arbeta med.

Induktionsbevis kan vara klurigt i början. Ett organiserat sätt att göra dem är följande sätt.

(i) Omformulera satsen så att den blir av typen: P_n är sann för alla n.

(ii) Bevisa P_0.

(iii) Anta att det finns ett tal n så att P_n är sant, och bevisa att P_{n+1} är sant.

Ett induktionsbevis bygger på att fallet $n+1$ kan beskrivas i termer av fallet n. Om man inte kan hitta en sådan reduktion, kommer ett induktionsbevis vara svårt att använda.

1.5 Olika satser och hur man bevisar dem

I föregående avsnitt diskuterade vi olika bevistekniker. Men vilka tekniker är lämpliga för vilka typer av satser?

- **Implikation:** Man säger att P implicerar Q om Q är sant när P är det. Ett exempel är Sats 1.1.4, som säger att om ett tal n är jämnt, så är talet $n + 1$ udda. Man brukar beteckna implikationer med en tjock pil $\implies$, så att P medför Q skrivs

  $$P \implies Q.$$

  En implikation kan bevisas med ett direkt bevis. Då antar man att P är sant, och sedan visar man att Q också måste vara sant (det är så vi bevisar Sats 1.1.4). Man kan också använda ett motsägelsebevis. Då antar man att P är sann och att Q är falsk, och bevisar en motsägelse.

Ett tredje sätt att bevisa att P implicerar Q är att bevisa att om Q är falsk, så är P falsk. Detta kallas för *omvändningen* av en implikation.

- **Ekvivalens:** En ekvivalens är när två påståenden P och Q implicerar varandra, alltså att om P så Q, och om Q så P. Man brukar använda frasen P om och endast om Q. Man använder tjocka dubbelpilar för att beteckna ekvivalenser, så att P om och endast om Q skrivs som

  $$P \iff Q.$$

  Ekvivalenser bevisas genom att första visa att P implicerar Q, och sedan att Q implicerar P.

- **Universalsats:** En universalsats säger att alla n i en mängd M uppfyller något villkor P. Universalsatser kan bevisas som implikationer, genom att omformulera universalsatsen som att om n ligger i mängden M, så uppfyller n villkoret P, det vill säga

  $$n \in M \implies n \text{ uppfyller } P$$

  Ifall M är mängden av alla naturliga tal kan man också använda ett induktionsbevis, som vi beskrev ovan.

  Man kan även bevisa en universalsats genom ett motsägelsebevis. Då antar man att det finns ett n i M som inte uppfyller P, och bevisar att det är omöjligt.

- **Existenssats:** En existenssats säger att finns ett objekt n som har egenskapen P. Den typiska existenssatsen är ekvationslösning. Att $x^2 = 3$ har en lösning är en existenssats, och kan omformuleras som att det finns ett tal x så att $x^2 = 3$.

  Ett sätt att bevisa en existenssats är att konstruera det sökta objektet utifrån objekt man redan vet finns. Till exempel så kan man bevisa att det finns ett udda kvadrattal genom att notera att $3^2 = 9$ är udda och ett kvadrattal.

  Man kan också använda ett motsägelsebevis. Då antar man att det inte existerar någon objekt med egenskapen P och visar att det leder till en motsägelse. Dessa bevis har fördelen att vi inte behöver beskriva hur objektet konstrueras. I gengäld kan bevisen vara mycket komplicerade.

En variant på universalsatsen är att inget n i M uppfyller P. Den kan omformuleras som att alla n i M saknar egenskapen P. För dessa typer av satser är motsägelsebevis ofta smidiga: man antar att det finns ett n i M som uppfyller P och härleder en motsägelse.

Universal- och existenssatser är duala till varandra, i bemärkelsen att om du ska bevisa en existenssats med hjälp av ett motsägelsebevis så antar du en universalsats, och vice versa, se bevisen av Sats 1.4.2, 1.4.3 och 1.4.4

Övningar

Övning 1.1. Lista elementen i följande mängder.

(i) $A = \{n \in \mathbb{N} \mid k < 5\}$.

(ii) $B = \{1, 2, \{2, 3\}\}$.

(iii) $C = \{k \in \mathbb{Z} \mid k^2 < 16\}$.

(iv) $A \cap B$.

(v) $(C \setminus A) \cup B$.

Övning 1.2. Lista elementen i följande mängder.

(i) $A = \{x \in \mathbb{Q} \mid x^2 = 2\}$.

(ii) $B = \{0, 1, 2, 3\}$.

(iii) $C = \{p/q \mid p, q \in \mathbb{N}, 0 \leqslant p < 3 \text{ och } 1 \leqslant q < 3\}$.

(iv) $(B \cup A) \cap C$.

(v) $(C \cap B) \setminus A$.

Övning 1.3. För nedanstående par av mängder A och B, avgör om A och B är lika, disjunkta, någon av dem är en äkta delmängd av den andra eller ingetdera.

(i) $A = \{1, 2, 3\}$ och $B = \{1, 1, 2\}$.

(ii) $A = \{0, 1, 2\}$ och $B = \{n \in \mathbb{N} \mid n^2 < 9\}$.

(iii) $A = \{\{\}\}$ och $B = \{x \in \mathbb{N} \mid 2x = -2\}$.

(iv) $A = \{x \in \mathbb{R} \mid |x| < 1\}$ och $B = \{x \in \mathbb{R} \mid |x - 1| < 1\}$.

(v) $A = \{x \in \mathbb{Q} \mid x^2 = 2\}$ och $B = \{x \in \mathbb{R} \mid x^2 = 2\}$.

Övning 1.4. För nedanstående par av mängder A och B, avgör om A och B är lika, disjunkta eller någon av dem är en äkta delmängd av den andra.

(i) $A = \{-2, 0, 2\}$ och $B = \{x \in \mathbb{Z} \mid |x| < 3 \text{ och } x \text{ är jämnt}\}$.

(ii) $A = \{x \in \mathbb{R} \mid x^2 < 2\}$ och $B = \{x \in \mathbb{Q} \mid x^2 \geqslant 2\}$.

(iii) $A = \{x \in \mathbb{Z} \mid x \text{ är jämnt}\}$ och $B = \{x \in \mathbb{Z} \mid x \text{ är kvadrattal}\}$.

(iv) $A = \{x \in \mathbb{Z} \mid 2x = -2\}$ och $B = \{x \in \mathbb{N} \mid 2x = 2\}$.

(v) $A = \{\emptyset, \{\emptyset\}\}$ och $B = \{\emptyset\}$.

Övning 1.5. Använd mängdbyggaren för att definiera följande mängder.

(i) Mängden av jämna, positiva heltal.

(ii) Mängden av rationella tal r så att $2r$ är ett heltal.

(iii) Mängden av irrationella tal som ligger inom avstånd 1 från origo.

Övning 1.6. Använd mängdbyggaren för att definiera följande mängder.

(i) Mängden av alla kvadrattal som är större än 2.

(ii) Mängden av rationella lösningar till $x^4 + x^2 - 1 = 0$.

(iii) Mängden av rationella tal som är volymen av en kub med rationella sidor.

Övning 1.7. Ange möjlig definitionsmängd och målmängd för följande funktioner.

(i) Funktionen som ger det n:te kvadrattalet.

(ii) Funktionen som beräknar arean av triangel.

(iii) Funktionen beräknar derivatan av ett andragradspolynom.

Övning 1.8. Ange möjlig definitionsmängd och målmängd för följande funktioner.

(i) Funktionen som ger arean av cirkel med radie r.

(ii) Funktionen som ger avståndet mellan 1 och ett tal r på tallinjen.

(iii) Funktionen som ger de rationella nollställena till ett förstagradspolynom med rationella koefficienter.

Övning 1.9. Är följande funktioner eller inte? Om inte, motivera varför.

(i) $f : \mathbb{R} \to \mathbb{R}$ där

$$f(x) = \begin{cases} 1 \text{ om } x \in \mathbb{Q} \\ 0 \text{ om } x \notin \mathbb{Q}. \end{cases}$$

(ii) $f : \mathbb{N} \to \mathbb{Q}$ där $f(n) = \sqrt{n}$.

(iii) $f : \mathbb{R} \to \mathbb{R}$ så att $f(x) = 0$ med sannolikhet $1/2$ och $f(x) = 1$ med sannolikhet $1/2$.

(iv) $f : \{0\} \to \mathbb{R}$ där $f(0) = 1$ om ordet Balkong börjar på B.

Övning 1.10. Är följande funktioner eller inte? Om inte, motivera varför.

(i) $f : \mathbb{Z} \to \mathbb{N}$, där $f(n)$ är siffersumman i det vanliga (decimala) talsystemet. Obs, alla *siffror* är icke-negativa.

(ii) $f : \mathbb{R} \to \mathbb{Q}$, där $f(x) = x/2$.

(iii) $f : \mathbb{N} \to \mathbb{R}$, där $f(n) = \sqrt[n+1]{n+1}$.

(iv) $f : \mathbb{Q} \to \mathbb{Z}$, där $f(p/q) = p$.

Övning 1.11. Avgör om följande funktioner är lika eller inte? Motivera varför.

(i) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$ och $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \sqrt{x^2}$.

(ii) $f : \mathbb{Z} \to \mathbb{Q}$, $f(n) = 1/n$ och $g : \mathbb{N} \to \mathbb{Q}$, $g(m) = 1/m$.

(iii) $f : \mathbb{N} \to \mathbb{Q}$, $f(n) = n/(n+1)$ och $g : \mathbb{N} \to \mathbb{R}$, $g(z) = z/(z+1)$.

Övning 1.12. Avgör om följande funktioner är lika eller inte? Om inte, motivera varför.

(i) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 + 2x + 1$ och $g : \mathbb{R} \to \mathbb{R}$, $g(x) = (x+1)^2$.

(ii) $f : \mathbb{Z} \to \mathbb{Z}$, $f(x) = x^2$ och $g : \mathbb{Z} \to \mathbb{Z}$, $g(x) = |x|^2$.

(iii) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \mathrm{frac}(x)$ och $g : \mathbb{Q} \to \mathbb{R}$, $g(x) = x - \lfloor x \rfloor$.

Övning 1.13 ($\star$). Använd ett direkt bevis för att bevisa att om n är ett jämnt tal så är $n + 1$ ett udda tal (detta är Sats 1.1.4).

Övning 1.14 ($\star$). Använd ett direkt bevis för att bevisa att om n är udda så är n^2 udda.

Övning 1.15 ($\star$). Använd ett motsägelsebevis för att bevisa att summan av ett irrationellt tal och ett rationellt tal är irrationellt.

Övning 1.16 ($\star$). Använd ett motsägelsebevis för att bevisa att om $a + b \geqslant c$ så är antingen $a \geqslant c/2$ eller $b \geqslant c/2$

Övning 1.17 ($\star$). Använd induktion för att bevisa att summan av de n första udda naturliga talen är n^2.

Övning 1.18 ($\star$). Använd induktion för att bevisa att summan av de n första naturliga talen är $n(n+1)/2$.

Satsen har också ett direkt bevis. Försök att hitta det.

Övning 1.19 ($\star\star$). Bevisa att om $ab = c$ så är $a \geqslant \sqrt{c}$ och $b \leqslant \sqrt{c}$, eller tvärtom.

Övning 1.20 ($\star\star$). Bevisa att vinkelsumman av en n-hörning är $180(n-2)$ grader

(i) med ett direkt bevis.

(ii) med induktion.

Övning 1.21 ($\star\star$). Ett sätt att definiera ordningen $\geqslant$ på naturliga tal är följande.

Ett naturligt tal n är *större eller lika med* ett naturligt tal m om det finns ett naturligt tal k så att $n = m + k$. Vi skriver detta som $n \geqslant m$.

Bevisa följande satser:

(i) $n \geqslant 0$ och $n \geqslant n$ för alla n.

(ii) Om $n \geqslant m$ och $m \geqslant k$ så gäller $n \geqslant k$.

(iii) Om $n \geqslant m$ och $m \geqslant n$, så är $n = m$.

(iv) Om n och m uppfyller $n \geqslant m$ så gäller $n + k \geqslant m + k$ för alla $k \geqslant 0$

(v) Om n och m uppfyller $n \geqslant m$ så gäller $nk \geqslant mk$ för alla $k \geqslant 0$.

Övning 1.22 ($\star$)**.** Bevisa att det finns två irrationella tal a och b så att a^b är rationellt. Tips: använd att $\sqrt{2}$ är irrationellt.

Övning 1.23 ($\star\star$)**.** Bevisa att

$$1^2 + 2^2 + \cdots + n^2 < n^3$$

för alla $n \geqslant 2$.

Övning 1.24 ($\star\star$)**.** Bevisa att det finns alla rationella tal kan skrivas som en produkt av två irrationella tal.

Övning 1.25 ($\star\star$)**.** Ett *minsta element* i en delmängd S av $\mathbb{N}$ ett tal $n \in S$ med egenskapen att $n \leqslant m$ för alla $m \in S$. Bevisa att alla icke-tomma delmängder av $\mathbb{N}$ har ett minsta element.

Ledtråd: Omformulera satsen och kombinera ett induktionsbevis med motsägelsebevis.

Övning 1.26 ($\star$)**.** Följande påståenden är alla sanna. Vilka bör vara satser och vilka bör vara definitioner? Motivera hur du tänker.

(i) Det finns både udda och jämna kvadrattal.

(ii) Ett tal n är ett kvadrattal om det finns en kvadrat vars area är n.

(iii) Ett tal n är ett kvadrattal om det finns ett heltal k så att $n = k^2$.

(iv) Det n:te kvadrattalet är summan av de n första udda talen.

Övning 1.27 ($\star$)**.** Följande påståenden är alla sanna. Vilka bör vara satser och vilka bör vara definitioner? Motivera hur du tänker.

(i) En rätvinklig triangel har vinkelsumma på 180 grader.

(ii) En rätvinklig triangel har två vinklar vars summa är 90 grader.

(iii) En rätvinklig triangel bildas när man drar en diagonal i en kvadrat.

(iv) En triangel är rätvinklig om den har en rät vinkel.

2 Euklidiska rum

När man för första gången introduceras till funktioner i matematik pratar man nästan enbart om funktioner av en variabel. I praktiken är dock detta ofta ett enkelt specialfall, då de flesta relationer ”i verkligheten” beror på mer än en variabel. Till exempel beror hur mycket pengar ett stort företag omsätter under ett år på näst intill oändligt många variabler, vilket gör att man med nödvändighet måste analysera en förenklad modell för att kunna dra några slutsatser. Men även en förenklad modell lär bero på väldigt många variabler, som till exempel antalet sålda produkter, pris på produkterna, antalet anställda, momssatser, arbetsgivaravgifter, inkomstskatter etc. Även enklare formler från fysik innehåller ofta mer än en variabel. Till exempel ges den kinetiska energin för en punktformig kropp med massan m och hastigheten v av $mv^2/2$.

I den här kursen kommer vi studera operationer som tar funktioner som *argument*; som till exempel operationen I som tar en funktion $f(x)$ och ger integralen på $[0, 1]$ av dess absolutbelopp i kvadrat:

$$I : f(x) \mapsto \int_0^1 |f(x)|^2 dx.$$

Men för att förstå dessa operationer behöver vi först analysera och förstå funktioner som beror av flera reella variabler; som till exempel $f(x, y, z) = e^{xy}/z$ eller $g(x, y, z) = xyz^2$, som då båda är funktioner av 3 reella variabler.

För vilka värden på (x, y, z) är funktionerna f och g definierade? I vilken punkt antar $f(x, y, z)$ sitt största respektive minsta värde? Finns det ens något största respektive minsta värde? Vad är funktionens maximum, om det nu existerar, om vi antar att $1/2 \leqslant x \leqslant y \leqslant z \leqslant 1$. Om vi börjar i punkten $(x, y, z) = (1, 1, 1)$, åt vilket håll ska vi då röra oss för att funktionen ska växa så snabbt som möjligt?

För att kunna analysera funktioner av flera variabler ordentligt och besvara frågorna ovan måste vi till att börja med förstå mängderna på vilka de är definierade.

2.1 Rummet $\mathbb{R}^n$

I den här kursen, och i *flervariabelanalys* i allmänhet, betraktar vi huvudsakligen funktioner som tar en *reell n-tupel* som argument, det vill säga argumentet är på formen $(x_1, \dots, x_n)$, där var och en av de n komponenterna x_j, $j = 1, 2, \dots, n$ är ett reellt tal, det vill säga $x_j \in \mathbb{R}$.

Vi använder beteckningen $\mathbb{R}^n$ för att beteckna mängden

$$\{(x_1, \dots, x_n) : x_j \in \mathbb{R}, \quad j = 1, \dots n\}.$$

Normalt skriver vi dock bara $\mathbb{R}$ istället för $\mathbb{R}^1$. Element i $\mathbb{R}$ kallas ofta för *skalärer*.

Ibland vill man av olika skäl skriva ut komponenterna i en kolonn istället för på en rad. Det vill säga man skriver

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Ibland kommer vi även använda fet stil för att beteckna punkter i $\mathbb{R}^n$. Då är det underförstått att den j:te komponenten i $\mathbf{x}$ är x_j, och att x_j då är ett reellt tal. Ibland skriver vi ändå ut $\mathbf{x} = (x_1, \ldots, x_n)$ för att vara extra tydliga.

Däremot är inte alla funktioner i n variabler definierade på hela $\mathbb{R}^n$. Till exempel är funktionen från vårt föregående exempel, $f(x, y, z) = e^{xy}/z$, inte definierad om $z = 0$. Precis som tidigare är *definitionsmängden*, som vi vanligtvis kommer beteckna med D_f, en viktig del av funktionen, och en funktions egenskaper beror i hög grad på såväl definitionsmängden som på själva regeln.

I kapitel 4 då vi mer noggrant undersöker funktioner av flera variabler och deras egenskaper kommer vi att definiera kontinuitet ordentligt, men följande exempel är förhoppningsvis tydligt ändå, även om vi tills vidare nöjer oss med en intuitiv känlsa av vad kontinuitet innebär.

Exempel 2.1.1. Om vi till exempel betraktar funktionen

$$f(x, y) = \begin{cases} \frac{x}{y} \text{ om } y \neq 0 \\ 0 \text{ om } y = 0 \end{cases}$$

så ser vi att den är kontinuerlig för en mängd som inte innehåller någon punkt där $y = 0$, det vill säga någon punkt på formen $(x, 0)$. Till exempel är funktionen kontinuerlig på mängden

$$D_f = \{(x, y) : 1 < x < 2, 1 < y < 2\}.$$

Däremot är funktionen inte kontinuerlig i alla punkter på mängden

$$D_f = \{(x, y) : -2 < x < 2, -2 < y < 2\}.$$

Till exempel är ju

$$\lim_{y \to 0} f(1, y) = \lim_{y \to 0} \frac{1}{y} \neq 0 = f(1, 0).$$

Huruvida en funktion är kontinuerlig eller ej beror alltså till viss del på definitionsmängden. ▲

Såväl $\mathbb{R}^n$ som delmängder till $\mathbb{R}^n$ är i grunden bara mängder, det vill säga samlingar med punkter. Men väldigt ofta väljer vi att lägga till ytterligare *struktur* på dessa mängder. På samma sätt som vi har matematiska operationer på reella tal — som att vi till exempel kan addera reella tal med varandra och få ett nytt reellt tal, och multiplicera reella tal med varandra och få ett nytt reellt tal — kan vi *definiera* operationer på n-tuplar i $\mathbb{R}^n$.

Till att börja med definierar vi tre operationer i $\mathbb{R}^n$, nämligen *addition*, *multiplikation* med ett reellt tal $\lambda \in \mathbb{R}$, samt *skalärmultiplikation*.

Om $\mathbf{x} = (x_1, \ldots, x_n)$ och $\mathbf{y} = (y_1, \ldots, y_n)$ definierar vi

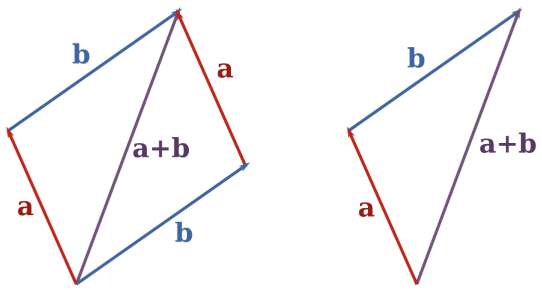
$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n).$$

$$\lambda\mathbf{x} = (\lambda x_1, \ldots, \lambda x_n).$$

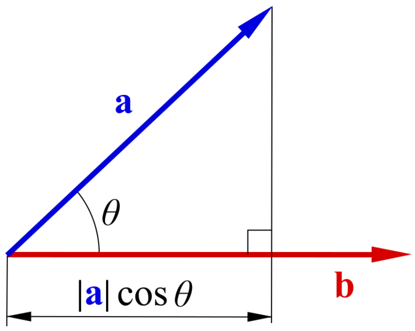
samt

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

Notera att addition och multiplikation med ett reellt tal ger nya element i $\mathbb{R}^n$, medan skalärprodukten av två element i $\mathbb{R}^n$ blir ett tal i $\mathbb{R}$ (det är därför operationen kallas just *skalär*produkt). En bild som visar hur det ser ut när vi adderar 2 vektorer finns i Figur 2.1 och en bild som visar vad skalärprodukten representerar finns i Figur 2.2.



Figur 2.1: Addition av 2 element $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$, bild från Wikipedia



Figur 2.2: Skalärprodukten $\mathbf{a} \cdot \mathbf{b}$ av 2 element, $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$, bild från Wikipedia

Det är värt att verifiera att vanliga formler för multiplikation och addition av reella tal också gäller med operationerna ovan. Till exempel är

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}, \quad \lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}, \quad \mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}.$$

I två och tre variabler kan punkter i $\mathbb{R}^n$ realiseras geometriskt som pilar i planet respektive rummet. Med denna geometriska tolkning kan addition av två element i $\mathbb{R}^n$ tolkas som att man helt enkelt klistrar fast den andra pilens början på den första pilens spets (se Figur 2.1), och multiplikation med ett positivt reellt tal skalar bara om längden på pilen. Till exempel kommer $\mathbf{x}$ och $2\mathbf{x}$ båda vara pilar som pekar åt samma håll, bara att $2\mathbf{x}$ är dubbelt så lång. Multiplicerar vi istället med ett negativt tal λ kommer vi att få en pil som pekar i motsatt riktning, och vars längd har blivit omskalad proportionerligt med absolutbeloppet av λ.

Vi behåller dessutom terminologin från $\mathbb{R}^2$ och $\mathbb{R}^3$ och säger att två element $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ är *parallella* om det finns ett reellt tal λ så att

$$\lambda \mathbf{x} = \mathbf{y}.$$

Om $\lambda \geqslant 0$ säger vi dessutom att $\mathbf{x}$ och $\mathbf{y}$ har *samma riktning* eller är *lika riktade*.

Vanligtvis kommer vi bara skriva 0 för *nollvektorn* i $\mathbb{R}^n$, alltså $(0, \dots, 0)$. Det bör alltid vara uppenbart från sammanhanget om 0 syftar på det reella talet 0 eller en n-tupel vars komponenter alla är noll. Notera att med definitionen ovan är alla element i $\mathbb{R}^n$ parallella och lika riktade med 0. Vissa väljer dock att i definitionen kräva att två parallella vektorer båda ska vara nollskilda just för att undvika detta.

Skalärproduktens geometriska tolkning är inte lika uppenbar, och den kommer inte riktigt behövas i den här kursen. Men det kan vara värt att veta att skalärprodukten är relaterad till vinkeln mellan två vektorer och deras längder enligt formeln $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos(\theta)$, där θ är vinkeln mellan vektorerna $\mathbf{x}$ och $\mathbf{y}$ (se Figur 2.2).

Ett viktigt specialfall från den geometriska tolkningen med pilar i två och tre variabler är att

$$\mathbf{x} \cdot \mathbf{x} = x_1^2 + x_2^2 (+x_3^2),$$

vilket vi känner igen från Pythagoras sats som kvadraten av längden av pilen.

Med detta som motivation definierar vi nu *längden* eller *beloppet* av ett element i $\mathbf{x} \in \mathbb{R}^n$ som

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Notera att skalärprodukten av ett element $\mathbf{x}$ med sig själv alltid är större än eller lika med 0, med likhet om och endast om $\mathbf{x} = 0$. Däremot är skalärprodukten av två olika element inte nödvändigtvis positiv.

Exempel 2.1.2. Låt $\mathbf{x} = (1, 2, 3, 4)$ och $\mathbf{y} = (-1, -1, -1, -1)$. Båda dessa vektorer är uppenbarligen nollskilda, men de är inte parallella. Vidare är

$$|\mathbf{x}| = \sqrt{1^2 + 2^2 + 3^2 + 4^2} = \sqrt{1 + 4 + 9 + 16} = \sqrt{30},$$

och

$$|\mathbf{y}| = \sqrt{(-1)^2 + (-1)^2 + (-1)^2 + (-1)^2} = \sqrt{4} = 2.$$

Slutligen är

$$\mathbf{x} \cdot \mathbf{y} = -1 - 2 - 3 - 4 = -10.$$

▲

Slutligen definierar vi avståndet mellan två punkter $\mathbf{x}$ och $\mathbf{y}$ som

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

Notera att detta överensstämmer med hur vi mäter avstånd i $\mathbb{R}^2$ och $\mathbb{R}^3$ med hjälp av Pythagoras sats.

När vi tänker på element i $\mathbb{R}^n$ utrustade med *strukturen* ovan — det vill säga att vi kan mäta deras längd, vi kan addera dem, vi kan multiplicera dem med reella tal, och vi kan ta skalärprodukter — så säger vi att dessa element är *Euklidiska vektorer*, eller bara *vektorer*. En vektor är alltså inte bara en punkt i rummet, utan den har mer struktur, som till exempel en längd och en riktning.

2.2 Några viktiga olikheter

En oerhört viktig olikhet, som vi bland annat kommer att behöva för att kunna avgöra i vilken riktining en funktion växer snabbast, är *Cauchy–Schwarz' olikhet.*

Sats 2.2.1. *(Cauchy–Schwarz' olikhet)*
I $\mathbb{R}^n$ gäller

$$|\mathbf{x}\cdot\mathbf{y}| \leqslant |\mathbf{x}||\mathbf{y}|, \tag{2.1}$$

eller utskrivet

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

Likhet inträffar om och endast om $\mathbf{x}$ *och* $\mathbf{y}$ *är parallella.*

Bevis. Vi kan anta att $\mathbf{x} \neq 0$, för om $\mathbf{x} = 0$ är båda leden i (2.1) lika med 0, och $\mathbf{x}$ och $\mathbf{y}$ är dessutom parallella.

Den viktiga insikten som beviset bygger på är det enkla faktum att skalärprodukten av *vilken vektor som helst* med sig själv är större än eller lika med 0.

Låt t vara ett reellt tal och betrakta vektorn $t\mathbf{x}+\mathbf{y}$. Enligt räknereglerna för skalärprodukten är

$$0 \leqslant (t\mathbf{x}+\mathbf{y})\cdot(t\mathbf{x}+\mathbf{y}) = t^2|\mathbf{x}|^2 + 2t\mathbf{x}\cdot\mathbf{y} + |\mathbf{y}|^2.$$

Eftersom $|\mathbf{x}|^2 \neq 0$ kan vi bryta ut $|\mathbf{x}|^2$ och kvadratkomplettera med avseende på t, vilket ger

$$0 \leqslant |\mathbf{x}|^2\left(t^2 + 2\frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}|^2}t + \frac{|\mathbf{y}|^2}{|\mathbf{x}|^2}\right) = |\mathbf{x}|^2\left(t + \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}|^2}\right)^2 + |\mathbf{y}|^2 - \frac{(\mathbf{x}\cdot\mathbf{y})^2}{|\mathbf{x}|^2}.$$

Notera att denna olikhet håller för *alla* reella värden på t. Om vi nu väljer ett lämpligt värde på t, nämligen

$$t = -\frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}|^2},$$

så försvinner den första termen i uttrycket ovan och vi får då

$$0 \leqslant |\mathbf{y}|^2 - \frac{(\mathbf{x}\cdot\mathbf{y})^2}{|\mathbf{x}|^2},$$

Vilket kan skrivas om till (2.1).

Från beviset framgår det även att vi har likhet i (2.1) precis då skalärprodukten av $(t\mathbf{x}+\mathbf{y})$ med sig själv är noll. Men det händer bara om $(t\mathbf{x}+\mathbf{y}) = 0$, vilket innebär att

$$\mathbf{y} = -t\mathbf{x}.$$

Så $\mathbf{x}$ och $\mathbf{y}$ måste alltså vara parallella om vi har likhet. $\square$

En annan väldigt viktig olikhet i $\mathbb{R}^n$ är följande generalisering av Sats 1.4.2 som säger att längden av hypotenusan i en rätvinklig triangel är kortare än summan av kateternas längder.

Sats 2.2.2. *(Triangelolikheten i $\mathbb{R}^n$)*
I $\mathbb{R}^n$ gäller

$$|\mathbf{x}+\mathbf{y}| \leqslant |\mathbf{x}| + |\mathbf{y}|,$$

med likhet precis när $\mathbf{x}$ och $\mathbf{y}$ är parallella och lika riktade.

Bevis. Genom att använda räknereglerna för skalärprodukten ser vi att

$$|\mathbf{x}+\mathbf{y}|^2 = (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) = |\mathbf{x}|^2 + 2\mathbf{x}\cdot\mathbf{y} + |\mathbf{y}|^2. \tag{2.2}$$

Genom att använda Cauchy–Schwarz' olikhet ser vi att

$$\mathbf{x}\cdot\mathbf{y} \leqslant |\mathbf{x}\cdot\mathbf{y}| \leqslant |\mathbf{x}||\mathbf{y}|,$$

Insättning i (2.2) ger nu att

$$|\mathbf{x}+\mathbf{y}|^2 \leqslant |\mathbf{x}|^2 + 2|\mathbf{x}||\mathbf{y}| + |\mathbf{y}|^2 = (|\mathbf{x}| + |\mathbf{y}|)^2,$$

vilket är ekvivalent med den sökta olikheten.

Vidare har vi likhet i triangelolikheten om och endast om vi har likhet i Cauchy–Schwarz' olikhet och

$$\mathbf{x}\cdot\mathbf{y} = |\mathbf{x}\cdot\mathbf{y}|.$$

Vi vet att vi har likhet i Cauchy–Schwarz' olikhet exakt då $\mathbf{x}$ och $\mathbf{y}$ är parallella, det vill säga då $\mathbf{x} = \lambda\mathbf{y}$ för något $\lambda \in \mathbf{R}$.

Villkoret att

$$\mathbf{x}\cdot\mathbf{y} = |\mathbf{x}\cdot\mathbf{y}|$$

betyder alltså att

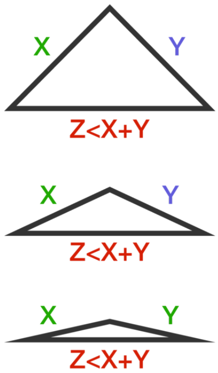
$$\mathbf{x}\cdot\mathbf{y} = \lambda\mathbf{y}\cdot\mathbf{y} = \lambda|\mathbf{y}|^2 = |\lambda||\mathbf{y}|^2 = |\lambda|\mathbf{y}|^2| = |\mathbf{x}\cdot\mathbf{y}|,$$

vilket håller precis då $\lambda = |\lambda|$, alltså när $\lambda \geqslant 0$.

Vi har alltså likhet i triangelolikheten precis då

$$\mathbf{x} = \lambda\mathbf{y}, \quad \lambda \geqslant 0,$$

det vill säga då $\mathbf{x}$ och $\mathbf{y}$ är parallella och lika riktade. $\square$



Figur 2.3: Triangelolikheten för $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, bild från `brilliant.org`

Figur 2.3 visar triangelolikheten med 2 vektorer $\mathbf{x}$ och $\mathbf{y}$. Ett viktigt specialfall av triangelolikheten är den så kallade *omvända triangelolikheten*.

Eftersom

$$|\mathbf{x}| = |(\mathbf{x} + \mathbf{y}) - \mathbf{y}| \leqslant |\mathbf{x} + \mathbf{y}| + |\mathbf{y}|$$

är

$$|\mathbf{x}| - |\mathbf{y}| \leqslant |\mathbf{x} + \mathbf{y}|.$$

På samma sätt visas att

$$|\mathbf{y}| - |\mathbf{x}| \leqslant |\mathbf{x} + \mathbf{y}|,$$

och dessa två olikheter ger tillsammans att

$$||\mathbf{x}| - |\mathbf{y}|| \leqslant |\mathbf{x} + \mathbf{y}|. \tag{2.3}$$

Olikheten (2.3) kallas ofta för den *omvända triangelolikheten*.

Med hjälp av induktion kan dessutom Sats 2.2.2 generaliseras till fler än två termer:

$$|\mathbf{x_1} + \ldots + \mathbf{x_n}| \leqslant |\mathbf{x_1}| + \ldots + |\mathbf{x_n}|.$$

Vi lämnar beviset av denna generalisering som en övningsuppgift.

2.3 Delmängder av $\mathbb{R}^n$

I många fall är man intresserade av att analysera en *delmängd* av $\mathbb{R}^n$. Till exempel kan man vara intresserad av att hitta den kortaste vägen att gå för att ta sig över ett bergspass, och i det fallet är den naturliga mängden att utföra sina analyser på den tvådimensionella delmängden till $\mathbb{R}^3$ som beskriver bergsytan av bergspasset.

Notera att vi i exemplet ovan är intresserde av att betrakta operationen som *tar* en kurva definierad på bergsytan och som förbinder två på förhand valda punkter på varsin sida av bergspasset, och som *ger* längden av den kurvan. Därefter vill vi hitta minimum av den operationen. I kapitel 7 härleder vi en metod för att hitta minimum och maximum till denna typ av operation.

I den här kursen är vi dock huvudsakligen intresserade av delmängder till $\mathbb{R}^n$ som uppkommer som naturliga definitionsmängder till funktioner vi är intresserade av att studera.

Exempel 2.3.1. Betrakta funktionen

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

Denna funktion är inte definierad för $x = 0, x = 1, y = 0$ och $y = 1$. Vi skulle till exempel kunna välja definitionsmängden som hela $\mathbb{R}^2$, bortsett från de fyra problemlinjerna $x = 0, x = 1, y = 0, y = 1$.

Den mängden är dock inte *sammanhängande* (notera att vi inte har definierat detta koncept ordentligt, så här får vi förlita oss på vår intuition om vad en sammanhängande mängd bör vara). Om vi vill ha en definitionsmängd som är sammanhängande skulle vi istället kunna betrakta mängden

$$D_f = \{(x,y) : 1 < x < 2, 1 < y < 2\}.$$

Notera att denna definitionsmängd inte är maximal eftersom vi kan *utvidga* definitionsmängden och få en strikt större definitionsmängd som också är sammanhängande. Till exempel kan vi betrakta

$$D_f = \{(x,y) : 1 < x < 3, 1 < y < 3\}.$$

Ytterligare en alternativ definitionsmängd är

$$D_f = \{(x,y) : 0 < x < 1, 0 < y < 1\}.$$

Denna definitionsmängd är dock maximal eftersom $f(x,y)$ inte är väldefinierad överallt på någon sammanhängande mängd som innehåller D_f som äkta delmängd. Problemet är att en sådan mängd med nödvändighet måste innehålla någon punkt vars x- eller y-koordinat är antingen 0 eller 1.

Det är dock viktigt att vara medveten om att vi *inte* har bevisat påståendet ovan. Vi har ju inte ens definierat vad vi menar med sammanhängande mängd ordentligt! ▲

Några av de vanligaste mängderna man intresserar sig för är *sfärer* och *klot*. Definitionen av dessa motiveras av att en cirkel med radie r och centrum a i två variabler beskrivs som mängden av punkter vars avstånd till a är precis r, och motsvarande (öppna) disk består av de punkter vars avstånd till a är strikt mindre än r.

Definition 2.3.2. I $\mathbb{R}^n$ definieras ett öppet klot som en mängd som kan beskrivas som

$$\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{a}| < r\}$$

för något $r > 0$ och någon punkt $\mathbf{a} \in \mathbb{R}^n$. I så fall kallas r för klotets radie och $\mathbf{a}$ för klotets medelpunkt.

Vidare definieras en sfär som en mängd som kan beskrivas som

$$\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{a}| = r\}$$

för något $r > 0$ och någon punkt $\mathbf{a} \in \mathbb{R}^n$. På samma sätt som för sfären kallas i så fall r för sfärens radie och $\mathbf{a}$ för sfärens medelpunkt.

Vi kommer ibland använda notationen $B_r(\mathbf{a})$ och $S_r(\mathbf{a})$ för att beteckna det öppna klotet respektive sfären med radie r och centrum $\mathbf{a}$. $\triangle$

Notera att en öppen sfär med radie $r > 0$ och centrum a i $\mathbb{R}^1$ bara blir det öppna intervallet $(a - r, a + r)$. Mer specifikt blir *det öppna enhetsklotet med centrum i origo* bara intervallet $(-1, 1)$.

Många viktiga definitioner i analys kräver att vi inte bara kan undersöka en funktion i en punkt, utan att vi kan uttala oss om funktionens beteende i ett litet område runt punkten. Till exempel kräver derivatans definition

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

att vi kan säga något om $f(x + h)$, i alla fall för *tillräckligt* små värden på h.

Om vi har en funktion $f(x)$ som bara är definierad på intervallet $[0, 1]$ så kan vi inte uttala oss om derivatan i punkten 1, för $f(1 + h)$ är inte definierat för *något* $h > 0$. I detta fall kan vi förvisso betrakta gränsvärdet från vänster ($h < 0$), men om vi har en funktion som är definierad på $[0, 1] \cup \{2\}$ finns det inget meningsfullt sätt att prata om en derivata i punkten 2.

Vi behöver alltså kunna prata om punkter i en mängd, som har egenskapen att ett litet område runt punkten också ligger i mängden.

Definition 2.3.3. Låt $M \subset \mathbb{R}^n$ och låt $\mathbf{a}$ vara en punkt i $\mathbb{R}^n$. Vi säger att:

- $\mathbf{a}$ är en *inre punkt* till M om det finns ett öppet klot med centrum i $\mathbf{a}$ som ligger helt i M.
- $\mathbf{a}$ är en *yttre punkt* till M om det finns ett öppet klot med centrum i $\mathbf{a}$ som ligger helt i komplementet till M.
- $\mathbf{a}$ är en *randpunkt* till M om *varje* öppet klot med centrum i $\mathbf{a}$ innehåller punkter från både M och komplementet till M.

$\triangle$

Notera att varje punkt i $\mathbb{R}^n$ tillhör precis en av de ovanstående tre kategorierna med avseende på en given mängd M.

Det är huvudsakligen konceptet av en inre punkt vi är intresserade av i den här kursen, då det är i inre punkter av en definitionsmängd vi (eventuellt) kan definiera koncept i stil med derivata.

Det är viktigt att vara medveten om att huruvida en punkt $\mathbf{a} \in M$ är en inre punkt eller ej inte bara beror på M, utan även på den omkringliggande mängden $\mathbb{R}^n$.

Låt till exempel $I \subset \mathbb{R}$ vara intervallet

$$I = \{x \in \mathbb{R} : 0 < x < 2\}.$$

Då är punkten $x = 1$ en inre punkt till M, för till exempel är det öppna klotet med centrum 1 och radie $1/2$ en delmängd av M.

Men för motsvarande intervall som en delmängd till $\mathbb{R}^2$

$$I = \{(x, 0) \in \mathbb{R}^2 : 0 < x < 2\}$$

har vi nu att *inget* öppet klot med centrum i $(1, 0)$ är en delmängd till I. Till exempel har vi för varje $\delta > 0$ att

$$(1, \delta/2) \in B_\delta(1),$$

men $(1, \delta/2) \notin I$.

Problemet är i någon bemärkelse att den omkringliggande mängden är större relativt I, ocn att öppna klot därför börjar innehålla en massa extra punkter utanför I.

Definition 2.3.4. En mängd $M \subset \mathbb{R}^n$ kallas *öppen* om alla punkter i M är *inre* punkter till M. Den kallas sluten om alla randpunkter till M ligger i M. △

Till exempel är det öppna intervallet $(0, 1) \subset \mathbb{R}$ en *öppen* mängd, och det slutna intervallet $[0, 1]$ är en *sluten* mängd då dess enda randpunkter är 0 och 1, och båda dessa ligger i $[0, 1]$. Att visa detta ordentligt lämnas som en övningsuppgift. Det är också värt att tänka igenom och verfiera att alla öppna klot faktiskt är öppna mängder enligt definitionen ovan. Att visa detta lämnas också som en övningsuppgift.

Notera dock att det finns mängder som varken är öppna eller slutna. Till exempel är $[0, 1)$ inte öppen då 0 inte är en inre punkt, men den är inte sluten heller då den har 1 som randpunkt, trots att 1 inte ligger i mängden.

Definition 2.3.5. Det *inre* av en mängd $M \subset \mathbb{R}^n$ är delmängden till M som består av alla inre punkter. Det vill säga mängden

$$\{x \in M : x \text{ är en inre punkt i } M\}.$$

Denna mängd betecknas ofta med M°. △

Exempel 2.3.6. Det inre av både det slutna och det halvöppna intervallen $[0, 1]$ och $[0, 1)$ är det öppna intervallet $(0, 1)$. Anledningen till detta är att alla punkter i $(0, 1)$ är inre, medan punkterna 0 och 1 är randpunkter. Att visa detta lämnas som en övning. ▲

Övningar

Övning 2.1. Låt $\mathbf{x} = (1, 3, 5, 7)$ och $\mathbf{y} = (-1, 3, -5, 1)$.

(a) Beräkna $|\mathbf{x}|$ och $|\mathbf{y}|$.

(b) Beräkna $\mathbf{x} \cdot \mathbf{y}$.

(c) Är $\mathbf{x}$ och $\mathbf{y}$ parallella?

Övning 2.2. Låt $\mathbf{x} = (1, 2, 3)$ och $\mathbf{y} = (-3, 2, -1)$.

(a) Verifiera Cauchy–Schwarz' olikhet för detta val av $\mathbf{x}$ och $\mathbf{y}$.

(b) Verifiera triangelolikheten för detta val av $\mathbf{x}$ och $\mathbf{y}$.

Övning 2.3 ($\star$)**.** Använd induktion för att bevisa att

$$|x_1 + \ldots + x_n| \leqslant |x_1| + \ldots + |x_n|,$$

för alla positiva heltal $n \geqslant 2$.

Övning 2.4 ($\star$)**.** Bevisa att det slutna intervallet $[0, 1] \subset \mathbb{R}$ är en *sluten* mängd.

Övning 2.5 ($\star$)**.** Bevisa att det öppna intervallet $(0, 1) \subset \mathbb{R}$ är en *öppen* mängd.

Övning 2.6 ($\star\star$)**.** Låt $r > 0$ och $\mathbf{a} \in \mathbb{R}^n$ och betrakta det öppna klotet $B_r(\mathbf{a})$. Visa att $B_r(\mathbf{a})$ är en *öppen* mängd.

Övning 2.7 ($\star\star$)**.** Beräkna det inre av den slutna enhetsskivan

$$D := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 1\}.$$

Övning 2.8 ($\star\star$)**.** Beräkna det inre av den slutna kvadraten

$$K := \{(x, y) \in \mathbb{R}^2 : |x| \leqslant 1, |y| \leqslant 1\}.$$

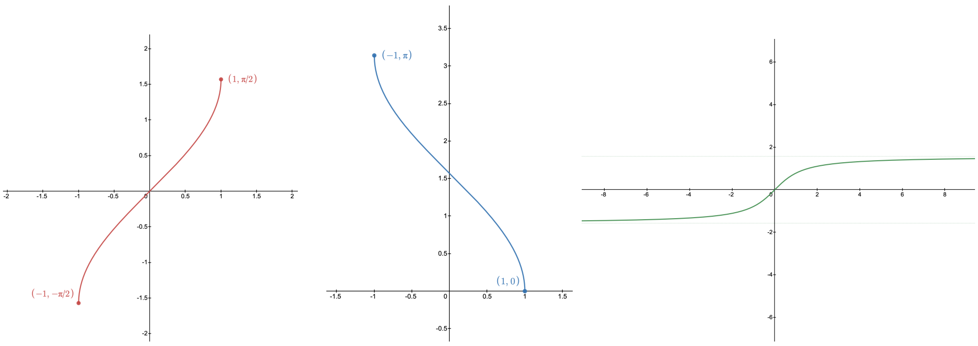
3 Lösningsmetoder för ordinära differentialekvationer

Målet med årets Matematiska Cirkel är att lära oss variationskalkyl för att lösa matematiska problem såsom Brachistochronproblemet. Variationskalkylen kommer att reducera våra problem, exempelvis Brachistochronproblemet, till att lösa en så kallad *differentialekvation*. I detta kapitel ser vi till att lägga en stabil grund för att kunna lösa differentialekvationerna som vi stöter på i senare kapitel. Vi utgår ifrån att en majoritet av eleverna som läser detta kapitel har stött på definitionen av kontinuitet, derivata och integral tidigare.

3.1 Repetition av envariabelanalys

Detta avsnitt är enbart ämnat för att repetera vissa definitioner och agera som ett formelblad för dem som har glömt eller inte sett begreppen förut. Vi kommer därav inte bevisa någonting i detta korta avsnitt. Om man inte hört talas om konceptet derivata förut så är övningstillfällena ämnade för er att komma och ställa frågor, så fråga om derivata då!

Anmärkning 3.1.1. Vi påminner läsaren om ett par speciella funktioner.



(a) arcsin **(b)** arccos **(c)** arctan

Vi listar deras definitions- och värdemängder i denna tabell.

	Definitionsmängd	Värdemängd
arcsin	$[-1, 1]$	$\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$
arccos	$[-1, 1]$	$[0, \pi]$
arctan	$\mathbb{R}$	$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

Dessa är de inversa funktioner på de vanliga trigonometriska funktioner på lämpliga delmängder, d.v.s.

$$\begin{aligned} x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] &\implies \arcsin(\sin(x)) = x \\ x \in [0, \pi] &\implies \arccos(\cos(x)) = x \\ x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) &\implies \arctan(\tan(x)) = x. \end{aligned}$$

Definition 3.1.2. Låt $f : I^\circ \to \mathbb{R}$, för något intervall I. En kontinuerlig funktion $F : I \to \mathbb{R}$ sägs vara en primitiv funktion till f på I om $F'(x) = f(x)$, för varje $x \in I^\circ$. △

Anmärkning 3.1.3. Intervallet I kan vara på formen

$$[a, b], [a, b), (a, b], (a, b), [a, \infty), (a, \infty), (-\infty, b], (-\infty, b), \mathbb{R}.$$

Exempel 3.1.4. Här är en lista på typiska funktioner med derivata och primitiv funktion, till var och en. Notera att alla andra primitiva funktioner erhålles genom att addera en konstant till den vi anger.

f'	f	F	Definitionsmängd
nx^{n-1}	x^n,	$\frac{x^{n+1}}{n+1}$	$x \in \mathbb{R} \quad n \neq -1$
e^x	e^x	e^x	$\mathbb{R}$
$\cos(x)$	$\sin(x)$	$-\cos(x)$	$\mathbb{R}$
$\frac{1}{x}$	$\ln(x)$	$x(\ln(x) - 1)$	x>0
$-\sin(x)$	$\cos(x)$	$\sin(x)$	$\mathbb{R}$
$1 + \tan^2(x)$	$\tan(x)$	$*$	$\|x\| < \frac{\pi}{2}$
$*$	$\frac{1}{1+x^2}$	$\arctan(x)$	$\mathbb{R}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$	$*$	$[-1, 1]$

Vi lämnar det som övning i slutet av kapitlet att fylla i de tomma rutorna i tabellen. ▲

Exempel 3.1.5. Notera att $F(x) = \sqrt{x}$ är en primitiv funktion till

$$f(x) = \begin{cases} \frac{1}{2\sqrt{x}} \text{ om } x > 0 \\ 0 \text{ om } x = 0 \end{cases}$$

på intervallet $I = [0, \infty)$. Detta då F är kontinuerlig på I och $F'(x) = f(x)$ för alla $x \in I^\circ = (0, \infty)$. Notera att värdet $f(0)$ inte spelar någon roll eftersom $x = 0$ inte är en inre punkt. ▲

Sats 3.1.6 (Analysens fundamentalsats). *Låt f vara en kontinuerlig funktion på intervallet $[a, b]$. Då gäller att*

$$F(x) := \int_a^x f(t)dt$$

är en primitiv funktion till f i $[a, b]$. Dessutom skiljer sig alla primitiva funktioner enbart åt med en additiv konstant.

Följdsats 3.1.7 (Variant av analysens fundamentalsats). *Låt f vara en kontinuerlig funktion på intervallet (a, b). Då är alla funktioner på formen*

$$F(x) = \int_c^x f(t)dt.$$

för $c \in (a, b)$ primitiva funktioner till f på (a, b).

Exempel 3.1.8. Antagandet att f är kontinuerlig på ett slutet intervall (dvs även i ändpunkterna) i analysens huvudsats är viktigt. Vi demonstrerar skillnaden i detta exempel. Låt

$$f(x) = \begin{cases} \frac{1}{x} \text{ om } x > 0 \\ 0 \text{ om } x = 0 \end{cases}$$

vara definerad på $[0, \infty)$, den är ej kontinuerlig i 0. Vi skulle kunna tro att

$$F(x) = \int_0^x \frac{1}{t} dt$$

är en primitiv funktion. Problemet är bara det att $\int_0^x \frac{1}{t} dt = \infty$ så det går inte att definera F såhär. Däremot vet vi att

$$F(x) = \int_1^x \frac{1}{t} dt = \ln(x)$$

är en primitiv funktion på intervallet $(0, \infty)$. ▲

Anmärkning 3.1.9. Från analysens huvudsats följer den så kallade *insättningsformeln* som man lär sig i gymnasiet. Om F är en primitiv funktion till f på $[a, b]$ så gäller

$$\int_a^b f(t)dt = F(b) - F(a).$$

Vi använder ofta notationen $[F(x)]_a^b := F(b) - F(a)$.

Hjälpsats 3.1.10. *Att integrera och derivera uppfyller följande räkneregler (så länge båda sidor är väldefinerade).*

- $\int_a^b f(t) + g(t)dt = \int_a^b f(t)dt + \int_a^b g(t)dt$
- $\forall \lambda \in \mathbb{R} \quad \int_a^b \lambda f(t)dt = \lambda \int_a^b f(t)dt$
- $\int_a^b f(t)dt + \int_b^c f(t)dt = \int_a^c f(t)dt$
- $\int_a^b f(t)dt = -\int_b^a f(t)dt$
- $(f + g)' = f' + g'$
- $\forall \lambda \in \mathbb{R} \quad (\lambda f)' = \lambda f'$
- $(f \cdot g)' = f' \cdot g + f \cdot g'$
- $(f(g(t))' = f'(g(t)) \cdot g'(t)$
- $\left(\frac{f}{g}\right)' = \frac{f' \cdot g - f \cdot g'}{g^2}.$

Exempel 3.1.11. Låt oss beräkna integralen

$$\int_0^1 x^2 + 3e^x dx.$$

På $[0, 1]$ har vi primitiva funktioner $\frac{x^3}{3}$ till x^2 och e^x till e^x. Då får vi,

$$\int_0^1 x^2 + 3e^x dx = \int_0^1 x^2 dx + \int_0^1 3e^x dx = \int_0^1 x^2 dx + 3\int_0^1 e^x dx$$
$$= \Big[x^3/3\Big]_0^1 + 3\Big[e^x\Big]_0^1 dx = \frac{1}{3} + 3(e-1).$$

▲

3.2 Tekniker för att integrera funktioner av en variabel

Många funktioner, och senare *funktionaler*, som vi kommer att stöta på i denna kurs defineras som primitiva funktioner och integraler av andra funktioner. Exempelvis så dyker många lösningar till differentialekvationer upp på detta vis. För att vi ska kunna explicit beräkna eller förstå denna typ av funktioner bättre så måste vi tillämpa ett flertal hjälpsatser och *integrationstekniker* som vi lär oss i detta kapitel.

Hjälpsats 3.2.1 (Partiell integration)**.** *Låt f, g och g' vara kontinuerliga i $[a, b]$ och låt F vara en primitiv funktion till f i $[a, b]$. Då gäller att*

$$\int_a^b f(x)g(x)dx = \Big[F(x)g(x)\Big]_a^b - \int_a^b F(x)g'(x)dx$$

Bevis. Analysens huvudsats ger oss att

$$\Big[F(x)g(x)\Big]_a^b = \int_a^b (F(x)g(x))'dx.$$

Med produktregeln för derivata kan vi ytterligare förenkla uttrycket till

$$\Big[F(x)g(x)\Big]_a^b = \int_a^b (F(x)g(x))'dx = \int_a^b f(x)g(x) + F(x)g'(x)dx =$$
$$\int_a^b f(x)g(x)dx + \int_a^b F(x)g'(x)dx.$$

Beviset är färdigt efter att man flyttar över ena termen i högerledet till vänsterledet. □

Exempel 3.2.2. Låt oss beräkna följande integral med hjälp av partiell integration,

$$\int_0^1 xe^x dx.$$

Vi låter $f(x) = e^x$ och $g(x) = x$. Då får vi $F(x) = e^x$ och $g'(x) = 1$, så när vi tillämpar partiell integration får vi

$$\int_0^1 xe^x dx = \Big[xe^x\Big]_0^1 - \int_0^1 e^x dx = \Big[xe^x\Big]_0^1 - \Big[e^x\Big]_0^1 = (e-0) - (e-1) = 1.$$

▲

En annan värdefull sats är följande:

Hjälpsats 3.2.3 (Variabelbyte)**.** *Låt g, g' vara kontinuerliga på $[a, b]$ och f kontinuerlig på värdemängden för g. Då gäller att*

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx.$$

Bevis. Analysens fundamentalsats och kedjeregeln för derivata ger oss att

$$\int_{g(a)}^{g(b)} f(t)dt = \Big[F(g(x))\Big]_a^b = \int_a^b F(g(x))'dx = \int_a^b f(g(x))g'(x)dx. \qquad \square$$

Anmärkning 3.2.4. Med g' kontinuerlig menar vi att g' har ett väldefinerat höger- respektive vänstergränsvärde i ändpunkterna på intervallet.

Exempel 3.2.5. Låt oss beräkna följande integral med hjälp av variabelbyte,

$$\int_0^3 2xe^{x^2}dx.$$

Vi låter $g(x) = x^2$ och $f(x) = e^x$. Då får vi $F(x) = e^x$, $g'(x) = 2x$, $g(0) = 0$ och $g(3) = 9$. Så när vi tillämpar variabelbyte får vi

$$\int_0^3 2xe^{x^2}dx = \int_0^9 e^t dt = e^9 - 1.$$

▲

Anmärkning 3.2.6. Variabelbyte i integral är ekvivalent med påståendet: om F är en primitiv funktion till f på värdemängden för g så är $F(g(x))$ primitiv till $f(g(x))g'(x)$ på $[a, b]$. När vi gör ett variabelbyte kommer vi att använda en lite annorlunda notation där vi använder bokstaven y istället för g. Såhär skriver vi då,

$$\int_{y(a)}^{y(b)} f(y)dy = \left\{ \substack{y=y(x) \\ x=a \implies y=y(a)} \quad \substack{dy=y'(x)dx \\ x=b \implies y=y(b)} \right\} = \int_a^b f(y(x))y'(x)dx.$$

Exempel 3.2.7. Låt oss beräkna följande integral med hjälp av variabelbyte,

$$\int_0^1 \frac{e^x}{1+e^{2x}}dx.$$

Vi låter $y(x) = e^x$ och $f(x) = \frac{1}{1+x^2}$. Då vet vi sedan tidigare att $F(x) = \arctan(x)$, $y'(x) = e^x$, $y(0) = 1$ och $y(1) = e$. Så när vi tillämpar variabelbyte får vi

$$\begin{aligned}\int_0^1 \frac{e^x}{1+e^{2x}}dx &= \int_1^e \frac{1}{1+y^2}dy = \Big[\arctan(y)\Big]_1^e \\ &= \arctan(e) - \arctan(1) = \arctan(e) - \frac{\pi}{4}.\end{aligned}$$

▲

3.3 Differentialekvationer

En ekvation $f(x) = 0$ har lösningar som motsvarar värden på variabeln x. I detta avsnitt studerar vi *differentialekvationer*, som beskriver förhållanden mellan en funktion, dess derivator och variabler. Dessa är i allmänhet svåra att lösa. Vi kommer att behandla funktioner $y(x) : I \to \mathbb{R}$, för något intervall I, och ännu mer specifikt differentialekvationer som kallas för *separabla* men först ska vi se på ett par enklare exempel.

Exempel 3.3.1. En differentialekvation som involverar en okänd funktion y som beror på enbart en variabel x kallas för *ordinär.* ▲

Anmärkning 3.3.2. Ofta är man lite slarvig med att specificera definitionsmängden av den okända funktionen $y(x)$ i differentialekvationer eftersom man inte vet vad den är. Ofta när man löser ekvationer, särskilt i denna kurs, kommer det naturligt fram vart lösningen är definerad eller ej.

Definition 3.3.3. En ordinär differentialekvation som kan skrivas på formen

$$y'(x) + g(x) \cdot y(x) = 0$$

där g är kontinuerlig kallas för en första ordningens homogen ordinär linjär differentialekvation. △

Hjälpsats 3.3.4. *Alla lösningar till en första ordningens homogen ordinär linjär differentialekvation är på formen*

$$y(x) = Ce^{-G(x)}$$

där G är en primitiv funktion till g och C en godtycklig konstant.

Bevis. En första ordningens ordinär linjär differentialekvation kan alltid skrivas om

$$\begin{aligned} y'(x) + g(x) \cdot y(x) = 0 &\iff \\ y'(x)e^{G(x)} + g(x)e^{G(x)} \cdot y(x) = 0 &\iff \\ (y(x)e^{G(x)})' = 0 &\iff \\ y(x)e^{G(x)} = C. &\quad \square \end{aligned}$$

Exempel 3.3.5. Några första ordningens ordinära linjära differentialekvationer med lösningar är

Ekvation	g(x)	G(x)	Lösning
$y'(x) - y(x) = 0$	-1	$-x$	$y(x) = Ce^{x}$
$x \cdot y'(x) + y(x) = 0$	$1/x$	$\ln(x)$	$y(x) = C/x$
$y'(x) + x \cdot y(x) = 0$	x	$x^2/2$	$y(x) = Ce^{-\frac{x^2}{2}}$
$y'(x) + \cos(x) \cdot y(x) = 0$	$\cos(x)$	$\sin(x)$	$y(x) = Ce^{-\sin(x)}$
$y'(x) + e^{x} \cdot y(x) = 0$	e^{x}	e^{x}	$y(x) = Ce^{-e^{x}}$

▲

Exempel 3.3.6. Några andra exempel på differentialekvationer är

$$f(x) \cdot f'(x) = 2x^3, \qquad g''(x) + 2g'(x) + g(x) = 0, \qquad \frac{y'(x)}{1 + (y(x))^2} = \tan(x).$$

Exempel på lösningar är

$$f(x) = x^2 \qquad g(x) = e^{-x} \qquad y(x) = -\tan(\log\cos(x))$$

▲

3.4 Separabla differentialekvationer

Definition 3.4.1. En ordinär differentialekvation kallas *separabel* om det finns två funktioner f, g i en variabel så att ekvationen kan skrivas om på formen

$$f(y(x)) \cdot y'(x) = g(x).$$

Här är x vår variabel och $y(x)$ funktionen vi försöker hitta. △

Anmärkning 3.4.2. Notera att valet av f och g inte alltid är unikt.

Exempel 3.4.3. Ett par exempel på separabla differentialekvationer är

$$y \cdot y' + x = 0, \qquad y' = y\ln(x) + y, \qquad \frac{y'}{1 + y^2} = \tan(x).$$

Exempel på val av f och g är

$$\begin{cases} f(y) = y \\ g(x) = -x \end{cases} \qquad \begin{cases} f(y) = \frac{1}{y} \\ g(x) = \ln(x) + 1 \end{cases} \qquad \begin{cases} f(y) = \frac{1}{1+y^2} \\ g(x) = \tan(x) \end{cases}.$$

▲

Definitionen av att en differentialekvation är separabel är konstruerad så att man kan lösa den med hjälp av integration vilket vi ser i följande hjälpsats. Den har ett kort bevis men där man måste ha väldig god kontroll över definitionsmängderna för funktionerna i fråga. Vi kommer att studera hur den används mer konkret i flera exempel.

Sats 3.4.4. *Givet en separabel differentialekvation på formen $f(y(x))y'(x) = g(x)$, låt F, G vara primitiva funktioner till f, g på intervallen I_y och I_x respektive. Då måste en lösning $y(x)$ på I_x som uppfyller $y(x) \in I_y$ även uppfylla*

$$F(y(x)) = G(x) + C,$$

för någon konstant C. Här är x vår variabel och $y(x)$ funktionen vi försöker hitta.

Bevis. Beviset är en direkt följd av ett lämpligt variabelbyte. Vi tar den separabla differentialekvationen och integrerar båda sidor och använder insättningsformeln för något $c \in I_x$,

$$\begin{aligned} f(y(x))y'(x) = g(x) &\iff \\ \int_c^x f(y(t))y'(t)dt = \int_c^x g(t)dt = G(x) - G(c) &\implies \text{variabelbyte} \\ \int_{y(c)}^{y(x)} f(y)dy = F(y(x)) - F(y(c)) = G(x) - G(c). \end{aligned}$$

Givet en fix lösning y så är C konstanten $F(y(c)) - G(c)$. □

Hjälpsats 3.4.5. *En ekvation på formen*

$$F(y(x)) = G(x),$$

har en unik lösning om F är inverterbar.

Bevis. Då F är inverterbar så har vi en unik lösning $y(x) = F^{-1}(G(x))$. □

Anmärkning 3.4.6. Notera att en funktion som inte är injektiv kan ha mer än en högerinvers. Ett typisk exempel är funktionen $\tan(x)$. Den har flera högerinverser $\arctan(x), \arctan(x) + \pi, \arctan(x) + 2\pi, \ldots$ Vi kommer att se vad detta har för inverkan i ett exempel nedan.

Anmärkning 3.4.7. Notera att en funktion är inverterbar om den är injektiv. För att visa att en funktion är injektiv är det tillräckligt att visa att den är strängt växande. Beviset lämnas till läsaren i övning 3.7.

Exempel 3.4.8. Låt oss använda sats 3.4.4 för att studera den separabla differentialekvationen

$$y(x)y'(x) + x = 0.$$

Vi har i tidigare exempel fastställt att $f(y) = y$ och $g(x) = -x$. Vi hittar primitiva funktioner

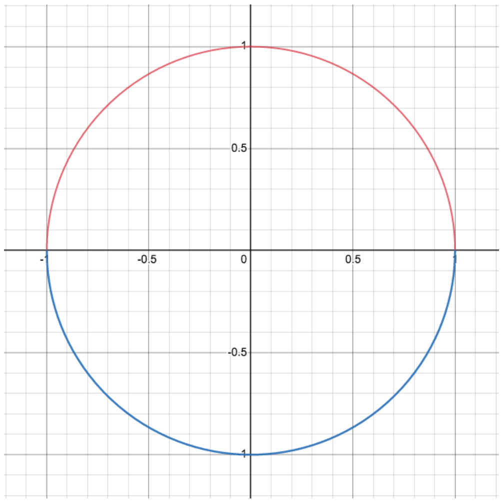
$$\begin{cases} F(y) = \frac{y^2}{2},\ y \in \mathbb{R} \\ G(x) = -\frac{x^2}{2}, x \in \mathbb{R} \end{cases}.$$

Oavsett vad $y(x)$ är så kommer $y(x)$ alltid ligga inom definitionsmängden för F och x i definitionsmängden för G. Vi kan då med gott samvete skriva ned

$$\forall x \in \mathbb{R} \qquad \frac{y^2}{2} = -\frac{x^2}{2} + C.$$

Det är tydligt att det inte finns någon vettig lösning om $C \leqslant 0$ eftersom $x^2, y^2 \geqslant 0$. Efter en liten omskrivning, $r^2 = 2C$ ser vi att detta är cirkelns ekvation

$$x^2 + y^2 = r^2.$$



Figur 3.2: En lösning till den givna differentialekvationen med $r^2 = 1$ måste alltså ligga inuti enhetscirkeln.

Som vi ser i bilden är inte detta en funktion som löser ekvationen, utan två stycken, den övre och den undre. Detta kan förklaras med att y^2 inte är injektiv! Däremot är den det om vi begränsar oss till $y \geqslant 0$ eller $y \leqslant 0$. Detta ger oss två olika lösningar för varje val av r^2,

$$y^2 = -x^2 + r^2 \implies y(x) = \pm\sqrt{-x^2 + r^2}.$$

Som vi ser är lösningarna bara definerade då $-x^2 + r^2 \geqslant 0$. ▲

Exempel 3.4.9. Låt oss använda sats 3.4.4 för att studera den separabla differentialekvationen

$$y' = y\ln(x) + y \qquad f(y) = \frac{1}{y} \qquad g(x) = \ln(x) + 1$$

Notera att f och g har en singularitet i $x = 0$ och $y = 0$. I den här kursen nöjer oss med att hitta lösningar för $I_y = (0, \infty)$ och $I_x = (0, \infty)$. Vi hittar primitiva funktioner

$$\begin{cases} F(y) = \ln(y),\ y > 0 \\ G(x) = x\ln(x), x > 0 \end{cases}.$$

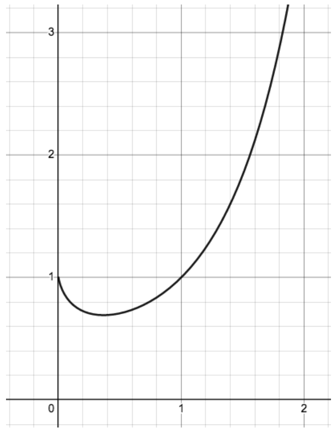
Vårt antagande $y(x) > 0$ garanterar att $y(x)$ alltid ligger inom definitionsmängden för F. Vi kan då skriva ned

$$x, y > 0 \implies \ln(y) = x\ln(x) + C.$$

Oavsett C får vi en vettig lösning som inte inskränker definitionsmängden på y. Med villkoret $y > 0$ har F en unik invers, $F^{-1}(t) = e^t$. Vi kan då skriva

$$\begin{aligned} x, y > 0 \implies \ln(y) = x\ln(x) + C &\iff \\ y = e^{x\ln(x)+C} &\iff \\ y = Dx^x \end{aligned}$$

för något $D = e^C > 0$. ▲



Figur 3.3: Lösningskruvan för $D = 1$.

Exempel 3.4.10. Låt oss använda sats 3.4.4 för att studera den separabla differentialekvationen

$$\frac{y'(x)}{1+(y(x))^2} = \tan(x) \qquad f(y) = \frac{1}{1+y^2} \qquad g(x) = \tan(x)$$

Notera att $\tan(x)$ har en singularitet i $x = \pm\frac{\pi}{2}$. Vi nöjer oss med att hitta en primitiv funktion på $(-\frac{\pi}{4}, \frac{\pi}{4})$. Vi hittar primitiva funktioner

$$\begin{cases} F(y) = \arctan(y),\ y \in \mathbb{R} \\ G(x) = -\ln(\cos(x)),\ x \in (-\frac{\pi}{4}, \frac{\pi}{4}) \end{cases}.$$

Att hitta $G(x)$ är ganska klurigt, se uppgift 3.3. Vi har även på det intervall vi betraktar en unik invers till $\arctan(y)$ som är tan(y) och vi har dessutom

$$|x| \leqslant \frac{\pi}{4} \implies \cos(x) \geqslant 1/2 \implies |\ln(\cos(x))| \leqslant \ln(2) < 1 < \frac{\pi}{2}.$$

I slutändan får vi ett exempel på en lösning,

$$y(x) = -\tan(\ln(\cos(x))).$$

▲

Exempel 3.4.11. Vi kommer senare i kapitel 8 att behöva lösa en differentialekvation på formen

$$(y')^2 = \frac{k^2}{y} - 1$$

för något $k \in \mathbb{R}$ och $y(x) \in [0, k^2]$. Den är inte separabel eftersom vi har en term $(y')^2$ som är knepig att bli av med. Låt oss för tillfället bara studera ekvationen och härleda några egenskaper hos lösningen genom att använda differentialekvationen som den löser och förvandla den till en separabel differentialekvation. Låt oss visa att $y''(x) \leqslant 0$, alltså att vår funktion är *konkav* vilket är samma sak som att y' är avtagande.

Vi ser att $y'(x) = 0 \iff y(x) = k^2$. Om $y(x) = k^2$ använder vi antagandet att $\forall t\ y(t) \leqslant k^2$, därför måste sådant x vara ett lokalt maximum vilket medför $y''(x) \leqslant 0$. Om istället $y'(x) \neq 0$ har vi att

$$(y')^2 = \frac{k^2}{y} - 1 \implies \frac{d}{dx}(y')^2 = \frac{d}{dx}\left(\frac{k^2}{y} - 1\right) \implies$$
$$2y'y'' = -\frac{k^2}{y^2}y' \overset{y'\neq 0}{\implies} y''(x) = -\frac{k^2}{2y^2} \leqslant 0.$$

Så vet vi att om $y'(x_0) = 0$ medför det

$$\begin{cases} x \geqslant x_0 \implies y'(x) \leqslant 0 \\ x \leqslant x_0 \implies y'(x) \geqslant 0 \end{cases}$$

Vi kan alltså dela upp vår differentialekvation i två stycken separabla differentialekvationer, en där $x \geqslant x_0$ och en där $x \leqslant x_0$ om det finns en punkt $y'(x_0) = 0$. Om $x \leqslant x_0$ kan vi skriva

$$y' = \sqrt{\frac{k^2}{y} - 1}.$$

På samma sätt får vi $y' = -\sqrt{\frac{k^2}{y} - 1}$ när $x \geqslant x_0$. ▲

Övningar

Övning 3.1. Beräkna integralen

$$\int_1^2 x^3 - 2x^2 + \pi dx.$$

Övning 3.2. Derivera

$$\frac{\text{arctan(x)}}{x}.$$

Övning 3.3 ($\star$)**.** Fyll i de två tomma rutorna i tabellen från exempel 3.1.4 och ange på vilka intervall de är definerade.
Ledtråd: För $\tan(x)$ *ställ upp uttrycket* $F(x) = \int_0^x \tan(\mathrm{t})dt$ *och gör variabelbytet* $y(x) = \cos(x)$ *och låt* $f(y) = \frac{1}{y}$*. För* arcsin *gör istället partiell integration med* $f(x) = 1$ *och* $g(x) = \arcsin(x)$*.*

Övning 3.4. Beräkna

$$\int_0^{\frac{\pi}{3}} 3t^2 - \cos(t) + \sin(2t)dt.$$

Övning 3.5. Beräkna en primitiv funktion till $\ln(x)$ då $x > 0$ genom att beräkna integralen

$$\int_1^x \ln(t)dt.$$

Ledtråd: Använd partiell integration med $g = \ln(t)$ *och* $f(t) = 1$*.*

Övning 3.6 ($\star$)**.** Studera differentialekvationen i exempel 3.4.11. Argumentera för att om y, y', y'' alla är kontinuerliga och $y'(x_0) = 0$ så måste $y''(x_0) = -\frac{1}{2k^2}$, givet att det finns ett öppet intervall I kring x_0 där $x \in T\backslash\{x_0\} \implies y'(x) \neq 0$.

Övning 3.7. Bevisa att en strängt växande funktion, alltså en funktion med egenskapen

$$x_1 < x_2 \implies f(x_1) < f(x_2)$$

är injektiv.

Övning 3.8. Antag att $y(x)$ är en lösning till

$$e^{(y'(\sqrt{x}))} + y(x) = 0 \qquad y(9) = -1.$$

Beräkna $y'(3)$.

Övning 3.9. Antag att $y(x)$ är en lösning till

$$(y')^2 = \frac{1}{y} - 1 \qquad y(\pi) = 1/2.$$

Vilka värden skulle $y'(\pi)$ kunna ha?

Övning 3.10. Antag att $y(x)$ är en lösning till

$$y'(x) + y'(x)^2 + y(x) + y(x)^2 = 2x \qquad y(1) = 1.$$

Beräkna vilka värden som $y'(1)$ kan anta.

Övning 3.11 ($\star$)**.** Verifiera att $F(x) = \frac{1}{2}\left(x\sqrt{1+x^2} + \ln(x + \sqrt{1+x^2})\right)$ är en primitiv funktion till $f(x) = \sqrt{1+x^2}$.

4 Derivering av funktioner av flera variabler

Ett av de viktigaste matematiska verktygen är *derivering*. Att verktyget är oerhört viktigt är nästan uppenbart då konceptet derivata i någon bemärkelse bara är den matematiska konkretiseringen av konceptet förändringshastighet. För en funktion av en variabel, $y(x)$, är frågan om förändring relativt otvetydigt, då förändringar av y relativt förändringar av x inte kan undersökas på särskillt många sätt: x kan göras större eller mindre. Redan för en funktion av två variabler, $f(x, y)$, är frågan betydligt svårare i och med att små förändringar av argumentet kan ske på många sätt. Det här kapitlet ger en introduktion till konceptet derivata och differentierbarhet för funktioner av flera variabler.

4.1 Partiella derivator

Vi börjar med att definiera begreppet kontinuitet för funktioner av flera variabler.

Definition 4.1.1. Låt $f(x_1, \ldots, x_n)$ vara en funktion av n variabler definierad på $D_f \subset \mathbb{R}^n$ och låt $\mathbf{a} \in D_f$ vara en inre punkt. Vi säger att f är *kontinuerlig i punkten* $\mathbf{a}$ om

$$\lim_{|\mathbf{h}|\to 0} f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}).$$

Om D_f är en öppen mängd och f är kontinuerlig i alla punkter i D_f så säger vi att f är en *kontinuerlig funktion.* $\triangle$

Notera att kontinuitet i flera variabler är krångligare än kontinuitet i en variabel eftersom det helt enkelt finns fler sätt som $\mathbf{h}$ kan gå mot 0 på.

Betrakta till exempel funktionen

$$f(x, y) = \begin{cases} \frac{x^2}{x^2+y^2} & \text{om } x, y \neq 0 \\ 0 & \text{om } x = y = 0. \end{cases}$$

Om vi nu betraktar $f(0 + \mathbf{h})$ där $\mathbf{h} = (h, 0)$ får vi

$$\lim_{|\mathbf{h}|\to 0} f(0 + \mathbf{h}) = \lim_{h\to 0} \frac{h^2}{h^2} = 1.$$

Om vi istället sätter $\mathbf{h} = (0, h)$ får vi

$$\lim_{|\mathbf{h}|\to 0} f(0 + \mathbf{h}) = \lim_{h\to 0} \frac{0^2}{h^2} = 0.$$

Och om vi sätter $\mathbf{h} = (h, h)$ får vi

$$\lim_{|\mathbf{h}|\to 0} f(0 + \mathbf{h}) = \lim_{h\to 0} \frac{h^2}{h^2 + h^2} = \frac{1}{2}.$$

Som vi ser kan vi få helt olika gränsvärden beroende på hur vi närmar oss 0, och som det sista exemplet visar räcker det inte att bara undersöka koordinataxlarna!

Låt nu $f(x_1, \ldots, x_n)$ vara en funktion av n variabler. Det är naturligt att fråga sig hur funktionens värde förändras när *en* variabel förändras, men de andra variablerna hålls konstanta.

Om till exempel $T(x, y, z)$ är en funktion som beskriver temperaturen på olika platser i ett rum, och x, y, z då alltså är koordinater som beskriver positionen i rummet, så är det naturligt att fråga sig hur temperaturen förändras om vi, till exempel, rör oss rakt uppåt från en punkt.

Detta leder oss till definitionen av en *partiell derivata*.

Definition 4.1.2. Låt $f(x_1, \ldots, x_n)$ vara en funktion definierad på $D_f \subset \mathbb{R}^n$ och låt $(a_1, \ldots, a_n)$ vara en *inre punkt* i D_f. Om gränsvärdet

$$\lim_{h \to 0} \frac{f(a_1, \ldots, a_{j-1}, a_j + h, a_{j+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}$$

existerar i punkten $(a_1, \ldots, a_n)$ säger vi att f är *partiellt deriverbar* med avseende på x_j i punkten $(a_1, \ldots, a_n)$. Själva gränsvärdet kallar vi för den *partiella derivatan* av f med avseende på x_j, och denna kommer vi beteckna med

$$\frac{\partial f}{\partial x_j}(a_1, \ldots, a_n), \quad \text{eller} \quad f_j'(a_1, \ldots, a_n).$$

Om alla de partiella derivatorna existerar i punkten $(a_1, \ldots, a_n)$ säger vi att f är partiellt deriverbar i punkten $(a_1, \ldots, a_n)$. Vidare säger vi att en funktion som är partiellt deriverbar i alla punkter i sin definitionsmängd helt enkelt är *partiellt deriverbar*. △

Det är värt att notera att om en funktion är partiellt deriverbar måste definitionsmängden vara öppen. Detta på grund av att definitionen av en partiell derivata förutsätter att det för varje punkt $(a_1, \ldots, a_n) \in D_f$ finns något tillräckligt litet $\delta > 0$ så att

$$|h| < \delta \Rightarrow (a_1, \ldots, a_{j-1}, a_j + h, a_{j+1}, \ldots, a_n) \in D_f.$$

Om detta inte är fallet kommer ju differenskvoten

$$\frac{f(a_1, \ldots, a_{j-1}, a_j + h, a_{j+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}$$

från definitionen av den partiella derivatan inte vara väldefinierad!

Detta är skälet till att vi framöver nästan alltid kommer kräva att en punkt $\mathbf{a}$ i vilken vi ska undersöka någon slags derivata är en *inre* punkt.

Exempel 4.1.3. Låt $f(x_1, x_2) = x_1 x_2^2$. Från definitionen ser vi då att

$$\begin{aligned} f_1'(x_1, x_2) &= \lim_{h \to 0} \frac{f(x_1 + h, x_2) - f(x_1, x_2))}{h} \\ &= \lim_{h \to 0} \frac{(x_1 + h)x_2^2 - x_1 x_2^2}{h} \\ &= \lim_{h \to 0} \frac{hx_2^2}{h} \\ &= \lim_{h \to 0} x_2^2 = x_2^2, \end{aligned}$$

och

$$
\begin{aligned}
f_2'(x_1,x_2) &= \lim_{h\to 0} \frac{f(x_1,x_2+h)-f(x_1,x_2))}{h} \\
&= \lim_{h\to 0} \frac{x_1(x_2+h)^2 - x_1x_2^2}{h} \\
&= \lim_{h\to 0} \frac{x_1(x_2^2+2x_2h+h^2) - x_1x_2^2}{h} \\
&= \lim_{h\to 0} \frac{2x_1x_2h + x_1h^2}{h} \\
&= \lim_{h\to 0} 2x_1x_2 + x_1h \\
&= 2x_1x_2.
\end{aligned}
$$

▲

Att beräkna partiella derivator från definitionen är dock oftast onödigt krångligt. Eftersom partiella derivator definieras utifrån ”samma” differenskvot som används i definitionen av derivatan för en funktion av en variabel, kan vi istället tänka att alla variabler utom den vi deriverar med avseende på är konstanter. Därefter kan vi använda kända formler för derivator av elementära funktioner, samt derivationsregler för summor, produkter, kvoter och sammansättningar.

Exempel 4.1.4. Låt $f(x_1,x_2,x_3) = x_1\cos(x_1x_2x_3)$ och låt $D_f = \mathbb{R}^3$. Vi vill beräkna $f_1'(x_1,x_2,x_3)$ och $f_2'(x_1,x_2,x_3)$.
I det första fallet tänker vi att

$$f(x_1,x_2,x_3) = x_1\cos(cx_1),$$

där konstanten c (konstant med avseende på x_1 då alltså) är x_2x_3. Genom att använda produktregeln och kedjeregeln ser vi att derivatan av $x_1\cos(cx_1)$ är

$$\cos(cx_1) - x_1c\sin(cx_1).$$

Eftersom $c = x_2x_3$ blir alltså

$$f_1'(x_1,x_2,x_3) = \cos(x_1x_2x_3) - x_1x_2x_3\sin(x_1x_2x_3).$$

När vi beräknar den partiella derivatan med avseende på x_2 är det istället alla uttryck i x_1 och x_3 som är konstanta istället. Vi tänker då istället att

$$f(x_1,x_2,x_3) = c_1\cos(c_2x_2),$$

där $c_1 = x_1$ och $c_2 = x_1x_3$. Den partiella derivatan med avseende på x_2 blir då alltså

$$-c_1c_2\sin(c_2x_2).$$

Genom att byta ut c_1 mot x_1 och c_2 mot x_1x_3 ser vi att

$$f_2'(x_1,x_2,x_3) = -x_1^2x_3\sin(x_1x_2x_3).$$

I många situationer räcker det inte att veta att en funktion *har* partiella derivator, utan man måste även kunna säga att dessa partiella derivator beter sig bra på något sätt. Vad som avses med att de partiella derivatorna beter sig bra kan variera, men det vanligaste är att man vill att de ska vara kontinuerliga, eller till och med partiellt deriverbara. Alltså, om $f(x_1, \dots, x_n)$ är en partiellt deriverbar funktion kan man betrakta de partiella derivatorna f'_j, och om dessa funktioner är partiellt deriverbara kan man bilda

$$\frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right),$$

där $j, k = 1, \dots, n$. Dessa kallas andra ordningens partiella derivator och brukar betecknas

$$\frac{\partial^2 f}{\partial x_k \partial x_j}.$$

Om $j = k$, det vill säga att vi deriverar med avseende på samma variabel två gånger, så skriver vi

$$\frac{\partial^2 f}{\partial x_k^2}.$$

På motsvarande sätt betecknas högre ordningens partiella derivator. Ordningen bestäms av antalet derivator. Till exempel är

$$\frac{\partial^7 f}{\partial x_1^2 \partial x_2 \partial x_3 \partial x_4^3},$$

en partiell derivata av sjunde ordningen.

Vi kommer ofta hamna i situationer där vi vill veta att funktionerna vi arbetar med har partiella derivator upp till och med en viss ordning, samt att alla dessa partiella derivator är kontinuerliga.

Definition 4.1.5. Låt f vara kontinuerlig på en öppen mängd $D \subset \mathbb{R}^n$. Vi säger att f är av klass C^k, eller att f tillhör $C^k(D)$, om alla partiella derivator till och med ordning k existerar och är kontinuerliga på D.

Om f är definierad på en mängd $D \subset \mathbb{R}^n$ som inte nödvändigtvis är öppen så säger vi att f ligger i $C^k(D)$ om f är kontinuerlig på D och $f \in C^k(D^\circ)$, där D° betecknar det *inre* av D. $\triangle$

Anledningen att vi vill kräva att en funktion f ligger i till exempel C^2 istället för att bara kräva att f har partiella derivator upp till och med ordning 2 är teknisk, men skillnaden är stor. En viktig skillnad handlar om i vilken ordning vi tar de partiella derivatorna. Är till exempel

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_2}?$$

Om $f \in C^2$ är svaret alltid ja, men om vi bara vet att f *har* partiella derivator av ordning 2, men vi inte vet något mer om dessa, så kan svaret faktiskt vara nej.

När vi säger $f \in C^0$ menar vi att f är kontinuerlig. Det är värt att notera att $C^{k+1} \subset C^k$ och att om $f, g \in C^k$ så är även $f + g \in C^k$.

Exempel 4.1.6. Betrakta funktionerna x^8 och $1/x$ definierade på det öppna intervallet $(0,1)$. Båda dessa funktioner har derivator upp till och med ordning 2 (i själva verket är båda funktionerna oändligt många gånger differentierbara) och dessa derivator är kontinuerliga. Det följer alltså att båda funktionerna ligger i $C^2((0,1))$.

Funktionen x^8 kan utan vidare betraktas som en funktion på det slutna intervallet $[0,1]$, och denna funktion ligger i $C^2([0,1])$ eftersom x^8 är kontinuerlig på det slutna intervallet $[0,1]$ och x^8 är två gånger kontinuerligt deriverbar på det inre av $[0,1]$, det vill säga på $(0,1)$.

Funktionen $1/x$ är dock svårare att betrakta som funktion på $[0,1]$ eftersom det är oklart vad funktionen ska ha för värde i punkten $x=0$. Vi skulle exempelvis kunna betrakta funktionen

$$f(x) = \begin{cases} 0 & \text{om } x = 0 \\ 1/x & \text{om } x \in (0,1]. \end{cases}$$

Denna funktion ligger fortfarande i $C^2([0,1]^\circ) = C^2((0,1))$, men den ligger ändå inte i $C^2([0,1])$ eftersom funktionen inte är kontinuerlig på $[0,1]$.

▲

4.2 Differentierbarhet

För många resultat är den naturliga generaliseringen av påståendet att en envariabelfunktion är *deriverbar* inte att en flervariabelfunktion har partiella derivator, utan ett aningen hårdare villkor. Nämligen att funktionen är *differentierbar*.

Låt $f(x)$ vara en funktion av en variabel. En ekvivalent formulering av påståendet att f är deriverbar i en punkt a är att säga att för något värde A gäller det att funktionen

$$g(h) = \frac{f(a+h) - f(a)}{h} - A$$

går mot 0 då $h \to 0$. En ekvivalent formulering är att säga att

$$f(a+h) - f(a) = Ah + hg(h)$$

för något värde A, och där $g(h)$ har egenskapen att

$$\lim_{h \to 0} g(h) = 0.$$

Detta är den formulering av deriverbarhet som av tekniska skäl bäst lämpar sig för att generalisera konceptet deriverbarhet till funktioner av flera variabler.

Definition 4.2.1. Låt f vara en funktion definierad på $D_f \subset \mathbb{R}^n$ och låt $\mathbf{a}$ vara en inre punkt i D_f. Vi säger att f är *differentierbar i punkten* $\mathbf{a}$ om det finns konstanter $A_1, \ldots, A_n$ och en funktion $g(\mathbf{h})$ så att

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = A_1 h_1 + \ldots + A_n h_n + |\mathbf{h}| g(\mathbf{h})$$

och

$$\lim_{\mathbf{h}\to 0} g(\mathbf{h}) = 0.$$

Om f är differentierbar i varje punkt $\mathbf{a} \in D_f$ säger vi att f är *differentierbar.* $\triangle$

Notera att genom att låta $\mathbf{h} = h\mathbf{e}_j$, där $\mathbf{e}_j$ är enhetsvektorn som har en 1a på position j och 0 på alla andra positioner, i definitionen av differentierbarhet följer det direkt att konstanterna A_j måste överensstämma med de partiella derivatorna $f'_j(\mathbf{a})$. Att vara differentierbar i en punkt är alltså ett hårdare krav än att ha partiella derivator i den punkten.

Den intresserade läsaren kan verifiera att ovanstående definition faktiskt fungerar exakt likandant även om vi skulle göra (det märkliga) antagandet att $g(0)$ är något annat än 0. Men för enkelhets skull kommer vi alltid gör det naturliga antagandet att $g(0) = 0$, det vill säga att g är kontinuerlig i origo.

I praktiken visar man sällan att en funktion är differentierbar utifrån definitionen, utan man använder istället följande sats.

Sats 4.2.2. *Varje funktion f av klass C^1 är differentierbar.*

Vi kommer dock inte att bevisa denna sats då det är för tekniskt avancerat och tidskrävande för denna kurs. Däremot kommer ett exempel ges i uppgifterna som visar att en funktion kan ha partiella derivator i en punkt *utan* att vara differentierbar. Att anta något mer, till exempel att de partiella derivatorna är kontinuerliga, behövs alltså för att vi ska kunna garantera att funktionen är differentierbar.

Den subtila, men viktiga skillnaden mellan differentierbarhet och att vara partiellt deriverbar är att de partiella derivatorna bara ger information om beteendet längs koordinataxlarna, medan differentierbarhet ställer krav på alla möjliga sätt att närma sig en punkt. Till exempel måste en differentierbar funktion vara kontinuerlig, men detta är faktiskt inte sant för en funktion som bara är partiellt deriverbar! Ett exempel på detta kommer att ges i övningarna.

4.3 Kedjeregeln

En av de viktigaste derivationsreglerna för funktioner av en variabel är kedjeregeln, som beskriver hur man deriverar sammansatta funktioner, $f(g(x))$. Ibland används även notationen $f \circ g(x)$ för att beskriva den sammansatta funktionen $f(g(x))$. Kedjeregeln i en variabel säger att

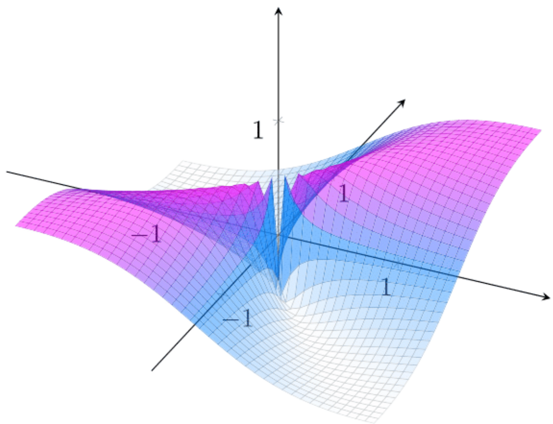
$$(f \circ g)'(x) = f'(g(x))g'(x),$$

eller

$$\frac{d}{dx} f(g(x)) = \frac{df}{dg}(g(x))\frac{dg}{dx}(x).$$

Denna regel är nödvändig för att till exempel beräkna

$$\frac{d}{dx}\sin(\cos(x)) = \cos(\cos(x))(-\sin(x)),$$



Figur 4.1: På bilden syns grafen för funktionen $xy/(x^2 + y^2)$. Om värdet för funktionen sätts till 0 i origo är funktionen partiellt deriverbar, men den är inte differentierbar i origo.

men även då vi beräknar partiella derivator, som till exempel

$$\frac{\partial}{\partial x}\sin(xy) = \cos(xy)y.$$

Funktionen $\sin(xy) : \mathbb{R}^2 \to \mathbb{R}$ är en sammansättning av $\sin(x) : \mathbb{R} \to \mathbb{R}$, och $xy : \mathbb{R}^2 \to \mathbb{R}$, men när vi tillämpar kedjeregeln för att beräkna den partiella derivatan erhåller vi en funktion från $\mathbb{R}$ till $\mathbb{R}$ genom att frysa den andra variabeln y genom att tänka på den som en konstant.

I den här kursen och i många andra sammanhang är man intresserad av att undersöka sammansättningar av typen

$$f \circ g : \mathbb{R}^m \to \mathbb{R}^n,$$

där $g : \mathbb{R}^m \to \mathbb{R}^k$ och $f : \mathbb{R}^k \to \mathbb{R}^n$, och $m, n, k \in \mathbb{N}$. För att undersöka dessa sammansättningar behövs en motsvarighet till kedjeregeln. En generell motsvarighet existerar, men är ganska teknisk, och i den här kursen behöver vi bara en motsvarighet till kedjeregeln för sammansättningar på formen

$$f \circ \mathbf{g} : \mathbb{R} \to \mathbb{R},$$

där $\mathbf{g} : \mathbb{R} \to \mathbb{R}^n$ och $f : \mathbb{R}^n \to \mathbb{R}$.

Det vill säga sammansättningar på formen

$$f(g_1(x), \dots, g_n(x)).$$

Vi begränsar oss därför till att bevisa följande sats.

Sats 4.3.1. *Låt $f(\mathbf{x}) = f(x_1, \dots, x_n)$ vara en differentierbar funktion av n variabler, och antag att funktionerna $g_1(t), \dots g_n(t)$ är deriverbara på det öppna intervallet (a, b). Då är den sammansatta funktionen $f(\mathbf{g}(t)) : (a, b) \to \mathbb{R}$ också deriverbar på intervallet (a, b) och*

$$\frac{d}{dt}f(\mathbf{g}(t)) = f_1'(\mathbf{g}(t))g_1'(t) + \dots + f_n'(\mathbf{g}(t))g_n'(t). \tag{4.1}$$

Anmärkning 4.3.2. Ofta arbetar man med funktionsbeteckningar som

$$f = f(\mathbf{x}) = f(x_1, \dots, x_n)$$

och

$$\mathbf{x} = \mathbf{x}(t),$$

och skriver bara

$$f = f(\mathbf{x}(t)) = f(x_1(t), \dots, x_n(t)).$$

Då skrivs ofta kedjeregeln mer kortfattat som

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}.$$

Med denna notation är det dock viktigt att komma ihåg att derivatorna av f, det vill säga $\partial f/\partial x_j$, evalueras i punkten $\mathbf{x}(t)$, medan x_j' evalueras i t.

Bevis. För att få enklare uttryck skriver vi bara ut beviset i fallet $n = 2$, det vill säga då f beror på två variabler. Beviset för generella n görs på samma sätt.

Enligt derivatans definition ska vi undersöka gränsvärdet

$$\lim_{k\to 0} \frac{f(\mathbf{g}(t+k)) - f(\mathbf{g}(t))}{k}.$$

Eftersom f enligt antagande är en differentierbar funktion ger definitionen av differentierbarhet att

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = f_1'(\mathbf{x})h_1 + f_2'(\mathbf{x})h_2 + |\mathbf{h}|\gamma(\mathbf{h}),$$

där $\lim_{\mathbf{h}\to 0} \gamma(\mathbf{h}) = 0$.

Genom att i ekvationen ovan sätta

$$\mathbf{x} = \mathbf{g}(t) \text{ och } \mathbf{h} = \mathbf{g}(t+k) - \mathbf{g}(t)$$

blir $\mathbf{x} + \mathbf{h} = \mathbf{g}(t+k)$, och efter division med k får vi

$$\frac{f(\mathbf{g}(t+k)) - f(\mathbf{g}(t))}{k} = \\ f_1'(\mathbf{g}(t))\frac{g_1(t+k) - g_1(t)}{k} + f_2'(\mathbf{g}(t))\frac{g_2(t+k) - g_2(t)}{k} + \frac{|\mathbf{h}|}{k}\gamma(\mathbf{h}). \quad (4.2)$$

Då vi låter $k \to 0$ i ekvationen ovan går de två första termerna enligt derivatans definition mot

$$f_1'(\mathbf{g}(t))g_1'(t) + f_2'(\mathbf{g}(t))g_2'(t),$$

vilket är resultatet vi slutligen vill få. Det återstår alltså bara att visa att den sista termen försvinner då $k \to 0$, alltså att

$$\lim_{k\to 0} \frac{|\mathbf{h}|}{k}\gamma(\mathbf{h}) = 0.$$

Eftersom g_1 och g_2 är deriverbara och

$$|\mathbf{h}| = \sqrt{h_1^2 + h_2^2} \leqslant |h_1| + |h_2| = |g_1(t+k) - g_1(t)| + |g_2(t+k) - g_2(t)|$$

följer det att $\lim_{k\to 0} |\mathbf{h}|/k$ är begränsat. Vidare är $\mathbf{g}$ kontinuerlig eftersom g_1 och g_2 är deriverbara, och därför har vi att

$$\mathbf{h} = \mathbf{g}(t+k) - \mathbf{g}(t) \to 0$$

då $k \to 0$. Detta innebär att

$$\lim_{k\to 0} \gamma(\mathbf{h}) = 0,$$

och eftersom vi ovan såg att $|\mathbf{h}|/k$ är begränsad då $k \to 0$ innebär detta att

$$\lim_{k\to 0} \frac{|\mathbf{h}|}{k}\gamma(\mathbf{h}) = 0,$$

vilket var det sista som behövde bevisas. □

Anmärkning 4.3.3. Vanligtvis kommer vi att använda kedjeregeln i situationer då vi kan anta att både f och funktionerna $g_1, \ldots, g_n$ är av klass C^1 (och inte bara differentierbara respektiva deriverbara, vilket är det som antas i satsformuleringen). Med dessa något starkare antaganden visar ekvation (4.1) att även sammansättningen $f \circ \mathbf{g}$ är av klass C_1, eftersom alla termer i högerledet i så fall är kontinuerliga.

Exempel 4.3.4. Låt $f : \mathbb{R}^2 \to \mathbb{R}$ ges av $f(x, y) = xy$ och låt $\mathbf{g} : \mathbb{R} \to \mathbb{R}^2$ ges av $\mathbf{g}(t) = (g_1(t), g_2(t)) = (\cos(t), \sin(t))$. Vi kan nu beräkna derivatan av sammansättningen $f \circ \mathbf{g} : \mathbb{R} \to \mathbb{R}$ på två sätt: Antingen genom att explicit uttrycka $f \circ \mathbf{g}$ som en funktion av en variabel, och därefter derivera denna, eller genom att använda ekvation (4.1).

Till att börja med är $f \circ \mathbf{g}(t) = (xy) \circ (\cos(t), \sin(t)) = \cos(t)\sin(t)$, och således är

$$\frac{d}{dt} f \circ \mathbf{g}(t) = -\sin(t)^2 + \cos(t)^2.$$

Om vi istället använder ekvation (4.1) ser vi att derivatan ges av

$$f_1'(\mathbf{g}(t))g_1'(t) + f_2'(\mathbf{g}(t))g_2'(t).$$

Genom att använda att $f_1' = y, f_2' = x, g_1'(t) = -\sin(t)$, och $g_2'(t) = \cos(t)$ ger ekvationen ovan också att

$$\frac{d}{dt} f \circ \mathbf{g}(t) = -\sin(t)^2 + \cos(t)^2.$$

▲

Kedjeregeln stora styrka ligger dock i att den hjälper oss att hitta uttryck för derivatorna av sammansättningar där inte alla funktioner är explicit kända.

Exempel 4.3.5. Låt $f : \mathbb{R}^3 \to \mathbb{R}$ ges av $f(x, y, z) = x^2 + y^2 + h(z)$ och låt $\mathbf{g} : \mathbb{R} \to \mathbb{R}^3$ ges av $\mathbf{g}(t) = (g_1(t), g_2(t), g_3(t))$.

Enligt kedjeregeln är

$$\frac{d}{dt} f(\mathbf{g}(t)) = f_1'(\mathbf{g}(t))g_1'(t) + f_2'(\mathbf{g}(t))g_2'(t) + f_3'(\mathbf{g}(t))g_3'(t).$$

Eftersom $f_1' = 2x$, $f_2' = 2y$, och $f_3' = h'(z)$ ges derivatan av

$$\frac{d}{dt} f(\mathbf{g}(t)) = 2g_1(t)g_1'(t) + 2g_2(t)g_2'(t) + h'(g_3(t))g_3'(t).$$

Om till exempel $h(z) = z^2$ och $g_1(t) = g_2(t) = g_3(t) = t^2$ ger insättning av derivatorna i ekvationen ovan att

$$\frac{d}{dt} f(\mathbf{g}(t)) = 12t^3,$$

vilket såklart även kan inses genom att explicit uttrycka sammansättningen $f \circ \mathbf{g}$ och derivera det erhållna uttrycket.

▲

Övningar

Övning 4.1. Beräkna de partiella derivatorna av $f(x, y) = \sin(\cos(xy))$.

Övning 4.2. Beräkna de partiella derivatorna av $f(x, y) = x^5y^3z + xyz$.

Övning 4.3 ($\star$). Använd definitionen av differentierbarhet för att visa att om en funktion f är differentierbar i en inre punkt $\mathbf{a} \in D_f$ så är f kontinuerlig i $\mathbf{a}$.

Övning 4.4 ($\star\star$). I den här uppgiften ska vi visa att det finns funktioner som är partiellt deriverbara, men som inte är differentierbara.

Betrakta funktionen

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{om } x, y \neq 0 \\ 0 & \text{om } x = y = 0. \end{cases}$$

(a) Visa att f är partiellt deriverbar i origo.

(b) Visa att f *inte* är kontinuerlig i origo.

(c) Använd (b) och resultatet av en tidigare övningsuppgift för att visa att f *inte* är differentierbar i origo.

Övning 4.5. Låt $f : \mathbb{R}^2 \to \mathbb{R}$ ges av $f(x, y) = \sin(x + y)$ och låt $\mathbf{g} : \mathbb{R} \to \mathbb{R}^2$ ges av $\mathbf{g}(t) = (g_1(t), g_2(t)) = (t^2, t^3)$.

Beräkna derivatan av sammansättningen $f \circ \mathbf{g} : \mathbb{R} \to \mathbb{R}$ på två sätt: genom att explicit uttrycka $f \circ \mathbf{g}$ som en funktion av en variabel, och därefter derivera denna; och genom att använda ekvation (4.1).

Övning 4.6. Låt $f : \mathbb{R}^2 \to \mathbb{R}$ ges av $f(x, y) = e^{x^2+y^2}$ och låt $\mathbf{g} : \mathbb{R} \to \mathbb{R}^2$ ges av $\mathbf{g}(t) = (g_1(t), g_2(t)) = (\sin(t), \cos(t))$.

Beräkna derivatan av sammansättningen $f \circ \mathbf{g} : \mathbb{R} \to \mathbb{R}$ på två sätt: genom att explicit uttrycka $f \circ \mathbf{g}$ som en funktion av en variabel, och därefter derivera denna; och genom att använda ekvation (4.1).

Övning 4.7. Låt $f : \mathbb{R}^3 \to \mathbb{R}$ ges av $f(x, y, z) = x + y^2 + z^3$ och låt $\mathbf{g} : \mathbb{R} \mapsto \mathbb{R}^3$ ges av $\mathbf{g}(t) = (g_1(t), g_2(t), g_3(t)) = (t^3, t^2, t)$.

Beräkna derivatan av sammansättningen $f \circ \mathbf{g} : \mathbb{R} \to \mathbb{R}$ på två sätt: genom att explicit uttrycka $f \circ \mathbf{g}$ som en funktion av en variabel, och därefter derivera denna; och genom att använda ekvation (4.1).

Övning 4.8. Beräkna alla partiella derivator upp till och med ordning 2 av funktionen $f(x, y) = x^2 + 2xy + y^2$.

Övning 4.9. Beräkna alla partiella derivator upp till och med ordning 2 av funktionen $f(x, y) = \sin(x^2 + y)$.

5 Optimering i Euklidiska rum

I väldigt många forskningsområden är man intresserad av att avgöra var en funktion eller operation antar sitt största respektive minsta värde, och målet i den här kursen är att studera en speciell typ av viktiga operationer, hitta ett villkor som måste uppfyllas för dessa operationers maximum och minimum, och använda detta villkor för att lösa ett svårt problem från klassisk mekanik. Men innan vi är redo att göra det behöver vi förstå motsvarande problem för vanliga reella funktioner av en och flera variabler. Vi kommer hitta ett villkor för när en funktion av en variabel antar sitt största respektive minsta värde, och därefter generaliserar vi detta villkor till funktioner av flera variabler. I kapitel 7 används idéerna från detta kapitel för att härleda ett motsvarande villkor för maximum och minimum av den speciella sorts operationer som nämndes ovan.

5.1 Optimering för funktioner av en variabel

Innan vi kan beskriva och bevisa villkor för när en funktion av en variabel har ett maximum eller minimum måste vi först rigoröst definiera vad det betyder att en funktion har ett maximum eller minimum. För att undvika repetition definierar vi maximum och minimum för funktioner av n variabler redan här. För att få de relevanta definitionerna för en funktion av en variabel är det bara att välja att $n = 1$.

Definition 5.1.1. Låt $f(x)$ vara en funktion definierad på mängden $D_f \subset \mathbb{R}^n$ och låt x_0 vara en punkt i D_f. Vi säger att f har ett *lokalt maximum* i x_0 om det finns ett tal $\delta > 0$ så att

$$x \in \{x \in D_f : |x - x_0| < \delta\} \Rightarrow f(x) \leqslant f(x_0).$$

Vi kallar då x_0 en *lokal maximipunkt* för f och funktionsvärdet $f(x_0)$ för ett *lokalt maximivärde.*

Om vi dessutom har att $f(x) < f(x_0)$ då $x \neq x_0$ säger vi istället att vi har en *sträng* lokal maximipunkt och ett *strängt* lokalt maximivärde.

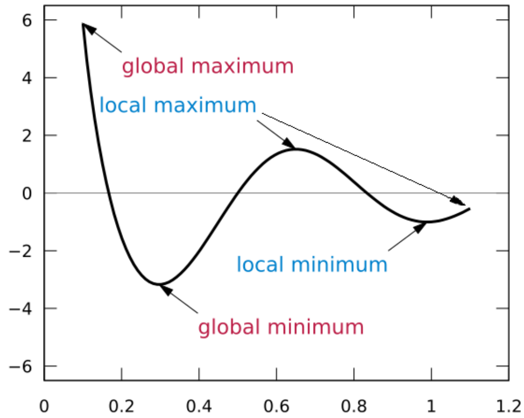
På motsvarande sätt (alltså med omvända olikhetstecken) definieras en (sträng) *lokal minimipunkt* och ett (strängt) *lokalt minimivärde.*

$\triangle$

Lokala maximipunkter och lokala minimipunkter kallas med ett gemensamt namn för *lokala extremvärden*, och vi säger att f har ett *lokalt extremvärde* i dessa punkter.

Det är viktigt att komma ihåg att huruvida en punkt x_0 är en lokal extrempunkt eller ej beror på såväl funktionen som på definitionsmängden. Låt till exempel $f(x) = x$. Om $D_f = [0, 1]$ är $x_0 = 1$ både en lokal och en global maximipunkt. Om $D_f = [0, 1] \cup [2, 3]$ är $x_0 = 1$ en lokal maximipunkt, men inte en global maximipunkt. Om $D_f = [0, 2]$ är $x_0 = 1$ inte ens en lokal maximipunkt.

Sats 5.1.2. *Om en funktion f av en variabel har ett lokalt extremvärde i en* inre *punkt $x_0 \in D_f \subset \mathbb{R}$ och om f är deriverbar i x_0 så är $f'(x_0) = 0$.*



Figur 5.1: Grafen för en funktion som har både lokala och globala extrempunkter. Bild från Wikipedia.

Bevis. Vi bevisar satsen i fallet då f har ett lokalt maximum i x_0. Beviset då f har ett lokalt minimum i x_0 är analogt.

Eftersom f har ett lokalt maximum vet vi att för något tillräckligt litet $\delta > 0$ gäller det att

$$|h| < \delta \Rightarrow f(x_0 + h) \leqslant f(x_0).$$

och att $x_0 + h \in D_f$ för alla $h \in [-\delta, \delta]$.

Från detta följer det att

$$\frac{f(x_0 + h) - f(x_0)}{h} \leqslant 0$$

för alla $0 < h < \delta$ eftersom täljaren är negativ och nämnaren är positiv.

Vidare har vi att

$$\frac{f(x_0 + h) - f(x_0)}{h} \geqslant 0$$

för alla $-\delta < h < 0$ eftersom både täljaren och nämnaren är negativa i det fallet.

Vi har alltså att

$$\lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h} \leqslant 0,$$

och

$$\lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h} \geqslant 0.$$

Eftersom vi antog att f är deriverbar i x_0 måste dessa gränsvärden överensstämma, och den enda möjligheten är då att $f'(x_0) = 0$. $\square$

Notera att det omvända påståendet *inte* gäller. Till exempel är $f'(0) = 0$ för $f(x) = x^3$ och $D_f = \mathbb{R}$, trots att $x_0 = 0$ inte är en lokal extrempunkt för $f(x) = x^3$. Satsen ger alltså ett *nödvändigt* men inte *tillräckligt* villkor för att en punkt x_0 ska vara en extrempunkt för en funktion f.

Notera även att vårt föregående exempel med den lokala extrempunkten $x_0 = 1$ för $f(x) = x$ och $D_f = [0, 1]$ visar att antagandet om att x_0 är en *inre* punkt i D_f är nödvändigt.

Vårt mål är att generalisera Sats (5.1.2) till funktioner som är definierade på $\mathbb{R}^n$, men för att kunna göra detta måste vi först förstå hur man definierar och undersöker förändringen av en flervariabelfunktion i en given *riktning.*

5.2 Gradienten

De partiella derivatorna ger var och en information om hur en funktion förändras längs en koordinatrikting, men det finns självklart mycket mer man kan vara intresserad av att undersöka. I vårt tidigare exempel där $T(x, y, z)$ beskriver temperaturen i ett rum, beskriver de partiella derivatorna hur temperaturen förändras om vi rör oss rakt framåt, rakt åt sidan, eller rakt uppåt. Men om man är intresserad av att veta hur temperaturen förändras när vi, till exempel, rör oss snett uppåt då? För att kunna svara på detta kommer vi sammanföra de partiella derivatorna till ett gemensamt begrepp.

Definition 5.2.1. För en partiellt deriverbar funktion $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ definierar vi *gradienten* av f i punkten $\mathbf{x}$ som vektorn

$$\operatorname{grad} f(\mathbf{x}) = \left(f_1'(\mathbf{x}), \ldots, f_n'(\mathbf{x})\right).$$

$\triangle$

Lägg märke till att $\operatorname{grad} f(\mathbf{x})$ är en vektor i $\mathbb{R}^n$, och att funktionen

$$\mathbf{x} \mapsto \operatorname{grad} f(\mathbf{x})$$

är en avbildning från $D_f \subset \mathbb{R}^n$ till $\mathbb{R}^n$. Denna funktion, som vi på vanligt sätt betecknar med $\operatorname{grad} f(\mathbf{x})$, är alltså ett n-dimensionellt *vektorfält*, som associerar en vektor i $\mathbb{R}^n$ till varje punkt i definitionsmängden till f.

Exempel 5.2.2. Låt $f(x, y, z) = xy^2z^3$. Derivering ger att

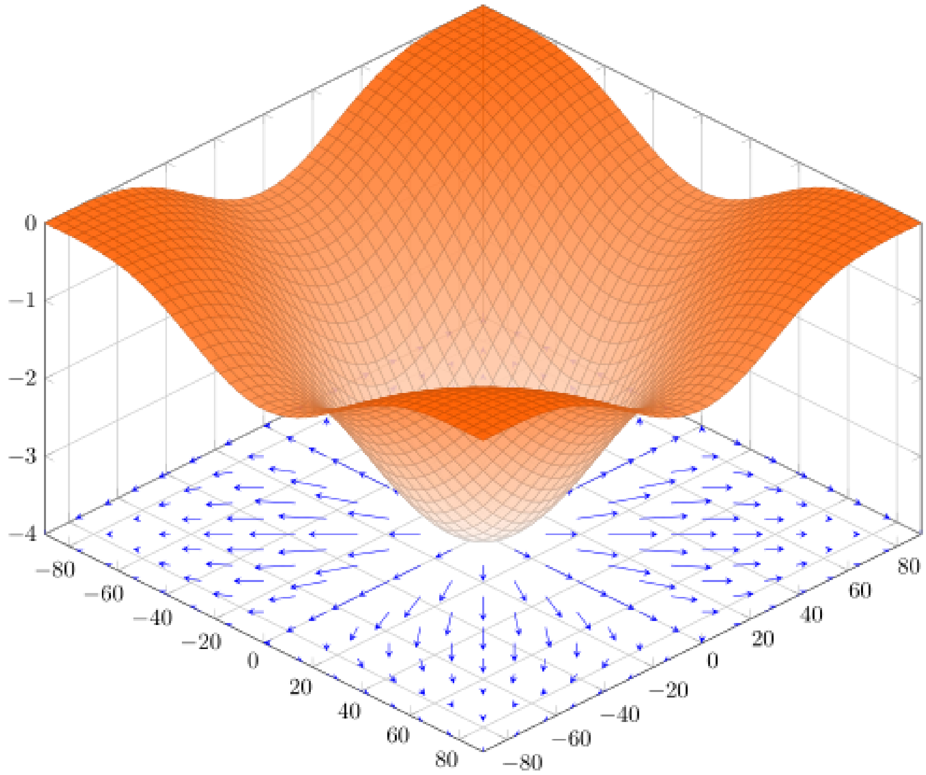
$$\frac{\partial f}{\partial x}(x, y, z) = y^2z^3, \quad \frac{\partial f}{\partial y}(x, y, z) = 2xyz^3, \quad \frac{\partial f}{\partial z}(x, y, z) = 3xy^2z^2.$$

Gradienten ges alltså av

$$\begin{aligned}\operatorname{grad} f(x, y, z) &= \left(\frac{\partial f}{\partial x}(x, y, z), \frac{\partial f}{\partial y}(x, y, z), \frac{\partial f}{\partial z}(x, y, z)\right)\\ &= (y^2z^3, 2xyz^3, 3xy^2z^2).\end{aligned}$$

Det här är alltså en avbildning från $\mathbb{R}^3$ till $\mathbb{R}^3$. Till exempel är

$$\operatorname{grad} f(1, 1, 1) = (1, 2, 3).$$



Figur 5.2: Gradienten för funktionen $f(x, y) = -(\cos(x)^2 + \cos(y)^2)^2$ utmarkerat som ett vektorfält i xy-planet under grafen för funktionen. Bild från Wikipedia.

5.3 Riktningsderivata

Vi är nu redo att undersöka tillväxten av en funktion f i en godtycklig inre punkt $\mathbf{a} \in D_f$ längs godtyckliga räta linjer $\mathbf{x} = \mathbf{a} + t\mathbf{v}$ som går genom punkten $\mathbf{a}$.

Här är $t \geqslant 0$ och $\mathbf{v} = (v_1, \ldots v_n)$ en *normerad* riktningsvektor, vilket innebär att

$$|\mathbf{v}| = \sqrt{v_1^2 + \ldots + v_n^2} = 1.$$

Detta innebär att t mäter avståndet från punkten $\mathbf{a}$ till punkten $\mathbf{a}+t\mathbf{v}$, eftersom

$$|(\mathbf{a} + t\mathbf{v}) - \mathbf{a}| = |t\mathbf{v}| = t|\mathbf{v}| = t.$$

Definition 5.3.1. Låt $\mathbf{v}$ vara en normerad riktningsvektor. Om gränsvärdet existerar kallar vi

$$f'_{\mathbf{v}}(\mathbf{a}) = \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t}$$

för *derivatan* av $f(x)$ i punkten $\mathbf{a}$ med avseende på riktiningen $\mathbf{v}$.

△

Om $\mathbf{v} = \mathbf{e}_j$ så får vi bara tillbaka den partiella derivatan $f'_j(\mathbf{a})$. Notera även att

$$f'_{-\mathbf{v}}(\mathbf{a}) = -f'_{\mathbf{v}}(\mathbf{a}). \tag{5.1}$$

I praktiken beräknar man sällan riktingsderivator från definitionen, utan istället använder man följande sats.

Sats 5.3.2. *Om f är en differentierbar funktion och $\mathbf{v}$ är en normerad riktningsvektor så är*

$$f'_{\mathbf{v}}(\mathbf{a}) = (\operatorname{grad} f(\mathbf{a})) \cdot \mathbf{v}.$$

Bevis. Vi vill visa att

$$f'_{\mathbf{v}}(\mathbf{a}) = \lim_{t\to 0} \frac{f(\mathbf{a}+t\mathbf{v}) - f(\mathbf{a})}{t} = \sum_{j=1}^{n} f'_j(\mathbf{a})v_j.$$

Eftersom f är differentierbar gäller det enligt definitionen att

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = |\mathbf{h}|g(\mathbf{h}) + \sum_{j=1}^{n} f'_j(\mathbf{x})h_j \tag{5.2}$$

där $g(\mathbf{h}) \to 0$ då $\mathbf{h} \to 0$.

Genom att sätta $\mathbf{x} = \mathbf{a}$ och $\mathbf{h} = t\mathbf{v} = (tv_1, \ldots tv_n)$ i (5.2) och dividera med t får vi

$$\frac{f(\mathbf{a}+t\mathbf{v}) - f(\mathbf{a})}{t} = \frac{|t\mathbf{v}|g(t\mathbf{v}) + \sum_{j=1}^{n} f'_j(\mathbf{a})tv_j}{t} = \pm|\mathbf{v}|g(t\mathbf{v}) + \sum_{j=1}^{n} f'_j(\mathbf{a})v_j,$$

där tecknet framför $|\mathbf{v}|g(t\mathbf{v})$ beror på om t är positivt eller negativt.

Eftersom $|\mathbf{v}| = 1$ och $\lim_{t\to 0} g(t\mathbf{v}) = 0$ är

$$\lim_{t\to 0} \frac{f(\mathbf{a}+t\mathbf{v}) - f(\mathbf{a})}{t} = \lim_{t\to 0} \pm|\mathbf{v}|g(t\mathbf{v}) + \sum_{j=1}^{n} f'_j(\mathbf{a})v_j = \sum_{j=1}^{n} f'_j(\mathbf{a})v_j. \quad \square$$

Exempel 5.3.3. Vi vill bestämma derivatan av $f(x, y) = x^2y^3$ i punkten (1,1), längs riktningen $\mathbf{v} = (1/\sqrt{2}, 1/\sqrt{2})$.

Notera till att börja med att

$$|\mathbf{v}| = \sqrt{\frac{1}{2} + \frac{1}{2}} = \sqrt{1} = 1,$$

så v är en normerad riktningsvektor.

Vi har att

$$\operatorname{grad} f(x, y) = (2xy^3, 3x^2y^2),$$

så

$$\operatorname{grad} f(1, 1) = (2, 3).$$

Det följer att

$$f'_{\mathbf{v}}(1, 1) = (2, 3) \cdot (1/\sqrt{2}, 1/\sqrt{2}) = \frac{2}{\sqrt{2}} + \frac{3}{\sqrt{2}} = \frac{5}{\sqrt{2}},$$

vilket ger den önskade riktingsderivatan. ▲

Sats 5.3.2 är viktig dels för att den ger oss en enkel metod för att beräkna riktningsderivator — vi tar bara skalärprodukten mellan den normerade riktningsvektorn och gradienten — men även för att den ger oss ett verktyg för att härleda viktiga teoretiska resultat. Följande sats är ett bra exempel på detta.

Sats 5.3.4. *Gradienten* $\operatorname{grad} f(\mathbf{a})$ *pekar i den riktning i vilken funktionen* f *förändras snabbast i punkten* $\mathbf{a}$. *Vidare ges mätetalet på den maximala förändringshastigheten av* $|\operatorname{grad} f(\mathbf{a})|$.

Bevis. Påståendet är en direkt konsekvens av Sats 5.3.2 och Cauchy–Schwarz' olikhet Sats 2.2.1 eftersom

$$|f'_{\mathbf{v}}(\mathbf{a})| = |\operatorname{grad} f(\mathbf{a}) \cdot \mathbf{v}| \leqslant |\operatorname{grad} f(\mathbf{a})||\mathbf{v}| = |\operatorname{grad} f(\mathbf{a})|.$$

Vidare säger Cauchy–Schwarz' olikhet också att likhet i olikheten ovan inträffar om och endast om de två vektorerna är parallella, det vill säga då

$$\mathbf{v} = \frac{1}{|\operatorname{grad} f(\mathbf{a})|} \operatorname{grad} f(\mathbf{a}).$$

Det här innebär precis att den maximala tillväxten erhålls i gradientens riktning, och att den maximala tillväxten ges av $|\operatorname{grad} f(\mathbf{a})|$. □

5.4 Optimering för funktioner av flera variabler

Kom ihåg definitionen för lokala extrempunkter som gavs i definition 5.1.1. I flera variabler gäller följande generalisering av sats 5.1.2.

Sats 5.4.1. *Om funktionen* f *har ett lokalt extremvärde i en* inre *punkt* $\mathbf{a}$ *i definitionsmängden* $D_f \subset \mathbb{R}^n$ *till* f *och om* f *är partiellt deriverbar i* $\mathbf{a}$ *så är*

$$f'_j(\mathbf{a}) = 0, \quad j = 1, \ldots, n.$$

Detta kan ekvivalent skrivas som att

$$\operatorname{grad} f(\mathbf{a}) = 0.$$

Bevis. Detta följer av motsvarande resultat i en variabel — det vill säga Sats 5.1.2 — genom att betrakta restriktionerna av f till de räta linjer genom $\mathbf{a}$ som är parallella med koordinataxlarna.

Eftersom f har ett lokalt extremvärde i $\mathbf{a}$ måste även envariabelfunktionen

$$x_j \mapsto f(a_1, \ldots, x_j, \ldots, a_n)$$

ha ett lokalt extremvärde för $x_j = a_j$, och därmed är dess derivata 0 för $x_j = a_j$. Men detta betyder exakt att

$$f'_j(\mathbf{a}) = 0$$

för $j = 1, \ldots, n$. □

Övningar

Övning 5.1. Beräkna gradienten av $f(x,y) = e^{x^2y}$.

Övning 5.2. Beräkna gradienten av $f(x,y,z) = x^3 + xy + y^3$.

Övning 5.3. Använd definitionen av globalt minimum för att visa att

$$f(x,y,z) = x^4 + y^2 + z^2$$

har ett lokalt minimum i origo.

Övning 5.4. Beräkna riktningsderivatan i riktningen $(1,1,1)/\sqrt{3}$ av funktionen $f(x,y,z) = x^2 + y^3 + z^2x$.

Övning 5.5. Beräkna riktningsderivatan i riktningen $v = (1,2)$ av funktionen $f(x,y) = x^2y$.

Övning 5.6 ($\star$). Låt $f(x,y) = x^4 - y^4$. Visa att $\mathrm{grad}(f)(x,y) = 0$ för $(x,y) = 0$, men visa att f varken har ett lokalt maximum eller minimum i origo.

Detta visar att villkoret att gradienten är 0 i en punkt är ett *nödvändigt* villkor för att punkten ska vara en lokal extrempunkt, men *inte* ett *tillräckligt* villkor.

Övning 5.7. Beräkna gradienten av $f(x,y) = e^{x\cos(y)}$ och visa att f inte har ett lokalt maximum i origo.

Övning 5.8. Betrakta funktionen $\sin(xyz^2)$. I vilken riktning förändras funktionen snabbast i punkten $(0,1,2)$? Vad är mätetalet på den maximala förändringshastigheten?

Övning 5.9. Betrakta funktionen x^2y^2z. I vilken riktning förändras funktionen snabbast i punkten $(-1,-1,3)$? Vad är mätetalet på den maximala förändringshastigheten?

Övning 5.10. Låt $T(x,y,z) = z/(x^2+y^2)$ beskriva temperaturen i ett rum som bland annat innehåller punkten $(1,1,1)$. I vilken riktning ska man röra sig från punkten $(1,1,1)$ för att temperaturen ska förändras så snabbt som möjligt? Hur snabbt förändras temperaturen i den riktningen?

6 Funktionaler

I detta kapitel ska vi formulera och definera de nödvändiga delarna som krävs för att formulera det så kallade *Brachistochronproblemet*.

För att formulera Brachistochronproblemet behöver vi en bättre förståelse för konceptet av en funktional. Ordet funktional kan ha olika betydelse för olika personer. Ofta syftar det på en funktion vars definitionsmängd är en mängd av funktioner såsom $C^1([a,b])$ (kom ihåg definition 4.1.5). I denna kurs definierar vi dem på följande vis,

Definition 6.0.1. Låt $\mathcal{F}$ vara en mängd funktioner. En *funktional* på $\mathcal{F}$ är en funktion

$$I : \mathcal{F} \to \mathbb{R}.$$

△

Exempel 6.0.2. Följande är ett exempel på en funktional på $C^2([0,2])$,

$$I : C^2([0,2]) \to \mathbb{R}$$
$$f \mapsto \sqrt{f(0) \cdot f'(1) \cdot f''\left(\frac{3}{2}\right)}$$

Till exempel så kan vi beräkna

$$I(e^x) = \sqrt{1 \cdot e^1 \cdot e^{\frac{3}{2}}} = e^{\frac{5}{4}}$$
$$I(3 + x^3) = \sqrt{(3 + 0^3) \cdot (3 \cdot 1^2) \cdot \left(6 \cdot \frac{3}{2}\right)} = \sqrt{3 \cdot 3 \cdot 9} = 9.$$

Exempel 6.0.3. Följande är ett exempel på en funktional på $C^0([0,1])$.

$$I : C^0([0,1]) \to \mathbb{R}$$
$$f \mapsto \int_0^1 f dx.$$

Till exempel så kan vi beräkna

$$I(x^2) = \int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$$
$$I\left((2x-1)^3\right) = \int_0^1 (2x-1)^3 dx = \left[\frac{(2x-1)^4}{8}\right]_0^1 = 0.$$

Exempel 6.0.4. Följande är ett exempel på en funktional på $C^1([0,1])$,

$$I : C^1([0,1]) \to \mathbb{R}$$
$$I(f) = \int_0^1 x f(x) f'(x) dx.$$

Till exempel så kan vi beräkna

$$I(x^3) = \int_0^1 x \cdot x^3 \cdot (3x^2) dx = 3 \int_0^1 x^6 dx = \frac{3}{7}.$$

Exempel 6.0.5. Notera att $\int_0^1 f dx$ **inte** är en funktional på $C^0(0,1)$. Till exempel så är den odefinerad på $f(x) = 1/x$. Den nyfikna läsaren kan fundera på varför den är väldefinerad för $g(x) = 1/x^2$ trots att g också blir oändligt stor i punkten $x = 0$.

Exempel 6.0.6. Låt $g : \mathbb{R} \to \mathbb{R}$ och I vara en funktional på en familj av funktioner $\mathcal{F}$. Då är

$$\begin{aligned} g \circ I : \mathcal{F} &\to \mathbb{R} \\ f &\mapsto g(I(f)) \end{aligned}$$

också en funktional. ▲

Vi ska nu studera ett par betydligt mer komplicerade funktionaler en de exempel vi sett tidigare.

6.1 Kurvlängds-funktionalen

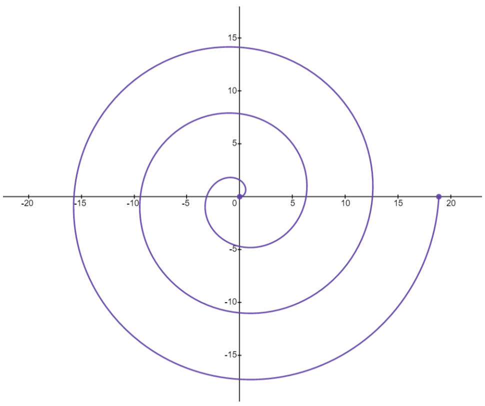
Föreställ dig en partikel i rörelse, exempelvis en elektron som flyger runt i ett elektriskt fält. Vi kan då föreställa oss att den sitter i ett koordinatsystem och att vi under T sekunder spelar in och skriver ned var partikeln befinner sig i koordinatsystemet för varje tidpunkt t. En naturlig fråga att ställa sig i denna situation är:

Hur lång är den bana/kurva som partikeln rört sig längs under tiden vi observerat den?

För att svara på denna fråga måste vi formulera om den mer matematiskt och beskriva dess rörelse med en *slät kurva*.

Definition 6.1.1. En *slät kurva* i $\mathbb{R}^n$ är en funktion $s : [0, T] \to \mathbb{R}^n$ vars komponenter är $C^2([0, T])$. Vi betecknar mängden släta kurvor i $\mathbb{R}^n$, definierade på $[0, T]$, för $\mathcal{S}_T^n$. △

Exempel 6.1.2. Här visar vi en bild på kurvan $s(t) = (t \cdot \cos(t), t \cdot \sin(t)) \in \mathcal{S}_{6\pi}^2$. ▲



Figur 6.1: Värdemängden för kurvan $(t \cdot \cos(t), t \cdot \sin(t))$. Bild gjord i Desmos.

Intuitionen bakom denna definition är att för en kurva i $\mathbb{R}^3$ så beskriver vektorn $s(t) = (x(t), y(t), z(t))$ var partikeln befinner sig i vårt koordinatsystem vid tidpunkten t i termer av dess x, y, z koordinater som var och en är en $C^2([0,T])$ funktion. Exempelvis säger $s(0)$ var partikeln befann sig när vi började observera den och $s(T)$ var den var när vi slutade observera den. Att s är kontinuerlig garanterar att partikeln inte 'teleporterar' sig från en plats till en annan. Att s är två gånger kontinuerligt differentierbar motsvarar att hastigheten och accelerationen också måste förändras kontinuerligt. För att svara på frågan i början av detta avsnitt måste vi förstå hur vi ska beräkna hastigheten av en partikel i varje tidpunkt t.

Definition 6.1.3. Givet $s \in \mathcal{S}_T^n$ så definerar vi *hastigheten* av s som $s'(t)$ och accelerationen som $s''(t)$.

Vi definerar *farten* av s som $v(t) = |s'(t)|$. △

Kom ihåg att när vi deriverar en funktion $s : [0,T] \to \mathbb{R}^n$ gör vi det komponentvis, se följande exempel.

Exempel 6.1.4. Låt $s(t) = (\cos(t), \sin(t)) \in \mathcal{S}_{2\pi}^2$. Vi beräknar hastigheten, farten och accelerationen.

$$s'(t) = \frac{d}{dt}\begin{bmatrix}\cos(t)\\ \sin(t)\end{bmatrix} = \begin{bmatrix}\frac{d}{dt}\cos(t)\\ \frac{d}{dt}\sin(t)\end{bmatrix} = \begin{bmatrix}-\sin(t)\\ \cos(t)\end{bmatrix}$$

$$v(t) = |s'(t)| = \sqrt{(-\sin(t))^2 + (\cos(t))^2} = \sqrt{1} = 1$$

$$s''(t) = \frac{d}{dt}s'(t) = \frac{d}{dt}\begin{bmatrix}-\sin(t)\\ \cos(t)\end{bmatrix} = \begin{bmatrix}-\cos(t)\\ -\sin(t)\end{bmatrix} = -s(t).$$

Det är bra att fundera en stund på varför accelerationen här är riktad i riktningen $-s(t)$. ▲

Exempel 6.1.5. Låt $s(t) = (t, \sqrt{t}) \in \mathcal{S}_1^2$. Vi beräknar hastigheten, farten och accelerationen.

$$s'(t) = \frac{d}{dt}\begin{bmatrix}t\\ \sqrt{t}\end{bmatrix} = \begin{bmatrix}\frac{d}{dt}t\\ \frac{d}{dt}\sqrt{t}\end{bmatrix} = \begin{bmatrix}1\\ \frac{1}{2\sqrt{t}}\end{bmatrix}$$

$$v(t) = |s'(t)| = \sqrt{1 + \frac{1}{4t}}$$

$$s''(t) = \frac{d}{dt}s'(t) = \frac{d}{dt}\begin{bmatrix}1\\ \frac{1}{2\sqrt{t}}\end{bmatrix} = \begin{bmatrix}0\\ \frac{-1}{4\sqrt{t^3}}\end{bmatrix}.$$

▲

Exempel 6.1.6. Låt $s(t) = (t, t^2, \frac{2t^3}{3}) \in \mathcal{S}_{10}^3$. Vi beräknar hastigheten, farten och accelerationen.

$$s'(t) = \frac{d}{dt}\begin{bmatrix}t\\ t^2\\ \frac{2t^3}{3}\end{bmatrix} = \begin{bmatrix}\frac{d}{dt}t\\ \frac{d}{dt}t^2\\ \frac{d}{dt}\frac{2t^3}{3}\end{bmatrix} = \begin{bmatrix}1\\ 2t\\ 2t^2\end{bmatrix}$$

$$v(t) = |s'(t)| = \sqrt{1^2 + 4t^2 + 4t^4} = \sqrt{(1 + 2t^2)^2} = 1 + 2t^2.$$

$$s''(t) = \frac{d}{dt} s'(t) = \frac{d}{dt} \begin{bmatrix} 1 \\ 2t \\ 2t^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 4t \end{bmatrix}.$$

▲

I grundskolan lärde vi oss att för en partikel som rör sig med konstant hastighet gäller

$$\text{fart} \cdot \text{tid} = \text{sträcka}.$$

Vi kommer nu ge en mer generell variant av denna formel som fungerar i högre dimensioner och när hastigheten in te är konstant.

Hjälpsats 6.1.7. *Givet $s \in \mathcal{S}_T^n$ gäller*

$$s(T) - s(0) = \int_0^T s'(t) dt.$$

Notera här att integrationen sker komponentvis.

Bevis. Vi inser att för varje koordinat kan vi tillämpa analysens fundamentalsats (häftets sats 3.1.6). Vi gör beräkningen i $\mathbb{R}^3$ men den är likadan oavsett dimension.

$$s(T) - s(0) = \begin{bmatrix} x(T) - x(0) \\ y(T) - y(0) \\ z(T) - z(0) \end{bmatrix} = \begin{bmatrix} \int_0^T v_x(t)dt \\ \int_0^T v_y(t)dt \\ \int_0^T v_z(t)dt \end{bmatrix} = \int_0^T s'(t) dt.$$

□

Detta är en bra formel men det är inte exakt det som vi behöver, vi skall demonstrera vad som saknas med denna formel i följande exempel.

Exempel 6.1.8. Betrakta en partikel som rör sig i en cirkel i $\mathbb{R}^2$. Vi beskriver dess rörelse med den $s \in \mathcal{S}_{2\pi}^2$ som ges av

$$s(t) = (\cos(t), \sin(t)).$$

Partikeln kommer att röra sig ett helt varv kring en cirkel med radie 1, så kurvlängden förväntas vara 2π längdenheter. Däremot så startar och slutar partikeln i samma punkt. Vi veriferar detta med den formella beräkingen,

$$s(2\pi) - s(0) = \int_0^{2\pi} \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} dt = \begin{bmatrix} \int_0^{2\pi} -\sin(t)dt \\ \int_0^{2\pi} \cos(t)dt \end{bmatrix} = \begin{bmatrix} \cos(2\pi) - \cos(0) \\ \sin(2\pi) - \sin(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Det föregående exemplet motiverar oss att inse att längden av den kurva en partikel färdas längs med bara beror på *farten* av partikeln, alltså hur fort den rör sig, och inte riktningen på hastigheten.

Definition 6.1.9. *Kurvlängd* är den funktional på $\mathcal{S}_T^n$ som defineras enligt följande formel,

$$L(s) = \int_0^T |s'(t)|dt.$$

$\triangle$

Exempel 6.1.10. Studera igen kurvan från exempel 6.1.8. Vi beräknar nu dess formella kurvlängd med hjälp av vår definition.

$$L(s) = \int_0^{2\pi} \left| \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} \right| dt = \int_0^{2\pi} \sqrt{\sin(t)^2 + \cos(t)^2} dt = \int_0^{2\pi} 1 dt = 2\pi.$$

Detta är det svar som vi förväntat oss. Om partikeln rör sig med en konstant fart på 1 längdenhet per tidsenhet under 2π tidsenheter så bör den rimligtvis ha rört sig 2π längdenheter. Detta antyder att det är en bra definition av kurvlängd vi har gjort.

För vår egen skull och för att övertyga oss själva om att vi har hittat en bra definition av kurvlängd bör vi bevisa att två partiklar som rör sig längs samma bana fast med olika fart har samma kurvlängd. Formellt kan man sammafatta detta i följande hjälpsats.

Definition 6.1.11. En *omparameterisering* av intervallet $[0, T]$ till intervallet $[0, \tau]$ är en funktion $\gamma \in C^2([0, \tau])$ som uppfyller

$$\begin{cases} \gamma(0) = 0 \\ \gamma(\tau) = T \\ \gamma'(t) > 0 \end{cases}.$$

Om s är en slät kurva definerad på $[0, T]$ säger vi att $s \circ \gamma$ är en omparameterisering av s. $\triangle$

Exempel 6.1.12. Funktionen x^2 är en omparameterisering av $[0, 2\pi]$ till $[0, \sqrt{2\pi}]$ och på samma sätt blir $(\cos(t^2), \sin(t^2)) \in \mathcal{S}^2_{\sqrt{2\pi}}$ en omparameterisering av $(\cos(t), \sin(t)) \in \mathcal{S}^2_{2\pi}$. Poängen är att desssa två kurvor, bland annat, har samma värdemängd, i detta fall enhetscirkeln. $\blacktriangle$

Hjälpsats 6.1.13. *Antag att $s \in \mathcal{S}_T^n$ är en slät kurva och låt γ vara en omparameterisering. Då gäller*

$$L(s) = L(s \circ \gamma).$$

Bevis. Detta bevis är en direkt tillämpning av ett lämpligt variabelbyte vid integrering i en variabel (hjälpsats 3.2.3), kedjeregeln för derivata i en variabel och att för $\lambda > 0$, $|\lambda \mathbf{x}| = \lambda |\mathbf{x}|$. Vi går genom beviset i $\mathbb{R}^3$ men det fungerar exakt likadant i fler eller färre dimensioner.

$$L(s \circ \gamma) = \int_0^\tau |(s \circ \gamma)'(t)| dt = \int_0^\tau \left| \begin{bmatrix} \frac{d}{dt} x \circ \gamma(t) \\ \frac{d}{dt} y \circ \gamma(t) \\ \frac{d}{dt} z \circ \gamma(t) \end{bmatrix} \right| dt$$

$$= \int_0^\tau \left| \begin{bmatrix} x'(\gamma(t)) \cdot \gamma'(t) \\ y'(\gamma(t)) \cdot \gamma'(t) \\ z'(\gamma(t)) \cdot \gamma'(t) \end{bmatrix} \right| dt = \int_0^\tau \left| s'(\gamma(t)) \cdot \gamma'(t) \right| dt = \int_0^\tau \left| s'(\gamma(t)) \right| \cdot \gamma'(t) dt$$

$$= \left\{ \begin{matrix} u=\gamma(t) & du=\gamma'(t)dt \\ u(0)=0 & u(\tau)=T \end{matrix} \right\} = \int_0^T \left| s'(u) \right| du = L(s) \quad \square$$

6.2 Brachistochronproblemet

Brachistochronproblemet lyder som följer:

Låt en kula i planet rulla friktionsfritt utan luftmotstånd från A till B längs med en slät kurva med enbart gravitationen som pådrivande kraft. Vilken kurva s minimerar tiden det tar att färdas från A till B?

En partikel som befinner sig i jordens gravitationsfält och som inte förlorar någon energi kommer att ha en konstant summa av potentiell- och kinetisk energi. Notera att i Brachistochronproblemet som vi nu studerar kan vi anta att partikeln startar i origo och står helt still, så summan av energierna är 0. Matematiskt så beskriver vi detta som

$$\forall t \in [0, T] \qquad mv^2/2 - mgy = 0$$

där $g \sim 9.82 m/s^2$ är tyngdaccelerationen vid jordens yta och $y \geqslant 0$ beskriver hur långt nedanför startpunkten kulan befinner sig. Alltså blir tecknet på den potentiella energin kanske lite förvirrande men det kommer att göra vissa beräkningar enklare senare. Vi kan därför hitta ett uttryck för farten v i termer av y.

$$v = \sqrt{2gy}$$

Låt oss nu börja bygga upp en mer formell matematisk formulering av problemet genom att beskriva tiden det tar för en kula att rulla från A till B längs med en kurva s.

6.3 Färdtidsfunktionalen

Färdtidsfunktionalen $\mathcal{T}$ är en intressant funktional som ser väldigt enkel att räkna ut. Om $s \in \mathcal{S}_T^n$ är en slät kurva i $\mathbb{R}^2$ så är, per definition, tiden det tar för partikeln att färdas längs med kurvan från start till slut exakt T. Det svåra med Brachistochronproblemet är att ofta så är $s(t)$ okänd. Det enda vi vet är formen på banan, men inte hur snabbt vi färdas längs med den. Formen av banan kan beskrivas med en funktion $y(x)$. Därför måste vi formulera om problemet så vi för varje val av bana, skapar en slät kurva s som vi sedan kan beräkna färdtiden utav.

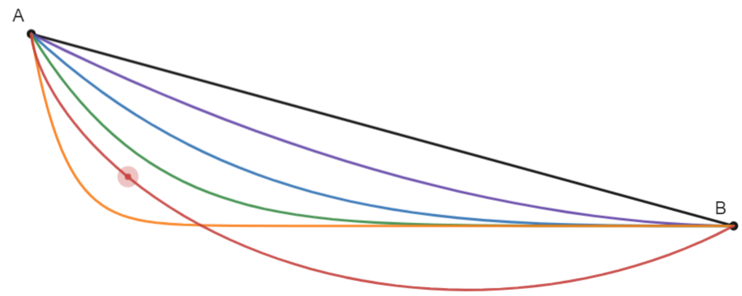
Definition 6.3.1. Färdtidsfunktionalen $\mathcal{T}$ är en funktional på mängden av alla släta kurvor i alla dimensioner, $\cup_{n\geqslant 1, T>0}\mathcal{S}_T^n$, som ges av

$$\mathcal{T} : \cup_{n\geqslant 1, T>0}\mathcal{S}_T^n \to \mathbb{R}$$
$$s \in \mathcal{S}_T^n \implies \mathcal{T}(s) = T.$$

△

Detta är en väldigt annorlunda funktional som beskriver de släta kurvornas definitionsmängd istället för något om deras värdemängd som det brukar vara i de tidigare exempel vi sett.

Vi går nu vidare och beskriver exakt hur det räcker att beskriva en bana $y(x)$ och hur den producerar en slät kurva när vi låter en kula rulla längs med den.



Figur 6.2: Exempel på olika funktioner y som beskriver en bana från A till B.

Definition 6.3.2. Ponera nu att vi har två punkter i planet, A och B. Vi kan införa ett koordinatsystem sådan att $A = (0, 0)$ och $B = (X, -Y)$ för två positiva tal $X, Y > 0$. Vi definerar en *lämplig funktion* y från A till B att vara en funktion av den horisontala förflyttningen x som uppfyller,

$$y \in C^2([0, X]) \quad y(0) = 0 \quad y(X) = Y$$
$$\lim_{x\to 0^+} \frac{dy}{dx}(0) > 0, \qquad x > 0 \implies y(x) > 0.$$

Denna kommer att beskriva en *'bana'* från A till B. △

Anmärkning 6.3.3. Antagandet att $\lim_{x\to 0^+} y'(0) > 0$ ska tolkas som att i starten av bana måste vi ha lite nedförsbacke för att kulan ska sättas i rullning. Vi kan även formellt ge ett ganska tekniskt bevis (som är överkurs) för att integralen $\int_0^x \sqrt{\frac{1+y'(u)^2}{2gy(u)}}du$ blir väldefinerad (konvergerar) och kontinuerlig även för $x = 0$ där själva integranden är oändligt stor. Idéen för beviset är att när vi har $\lim_{x\to 0^+} \frac{dy}{dx}(0) > 0$ kan vi för små värden på x kan skriva $c \leqslant y'(x)$ för något positivt tal c. Därmed får vi att $1 \leqslant (y'(x)/c)^2$ och

$$\int_0^x \sqrt{\frac{1+y'(u)^2}{y(u)}}du \leqslant \int_0^x \sqrt{\frac{(1+1/c^2)y'(u)^2}{y(u)}}du \leqslant \sqrt{(1+1/c^2)}\int_0^x \frac{y'(u)}{\sqrt{y(u)}}du =$$
$$=2\sqrt{(1+1/c^2)}\int_0^x \frac{d\sqrt{y(u)}}{du}du = 2\sqrt{(1+1/c^2)}\sqrt{y(x)} < \infty.$$

Antagandet $x > 0 \implies y(x) > 0$ har samma syfte, vi kommer aldrig fram om kulan åker upp till höjden $y = 0$ igen och stannar där. Mer formellt så handlar det villkoret också om att garantera att integralen konvergar.

Följande sats är teknisk men nödvändig för att vi med gått samvete skall kunna gå vidare med att lösa Brachistochronproblemet.

Sats 6.3.4. *Låt y vara en lämplig funktion från A till B. Då existerar ett tal T och en unik kurva $s(t) \in \mathcal{S}_T^2$ med egenskaperna:*

$$s(t) = \begin{bmatrix} x(t) \\ -y(x(t)) \end{bmatrix} \quad x(0) = 0, \quad x(T) = X, \quad \left\|s'(t)\right\| = \sqrt{2gy(x(t))}.$$

Vi kallar mängden av släta kurvor som uppstår på detta sätt för $S_{A,B}$.

Bevis. Givet y så måste s, om den existerar, per definition vara en omparameterisering av den släta kurvan $\begin{bmatrix} x \\ y(x) \end{bmatrix}$. Det vi måste göra är alltså att hitta en omparameterisering $x : [0,T] \to [0,X], \quad x'(t) > 0$. Vi utgår ifrån att s måste uppfylla

$$|s'(t)| = \sqrt{2gy(x(t))} \iff \sqrt{x'(t)^2 + (y'(x(t)) \cdot x'(t))^2} = \sqrt{2gy(x(t))}$$

$$\iff \sqrt{1+y'(x(t))^2} \cdot x'(t) = \sqrt{2gy(x(t))} \iff \sqrt{\frac{1+y'(x(t))^2}{2gy(x(t))}} \cdot x'(t) = 1. \tag{6.1}$$

Notera att när vi valt en specifik funktion y så blir detta en separabel differentialekvation där t är variabeln och $x(t)$ den okända funktionen. Oavsett y så kan vi med analysens huvudsats 3.1.6 konstruera en primitiv funktion till $\sqrt{\frac{1+y'(u)^2}{2gy(u)}}$.

$$F_y(x) = \int_0^x \sqrt{\frac{1+y'(u)^2}{2gy(u)}}du.$$

Enligt anmärkningen 6.3.3 så går F faktiskt att definera såhär och F_y blir en primitiv funktion till $\sqrt{\frac{1+y'(x)^2}{2gy(x)}}$ på $[0,X]$, trots att integranden inte uppfyller alla villkor i analysens fundamentalsats. Notera att med hjälp av kedjeregeln och ekvation 6.1 så ser vi att

$$\frac{d}{dt}F_y(x(t)) = F_y'(x(t)) \cdot x'(t) = \sqrt{\frac{1+y'(x(t))^2}{2gy(x(t))}} \cdot x'(t) = 1.$$

Vi kan då dra slutsatsen att

$$F_y(x(t)) = t$$

då $x(0) = 0$ måste vara sant och $F(0) = 0$. Vi inser också att $F_y(x)$ är strängt växande då den är väldefinerad för $x(t) \geqslant 0$ och dess derivata $\sqrt{\frac{1+y'(x)^2}{2gy(x)}} > 0$. Alla strängt växande funktioner är inverterbara enligt övningen 3.7. Alltså kommer vi ha en *unik* lösning till denna separabla differentialekvationen på formen $x(t) = F_y^{-1}(t)$ enligt hjälpsats 3.4.5. Vi kan alltså låta $T = F_y(X)$. Notera att $x(t)$ är två gånger deriverbar då

$$x'(t) = \sqrt{\frac{2gy(x(t))}{1+y'(x(t))^2}} \implies$$

$$x''(t) = \sqrt{\frac{1+y'(x(t))^2}{8gy(x(t))}} \frac{d}{dt}\left(\frac{2gy(x(t))}{1+y'(x(t))^2}\right) \implies$$

$$x''(t) = 2g\sqrt{\frac{1+y'(x(t))^2}{8gy(x(t))}} \frac{(1+y'(x(t))^2)y'(x(t)) - 2y''(x(t)) \cdot y'(x(t)) \cdot y(x(t))}{(1+y'(x(t))^2)^2} \cdot x'(t).$$

□

Vi kan alltså formulera Brachistochronproblemet som följande:
Hitta den kurva $s \in S_{A,B}$ som har den minsta färdtiden $\mathcal{T}(s)$.

Tanken är att $\mathcal{T}$ är exakt den tiden sådan att x tar sig hela vägen till $x(\mathcal{T}) = X$. Alltså skulle vi vilja skriva (utan hänsyn till några som helst detaljer)

$$\mathcal{T} = \int_0^{\mathcal{T}} 1dt = \int_0^X \frac{dt}{dx}dx = \int_0^X \frac{1}{dx/dt}dx = \int_0^X \frac{1}{x'}dx.$$

Vi skriver ner en hjälpsats som ger oss något rimligare att räkna på.

Hjälpsats 6.3.5. *För $s \in S_{A,B}$ gäller*

$$x'(t) = \sqrt{\frac{2gy(x(t))}{1+y'(x(t))^2}}.$$

Bevis. Eftersom $s \in S_{A,B}$ har vi att

$$\begin{aligned}
|s'(t)| &= \sqrt{2gy(x(t))} \iff \\
\left|\begin{bmatrix} x'(t) \\ y'(x(t)) \cdot x'(t) \end{bmatrix}\right| &= \sqrt{2gy(x(t))} \iff \\
x'(t) \cdot \left|\begin{bmatrix} 1 \\ y'(x(t)) \end{bmatrix}\right| &= \sqrt{2gy(x(t))} \iff \\
x'(t)\sqrt{1+y'(x(t))^2} &= \sqrt{2gy(x(t))} \iff \\
x'(t) &= \sqrt{\frac{2gy(x(t))}{1+y'(x(t))^2}}.
\end{aligned}$$

□

Sats 6.3.6. *Givet en lämplig funktion y från A till B, låt $s \in S_{A,B}$ vara den associerade släta kurvan. Då gäller det att*

$$\mathcal{T}(s) = \frac{1}{\sqrt{2g}} \int_0^X \sqrt{\frac{1+(y')^2}{y}} dx \tag{6.2}$$

Bevis. Detta är en direkt beräkning som använder variabelbyte

$$\begin{aligned}
\mathcal{T}(s) = \int_0^{\mathcal{T}(s)} 1dt = \int_0^{\mathcal{T}(s)} \sqrt{\frac{2gy(x(t))}{1+y'(x(t))^2}} \cdot \sqrt{\frac{1+y'(x(t))^2}{2gy(x(t))}} dt \\
\{\text{hjälpsats 6.3.5}\} = \int_0^{\mathcal{T}} \sqrt{\frac{1+y'(x(t))^2}{2gy(x(t))}} x'(t)dt \\
= \left\{ \begin{smallmatrix} x=x(t) & dx=x'(t)dt \\ x(0)=0 & x(\mathcal{T})=X \end{smallmatrix} \right\} = \int_0^X \sqrt{\frac{1+y'(x)^2}{2gy(x)}} dx.
\end{aligned}$$

□

Vi kan nu förenkla Brachistochronproblemet till att hitta den funktion $y(x)$ som minimerar

$$\mathcal{T}(y) = \frac{1}{\sqrt{2g}} \int_0^X \sqrt{\frac{1+(y')^2}{y}} dx \tag{6.3}$$

Övningar

Övning 6.1. Skissa kurvan $s(t) = (t - \sin(t), \cos(t) - 1)$ genom rita ut punkterna $s(t_k)$ för värdena $t_k = \frac{k\pi}{8}$ med heltalen $0 \leqslant k \leqslant 10$ och sedan för varje sådant $k < 10$ dra linjer $s(t_k) \to s(t_{k+1})$. Det är okej att använda miniräknare för att beräkna $s(t_k)$.

Övning 6.2. Skissa kurvan $s(t) = (\sin(t), \frac{1}{2}\sin(2t))$ genom rita ut punkterna s för värdena $t_k = \frac{k\pi}{6}$ med positiva heltala $k \leqslant 12$ och sedan för varje sådant k dra linjer $s(t_k) \to s(t_{k+1})$ samt $s(t_{12}) \to s(t_0)$. Det är okej att använda miniräknare för att beräkna $s(t_k)$.

Övning 6.3. Vilka av följande uttryck är funktionaler $I : C^2([-1,1]) \to \mathbb{R}$?

(a) $I(f) = \int_0^1 f dt$

(b) $I(f) = \int_0^1 f^2 dt$

(c) $I(f) = \int_0^x f' dt$

(d) $I(f) = \text{fisk}$

(e) $I(f) = e^{\left((\int_0^{\cos(x)} f(t)dt)'(0)\right)}$

(f) $I(f) = \int_0^1 \sqrt[3]{\frac{1+f'^2}{|f|+1}} dt$

(g) $I(f) = 1$.

(h) $I(f) = e^x$.

Övning 6.4. Låt $\cosh(x) = \frac{e^x+e^{-x}}{2}$ och $\sinh(x) = \frac{e^x-e^{-x}}{2}$. Bevisa att koordinaterna för kurvan $s(t) = (\cosh(t), \sinh(t)) \in \mathcal{S}_1^2$ uppfyller ekvationen $x^2-y^2 = 1$.

Övning 6.5 ($\star$). Beräkna kurvlängden av $s(t) = (e^t, \frac{e^{2t}}{2}) \in \mathcal{S}_1^2$. *Ledtråd: Använd övning 3.11.*

Övning 6.6 ($\star$). Hitta en slät kurva s som har farten

(a) $|s'(t)| = 7$

(b) $|s'(t)| = \sqrt{3}e^t$

(c) $|s'(t)| = f(t)$ för någon kontinuerlig funktion $f(t) \geqslant 0$ på $[0, T]$.

(d) $|s'(t)| = \sqrt{f(t)^2 + g(t)^2 + h(t)^2}$ för några kontinuerliga funktioner $f, g, h \geqslant 0$ på $[0, T]$.

Övning 6.7 ($\star$). Visa att varje kurva $s(t) = (x(t), y(t)) \in \mathcal{S}_T^2$ med egenskapen $x(0) = 0$, $x'(t) > 0$, är en omparameterisering av en slät kurva på formen

$$\tilde{s}(x) = (x, f_s(x)) \in \mathcal{S}_{x(T)}^2.$$

Följ dessa steg:

- Övertyga dig själv om att $x(t)$ är strängt växande och alltså inverterbar (använd övning 3.7).
- Verifiera att $x(t)$ är en omparameterisering av $[0, T]$ till $[0, x(T)]$.
- Låt $f_s = y \circ x^{-1}$. Visa att $s(t) = \tilde{s}(x(t))$ (du behöver inte visa att f_s är C^2 men det är den).

Övning 6.8 ($\star$). Studera kurvlängdsfunktionalen

$$L(s) := \int_0^T |s'(t)| dt.$$

Visa att för varje $s(t) = (x(t), y(t)) \in \mathcal{S}_T^2$ med egenskapen $x(0) = 0$, $x'(t) > 0$ så finns det en funktion f_s sådan att vi skriva

$$L(s) = \int_0^{x(T)} \sqrt{1 + f_s'(x)^2} dx.$$

Ledtråd: Kombinera övning 6.7 och hjälpsats 6.1.13

Övning 6.9 ($\star\star$). Motivera varför i princip vilken slät kurva som helst kan omparameteriseras så dess fart $v(t)$ bli vilken deriverbar funktion som helst.

7 Euler–Lagranges ekvation

I det här kapitlet ska vi härleda en metod för att hitta minimum, förutsatt att det existerar, till en speciell typ av funktional som är vanligt förekommande i många olika tillämpningar.

Låt

$$I : C^2([a,b]) \to \mathbb{R}$$

vara en funktional på formen

$$I : y(x) \mapsto \int_a^b f\big(x, y(x), y'(x)\big)\, dx, \tag{7.1}$$

där $f : \mathbb{R}^3 \to \mathbb{R}$ är en funktion i $C^2(\mathbb{R}^3)$, och argumentet y antas uppfylla något villkor på formen: $y(a) = y_1$ och $y(b) = y_2$, för några fixa reella tal y_1 och y_2.

Notera att enligt ekvation 6.3 från Sats 6.3.6 är färdtidsfunktionalen på formen (7.1).

Om till exempel $a = 0, b = 1$, och $f(x, y, z) = xyz$ så skulle funktionalen $I(y)$ ges av

$$I(y) = \int_0^1 xy(x)y'(x)dx.$$

För funktionen $y(x) = x$ ger funktionalen värdet

$$I(x) = \int_0^1 x \cdot x \cdot 1 = \int_0^1 x^2 dx = 1/3,$$

och för funktionen $y(x) = x^2$ ger funktionalen värdet

$$I(x^2) = \int_0^1 x \cdot x^2 \cdot 2x = 2\int_0^1 x^4 dx = 2/5.$$

Frågan vi är intresserade av är att avgöra vilken kurva $y(x)$ som minimerar/-maximerar I, förutsatt att en sådan kurva existerar.

Villkoren att $y \in C^2$ och $f \in C^2$ behövs av tekniska skäl för att härledningen som följer ska vara matematiskt korrekt. Vi kommer tyvärr inte kunna gå igenom alla detaljerade steg där dessa villkor krävs, men vi kommer att påpeka när villkoren används och vilka påståenden som inte nödvändigtvis är sanna om vi har svagare krav på y och f.

Vårt mål i det här kapitlet är att härleda ett resultat motsvarande Sats 5.4.1, men för funktionaler av formen ovan istället för funktioner från $\mathbb{R}^n$ till $\mathbb{R}$. Påståendet att gradienten måste försvinna i en inre extrempunkt kommer motsvaras av att en funktion $y(x)$ som är ett minimum eller maximum till I måste uppfylla en viss differentialekvation relaterad till I.

7.1 Härledning av Euler–Lagranges ekvation

Vårt mål är att bevisa följande.

Sats 7.1.1. *Låt*

$$I : C^2([a,b]) \to \mathbb{R}$$

vara en funktional på formen

$$I : y(x) \mapsto \int_a^b f\left(x, y(x), y'(x)\right) dx, \tag{7.2}$$

där $f : D_f \subset \mathbb{R}^3 \mapsto \mathbb{R}$ *är en funktion i* $C^2(D_f)$*. Om en kurva* $y(x) \in C^1[a,b]$ *maximerar/minimerar funktionalen* I *under det extra antagandet att* $y(a) = y_1$ *och* $y(b) = y_2$ *för några reella tal* y_1 *och* y_2*, så måste* $y(x)$ *uppfylla*

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0. \tag{7.3}$$

Idén är i stort sett densamma som för beviset av Sats 5.1.2, nämligen att om vi har ett max/min i en punkt, så kommer varje liten förändring — eller *variation* — i den punkten att resultera i ett mindre/större värde. För en vanlig, deriverbar funktion $f : \mathbb{R} \to \mathbb{R}$ innebar detta att derivatan i extrempunkten behövde vara 0, det vill säga i punkten x måste ekvationen $f'(x) = 0$ vara uppfylld. För funktionaler $I(y(x))$ på formen ovan kommer motsvarande observation resultera i slutsatsen att en extremal kurva $y(x)$ måste uppfylla differentialekvationen (7.3). Denna differentialekvation kallas för *Euler–Lagranges ekvation.*

Bevis. Vi bevisar satsen under antagandet att $y(x)$ är ett minimum till I. Beviset om $y(x)$ är ett maximum görs på samma sätt.

Låt $y(x)$ vara ett minimum till (7.2) och låt $\eta(x) \in C^2([a,b])$ vara en godtycklig funktion som uppfyller att $\eta(a) = \eta(b) = 0$. Om parametern α är ett litet reellt tal kommer

$$\tilde{y}_\alpha(x) := y(x) + \alpha\eta(x)$$

ge en familj av funktioner, eller *variationer*, parametriserade av α, som alla kan användas som argument för funktionalen I då alla ligger i C^2 och alla uppfyller randvillkoren: Eftersom $\eta(a) = \eta(b) = 0$ är $\tilde{y}_\alpha(a) = y(a) + \alpha\eta(a) = y_1$ och $\tilde{y}_\alpha(b) = y(b) + \alpha\eta(b) = y_2$ för alla α.

Notera att den vertikala skillnaden mellan $y(x)$ och $\tilde{y}_\alpha(x)$ är $\alpha\eta(x)$. Vi undersöker alltså funktionen som erhålls genom att ta funktionen $y(x)$, och därefter lägga på en perturbering $\alpha\eta(x)$. Vi kan tänka på detta som en perturbering i riktningen η och med magnituden α, precis som när vi betraktar vektorer i $\mathbb{R}^n$, $\mathbf{x} + \alpha\mathbf{v}$, när vi vill undersöka riktningsderivatan i punkten $\mathbf{x}$ i riktning $\mathbf{v}$.

Bevisidén är nu att beräkna derivatan med avseende på den reella variabeln α av funktionalen I applicerad på familjen av funktioner $y(x) + \alpha\eta(x)$. Man kan tänka på detta som att beräkna riktningsderivatan av I i riktningen η. Eftersom vi antar att I har ett minimum för funktionen y måste denna derivata vara 0, vilket vi kommer visa innebär att

$$0 = \int_a^b \eta(x)\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right) dx. \tag{7.4}$$

Eftersom ekvationen ovan måste hålla för *alla* val av funktionen η kan vi dra slutsatsen att

$$\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right) = 0,$$

vilket är det vi vill visa.

Vi fortsätter nu med härledningen av ekvation (7.4).

Låt η och $\tilde{y}_\alpha(x)$ tills vidare vara fixa, och substituera $y(x)$ och $y'(x)$ mot $\tilde{y}_\alpha(x)$ och $\tilde{y}'_\alpha(x) = y'(x) + \alpha\eta'(x)$ i (7.2). Detta ger oss

$$\begin{aligned} I_\eta(\alpha) := I(\tilde{y}_\alpha(x)) = \int_a^b f\left(x, \tilde{y}_\alpha(x), \tilde{y}'_\alpha(x)\right) dx \\ = \int_a^b f\left(x, y(x) + \alpha\eta(x), y'(x) + \alpha\eta'(x)\right) dx. \end{aligned}$$

Observera att funktionen $I_\eta(\alpha)$ är en vanlig funktion av ett reellt tal, nämligen parametern α.

När $\alpha = 0$ är $y(x) = \tilde{y}_\alpha(x)$, och eftersom vi antar att $y(x)$ minimerar I, så måste funktionen $I_\eta(\alpha)$ ha ett minimum för $\alpha = 0$. Från Sats 5.1.2 följer det därför att $I'_\eta(0) = 0$.

Eftersom vi antar att f, y och $\eta \in C^2$ kan vi derivera under integraltecknet, och alltså följer det att

$$I'_\eta(\alpha) = \frac{\partial}{\partial \alpha}\int_a^b f\left(x, \tilde{y}_\alpha(x), \tilde{y}'_\alpha(x)\right) dx = \int_a^b \frac{\partial}{\partial \alpha} f\left(x, \tilde{y}_\alpha(x), \tilde{y}'_\alpha(x)\right) dx. \tag{7.5}$$

Anmärkning 7.1.2. Det är värt att notera att ekvationer av typen (7.5) inte håller i allmänhet, utan vi behöver någon form av villkor på integranden — till exempel att den ligger i C^1 — för att derivatan av integralen ska överensstämma med integralen av derivatan. Att bevisa detta rigoröst är dock för tekniskt avancerat och tidskrävande för denna kurs.

Genom att använda kedjeregeln för funktioner av flera variabler, det vill säga Sats 4.3.1, ser vi att

$$\frac{\partial}{\partial \alpha} f\left(x, \tilde{y}_\alpha(x), \tilde{y}'_\alpha(x)\right) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial f}{\partial \tilde{y}_\alpha}\frac{\partial \tilde{y}_\alpha}{\partial \alpha} + \frac{\partial f}{\partial \tilde{y}'_\alpha}\frac{\partial \tilde{y}'_\alpha}{\partial \alpha} = \frac{\partial f}{\partial \tilde{y}_\alpha}\eta(x) + \frac{\partial f}{\partial \tilde{y}'_\alpha}\eta'(x), \tag{7.6}$$

där den sista ekvationen erhålls genom att notera att x inte beror på α, varav $(\partial x)/(\partial \alpha) = 0$, och att

$$\frac{\partial}{\partial \alpha}\tilde{y}_\alpha(x) = \frac{\partial}{\partial \alpha}\left(y(x) + \alpha\eta(x)\right) = \eta(x),$$

och

$$\frac{\partial}{\partial \alpha}\tilde{y}'_\alpha(x) = \frac{\partial}{\partial \alpha}\left(y'(x) + \alpha\eta'(x)\right) = \eta'(x).$$

Genom att substituera (7.6) i (7.5) ger detta att

$$I'_\eta(\alpha) = \int_a^b \frac{\partial f}{\partial \tilde{y}_\alpha}\eta(x) + \frac{\partial f}{\partial \tilde{y}'_\alpha}\eta'(x) dx. \tag{7.7}$$

Eftersom vi antar att $I'_\eta(\alpha) = 0$ då $\alpha = 0$ och vi alltid har att $\tilde{y}_\alpha = y$ då $\alpha = 0$, så följer det att

$$0 = I'_\eta(0) = \int_a^b \frac{\partial f}{\partial y}\eta(x) + \frac{\partial f}{\partial y'}\eta'(x)dx. \tag{7.8}$$

För att få ett uttryck som inte innehåller η' partialintegrerar vi den andra termen i integralen ovan. Eftersom $\eta(a) = \eta(b) = 0$ får vi att

$$\begin{aligned}\int_a^b \frac{\partial f}{\partial y'}\eta'(x)dx &= \left[\eta(x)\frac{\partial f}{\partial y'}\right]_a^b - \int_a^b \eta(x)\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)dx \\ &= -\int_a^b \eta(x)\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)dx.\end{aligned}$$

Tillsammans med ekvationen ovan innebär detta alltså att

$$0 = \int_a^b \eta(x)\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)dx. \tag{7.9}$$

Eftersom ekvationen ovan håller för *alla* funktioner $\eta(x) \in C^2([a,b])$ som uppfyller att $\eta(a) = \eta(b) = 0$ måste

$$\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right) = 0$$

överallt på intervallet $[a,b]$, vilket skulle bevisas. $\square$

Anmärkning 7.1.3. I slutet av beviset används påståendet att om $f(x)$ är kontinuerlig så gäller det att

$$\begin{aligned}\int_a^b \eta(x)f(x)dx = 0 \text{ för alla } \eta \in C^2([a,b]) \text{ som uppfyller att } \eta(a) = \eta(b) = 0 \\ \iff f(x) = 0,\end{aligned}$$

vilket inte är ett uppenbart påstående, utan det kräver egentligen ett separat bevis. Påståendet kallas ofta för *variationskalkylens huvudlemma* och kan till exempel bevisas genom ett motsägelsebevis. Idén är följande.

Om det finns någon punkt $x_0 \in [a,b]$ där $f(x_0) > 0$ (som vanligt hanteras situationen då $f(x_0) < 0$ på samma sätt) så innebär det att $f(x) > 0$ överallt i något litet intervall I_0 som innehåller x_0.

Vi kan nu konstruera en lämplig funktion η_0 som är positiv på något ännu mindre intervall, och 0 överallt utanför I_0.

För denna funktion η_0 kommer

$$\int_a^b f(x)\eta_0(x)dx > 0$$

på grund av det som har sagts ovan, men detta motsäger grundantagandet.

För att vara extra tydliga går vi ännu en gång igenom vad det innebär att en kurva $y(x)$ är en lösning till Euler–Lagranges ekvation. *Om* det finns någon kurva som minimerar eller maximerar funktionalen I så måste den kurvan uppfylla Euler–Lagranges ekvation. Men, bara för att vi har hittat en kurva som löser Euler–Lagranges ekvation så innebär inte detta att vi har hittat minimum eller maximum till I; det är inte säkert att det ens finns något minimum eller maximum.

Situationen är lik den för vanliga deriverbara funktioner $f : \mathbb{R} \to \mathbb{R}$. Om ett minimum eller maximum antas i en inre punkt x_0 måste x_0 vara en lösning till ekvationen $f'(x_0) = 0$, men den omvända implikationen gäller inte. Till exempel är derivatan till $f(x) := x^3$ noll i origo, men vi har inget maximum eller minimum eftersom funktionen är strängt växande.

Vidare kan det vara värt att betona att när vi, i Euler–Lagranges ekvation, deriverar f med avseende på x, y och y' så behandlas x, y och y' som *oberoende variabler*. I allmänhet är dock $\partial f/\partial y'$ en explicit funktion av x, som då därefter ska deriveras med avseende på x. Vi illustrerar detta med några exempel.

7.2 Exempel

Exempel 7.2.1. Låt återigen $a = 0$, $b = 1$, $f(x, y, z) = xyz$, och betrakta den associerade funktionalen

$$I : y(x) \mapsto \int_0^1 f(x, y, y')dx = \int_0^1 xy(x)y'(x)dx$$

Euler–Lagranges ekvation tillhörande funktionalen ovan ges av

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0,$$

där $f(x, y, y')$ återigen är xyy'.

Vi har att

$$\frac{\partial f}{\partial y} = \frac{\partial(xyy')}{\partial y} = xy',$$

och

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{d}{dx}\left(\frac{\partial(xyy')}{\partial y'}\right) = \frac{d}{dx}(xy) = y + xy',$$

så Euler-Lagrange ekvationen är alltså

$$(y + xy') - xy' = y = 0. \tag{7.10}$$

Detta innebär, till exempel, att en kurva $y(x)$ som minimerar I under det extra antagandet att $y(0) = y(1) = 1$ måste vara en lösning till differentialekvationen (7.10) som dessutom uppfyller randvillkoren $y(0) = y(1) = 1$.

Notera dock att det finns någon kurva som minimerar funktionalen under dessa villkor, då den enda funktionen som uppfyller (7.10) är funktionen $y(x) = 0$, och denna funktion uppfyller inte randvillkoren $y(0) = y(1) = 1$.

Exempel 7.2.2. Låt $a = 0$, $b = \pi$ och $f(x, y, z) = 2y\sin(x) - z^2$, och betrakta den associerade funktionalen

$$I : y(x) \mapsto \int_0^\pi f(x, y, y')dx = \int_0^\pi 2y(x)\sin(x) - y'(x)^2 dx.$$

Euler–Lagranges ekvation tillhörande funktionalen ovan ges av

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0,$$

där $f(x, y, y')$ återigen är $2y\sin(x) - y'^2$.

Vi har att

$$\frac{\partial f}{\partial y} = \frac{\partial\left(2y\sin(x) - y'^2\right)}{\partial y} = 2\sin(x),$$

och

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{d}{dx}\left(\frac{\partial\left(2y\sin(x) - y'^2\right)}{\partial y'}\right) = \frac{d}{dx}(-2y') = -2y'',$$

så Euler–Lagranges ekvation är alltså

$$-2y''(x) - 2\sin(x) = 0 \iff y''(x) = -\sin(x). \tag{7.11}$$

Den allmänna lösningen till denna differentialekvation ges av

$$y(x) = \sin(x) + ax + b,$$

Och om vi till exempel är intresserad av att hitta en extrempunkt till funktionalen I som dessutom uppfyller att $y(0) = y(\pi) = 0$ så innebär detta att $y(x) = \sin(x)$.

▲

Övningar

Övning 7.1 ($\star$).

(a) Skriv ut formeln för funktionalen $I : C^2([a, b]) \to \mathbb{R}$ associerad med $a = 0$, $b = \pi$ och $f(x, y, z) = y^2 - z^2$.

(b) Härled Euler–Lagranges ekvation för funktionalen från (a).

(c) Har Euler–Lagranges ekvation någon lösning som uppfyller bivillkoret att $y(a) = y(b) = 0$?

Om ja, hitta en lösning. Om nej, motivera varför en sådan ej kan existera.

Övning 7.2 ($\star$).

(a) Skriv ut formeln för funktionalen $I : C^2([a, b]) \to \mathbb{R}$ associerad med $a = 0$, $b = 1$ och $f(x, y, z) = z^2/2 - y$.

(b) Härled Euler–Lagranges ekvation för funktionalen från (a).

(c) Har Euler–Lagranges ekvation någon lösning som uppfyller bivillkoret att $y(a) = y(b) = 0$?

Om ja, hitta en lösning. Om nej, motivera varför en sådan ej kan existera.

Övning 7.3 ($\star\star$)**.**

(a) Skriv ut formeln för funktionalen $I : C^2([a, b]) \to \mathbb{R}$ associerad med $a = 0$, $b = 1$ och $f(x, y, z) = x^2 y$.

(b) Härled Euler–Lagranges ekvation för funktionalen från (a).

(c) Har Euler–Lagranges ekvation någon lösning som uppfyller bivillkoret att $y(a) = y(b) = 0$?

Om ja, hitta en lösning. Om nej, motivera varför en sådan ej kan existera.

(d) Har funktionalen något maximum då den begränsas till klassen av funktioner $y(x) \in C^2([a, b])$ som uppfyller att $y(a) = y(b) = 0$.

Om ja, motivera. Om nej, visa att det för varje tal $n \in \mathbb{N}$ existerar någon funktion i klassen så att $I(y) \geqslant n$.

Övning 7.4 ($\star\star$)**.**

(a) Skriv ut formeln för funktionalen $I : C^2([a, b]) \to \mathbb{R}$ associerad med $a = 0$, $b = 1$ och $f(x, y, z) = x + y + z$.

(b) Härled Euler–Lagranges ekvation för funktionalen från (a).

(c) Har Euler–Lagranges ekvation någon lösning som uppfyller bivillkoret att $y(a) = 0, y(b) = 1$?

Om ja, hitta en lösning. Om nej, motivera varför en sådan ej kan existera.

(d) Har funktionalen något maximum då den begränsas till klassen av funktioner $y(x) \in C^2([a, b])$ som uppfyller att $y(a) = 0, y(b) = 1$.

Om ja, motivera. Om nej, visa att det för varje tal $n \in \mathbb{N}$ existerar någon funktion i klassen så att $I(y) \geqslant n$.

Övning 7.5 ($\star$)**.** Skriv ut formeln för funktionalen $I : C^2([a, b]) \to \mathbb{R}$ associerad med $a = 0$, $b = 1$ och $f(x, y, z) = x + y + z^2$. Beräkna därefter $I(\tilde{y}_\alpha)$ då $y(x) = -x$ och $\eta(x) = x^2$ som en funktion av α. Beräkna slutligen

$$\frac{\partial}{\partial \alpha} I(\tilde{y}_\alpha),$$

och evaluera denna derivata då $\alpha = 0$.

Övning 7.6 ($\star$)**.** Skriv ut formeln för funktionalen $I : C^2([a, b]) \to \mathbb{R}$ associerad med $a = 0$, $b = \pi$ och $f(x, y, z) = y^2 + z^2$. Beräkna därefter $I(\tilde{y}_\alpha)$ då $y(x) = \sin(x)$ och $\eta(x) = x$ som en funktion av α. Beräkna slutligen

$$\frac{\partial}{\partial \alpha} I(\tilde{y}_\alpha),$$

och evaluera denna derivata då $\alpha = 0$.

Övning 7.7 ($\star\star$)**.** Låt $f(x, y)$ vara en differentierbar funktion och låt $\tilde{y}_\alpha := \alpha\eta(x)y(x)$. Använd kedjeregeln för att uttrycka

$$\frac{\partial}{\partial\alpha} f\left(\tilde{y}_\alpha(x), \tilde{y}'_\alpha(x)\right).$$

Vad ger detta om $\eta(x) = 1$?

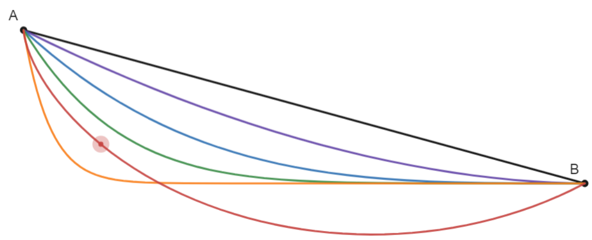
8 Lösningen av Brachistochronproblemet

I detta kapitel använder vi Euler–Lagranges ekvation för att lösa Brachistochronproblemet som vi introducerade i avsnitt 6.2 och som i avsnitt 6.3 visade sig vara ekvivalent med att minimera funktionalen i ekvation (6.3). Mer specifikt så kunde Brachistochronproblemet förenklas till att hitta en lämplig funktion $y(x)$ från $A = (0, 0)$ till $B = (X, -Y)$ som minimerar

$$\mathcal{T}(y) = \frac{1}{\sqrt{2g}} \int_0^X \sqrt{\frac{1 + y'^2}{y}} dx.$$

vilket motsvarar färdtiden för den släta kurvan

$$s(t) = \begin{bmatrix} x(t) \\ y(x(t)) \end{bmatrix} \quad x(0) = 0, \quad x(T) = X, \quad \|s'(t)\| = \sqrt{2gy(x(t))}.$$



Figur 8.1: Exempel på olika lämpliga funktioner y som beskriver en bana från A till B.

8.1 Användning av Euler–Lagranges ekvation

Vi kan nu verifiera att detta problem passar in i sats 7.1.1 genom att sätta $a = 0, b = X, y_1 = 0, y_2 = Y$ och

$$f(x, y, z) = \sqrt{\frac{1 + z^2}{y}}.$$

För att få göra det måste vi verifiera att f är 2 gånger kontinuerligt differentierbar. Vi beräknar

$$\operatorname{grad} f(x, y, z) = \left(0, -\frac{1}{2}\sqrt{\frac{1 + z^2}{y^3}}, \frac{z}{\sqrt{y(1 + z^2)}}\right)$$

$$\frac{\partial^2 f}{\partial z^2} = \frac{1}{\sqrt{y}} \frac{1}{(1 + z^2)^{3/2}}, \quad \frac{\partial^2 f}{\partial y^2} = \frac{3}{4}\sqrt{\frac{1 + z^2}{y^5}}, \quad \frac{\partial^2 f}{\partial z \partial y} = \frac{-z}{2\sqrt{y^3(1 + z^2)}}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial x \partial z} = 0.$$

Alltså är $f \in C^2(\{y \geqslant 0\})$, vilket låter oss använda Euler–Lagranges ekvation (sats 7.1.1). I detta fall har vi en extra snäll integrand i Euler–Lagranges ekvation. Vi kan utnyttja detta med hjälp av följande lemma.

Hjälpsats 8.1.1 (Beltramis identitet)**.** *Givet en funktional på formen*

$$I(y) = \int_a^b f(x, y(x), y'(x))dx$$

som uppfyller antagandena för Euler-Lagrange (sats 7.1.1) och dessutom,

$$\frac{\partial}{\partial x} f(x, y, z) = 0.$$

Då följer det från Euler–Lagranges ekvation att

$$f(x, y(x), y'(x)) - y'(x) \cdot \frac{\partial f}{\partial y'}(x) = C$$

för någon konstant $C \in \mathbb{R}$.

Bevis. Med hjälp av kedjeregeln i flera variabler (sats 4.1) så ser vi att

$$\begin{aligned} \frac{d}{dx} f(x, y(x), y'(x)) &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot y'(x) + \frac{\partial f}{\partial y'} \cdot y''(x) \\ &= \frac{\partial f}{\partial y} \cdot y'(x) + \frac{\partial f}{\partial y'} \cdot y''(x) \end{aligned}$$

I detta skede ger Euler–Lagrange (sats 7.1.1) att

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0.$$

vilket leder till

$$\frac{d}{dx} f(x, y(x), y'(x)) = y'(x) \cdot \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + \frac{\partial f}{\partial y'} \cdot y''(x) = \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \cdot y'\right).$$

Vi integrerar nu bägge sidor med avseende på x och erhåller

$$f(x, y(x), y'(x)) = \frac{\partial f}{\partial y'} \cdot y' + C.$$

□

Hjälpsats 8.1.2. *Den släta kurvan som löser Brachistochronproblemet motsvarar en funktion $y(x)$ (enligt sats 6.3.4 och definition 6.3.2) som uppfyller differentialekvationen*

$$(y')^2 = \frac{k^2}{y} - 1$$

för något tal $k \in \mathbb{R}$.

Bevis. Den optimala släta kurvan motsvarar en funktion y som måste uppfylla Beltramis identitet (hjälpsats 8.1.1), som vi sedan förenklar ytterligare.

$$\begin{aligned}
\sqrt{\frac{1+(y')^2}{y}} - y'\frac{y'}{\sqrt{y(1+(y')^2)}} &= C \iff \\
\frac{1}{\sqrt{y}}\left(\sqrt{1+(y')^2} - \frac{(y')^2}{\sqrt{(1+(y')^2)}}\right) &= C \iff \\
\sqrt{\frac{(1+(y')^2)}{y}}\left(1 - \frac{(y')^2}{1+(y')^2}\right) &= C \iff \\
\sqrt{\frac{(1+(y')^2)}{y}}\left(\frac{1+(y')^2}{1+(y')^2} - \frac{(y')^2}{1+(y')^2}\right) &= C \iff \\
\sqrt{\frac{(1+(y')^2)}{y}}(\frac{1}{1+(y')^2}) &= C \iff \\
\frac{1}{\sqrt{y(1+(y')^2)}} &= C.
\end{aligned}$$

Vi kvadrerar båda sidor av den senaste likheten och får,

$$\begin{aligned}
\frac{1}{\sqrt{y(1+(y')^2)}} &= C \iff \\
y(1+(y')^2) &= 1/C^2.
\end{aligned}$$

Notera att det är en ekvivalenspil då $y(1+(y')^2) \geqslant 0$. Kalla C för $1/k$ istället. Slutligen får vi

$$\begin{aligned}
y(1+(y')^2) &= k^2 \iff \\
1+(y')^2 &= k^2/y \iff \\
(y')^2 &= k^2/y - 1
\end{aligned}$$

□

Anmärkning 8.1.3. Kom ihåg exempel 3.4.11. Där visar vi att differentialekvationen från hjälpsats 8.1.2 som lösningen till Brachistochronproblemeti måste lösa. Differentialekvationen tvingar y' att vara en avtagande funktion! Idéen var att vi deriverade båda sidor av ekvationen $(y')^2 = \frac{k^2}{y} - 1$ och fick någonting i stilen med $y''(x) = -\frac{k^2}{2y^2} \leqslant 0$. Alltså kan differentialekvationen delas upp på olika intervall där vi vet att derivatan måste vara positiv på det ena och negativ på det andra.

Vi kan då fortsätta att förenkla ekvationen i 8.1.2. Vi förenklar den sådan att den blir separabel, så länge $y' \geqslant 0$.

$$\begin{aligned}
(y')^2 = k^2/y - 1 &\overset{y'\geq 0}{\iff} \\
y' = \sqrt{k^2/y - 1} = \sqrt{\frac{k^2-y}{y}} &\iff \\
\sqrt{\frac{y}{k^2-y}}y' &= 1.
\end{aligned}$$

Vi fortsätter nu att visa något ännu starkare, att om det finns en punkt x_0 med $y'(x_0) = 0$ så är y symmetrisk kring den.

Hjälpsats 8.1.4. *Om $y(x) \in [0, k^2], y(0) = 0$ är en lämplig funktion från A till B som löser differentialekvationen*

$$(y'(x))^2 = \frac{k^2}{y(x)} - 1$$

för något $k \in \mathbb{R}$ och det finns en punkt x_0 där $y(x_0) = k^2$. Då måste y vara symmetrisk kring x_0 d.v.s.

$$y(x_0) = k^2 \implies y(x_0 + x) = y(x_0 - x).$$

Bevis. Vi ser att $y' = 0 \iff y = k^2$ från den ursprungliga ekvationen. Alltså har vi att om y uppfyller alla antaganden i problemet så gäller det att båda funktionera $y(x_0 + x)$ och $y(x_0 - x)$ uppfyller den 'nya' differentialekvationen:

$$y'(x)^2 = \frac{k^2}{y(x)} - 1 \quad y(0) = k^2, \quad y'(0) = 0, \quad y \in [0, k^2].$$

Om vi kan visa att denna differentialekvation med begynnelsevillkor ($y(0) = k^2, y'(0) = 0$) har en unik lösning är vi färdiga. Det räcker med att visa det på en av sidorna om x_0 eftersom beviset är identiskt på båda sidor, vi behöver bara byta ut ordet *växande* mot *avtagande*. Låt oss studera $x \leqslant 0$ i den 'nya' differentialekvationen. Eftersom vi i exempel 3.4.11 visar att y konkav har vi att

$$\begin{aligned} y' = \sqrt{k^2/y - 1} = \sqrt{\frac{k^2 - y}{y}} &\iff \\ \sqrt{\frac{y}{k^2 - y}} y' = 1. \end{aligned}$$

Detta är en separabel differentialekvation där

$$f(y) = \sqrt{\frac{y}{k^2 - y}} > 0 \quad g(x) = 1.$$

Om vi låter

$$F(y) = \int_{k^2}^{y} \sqrt{\frac{s}{k^2 - s}} ds$$

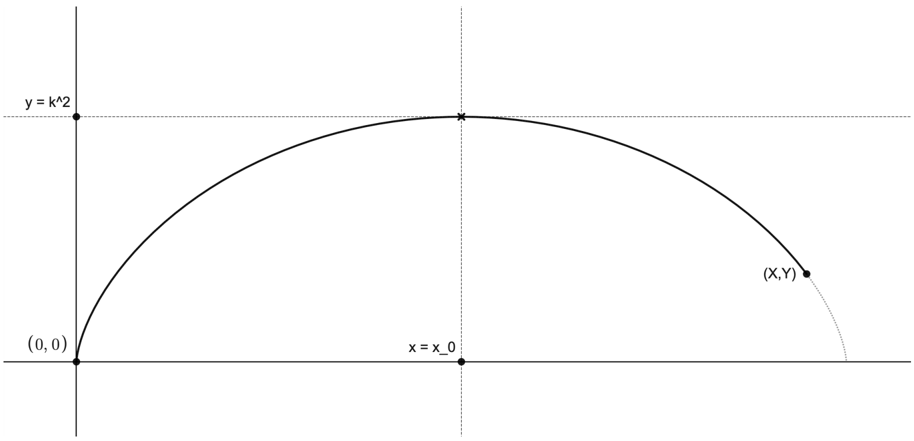
så att vi får

$$F(y(x)) = x$$

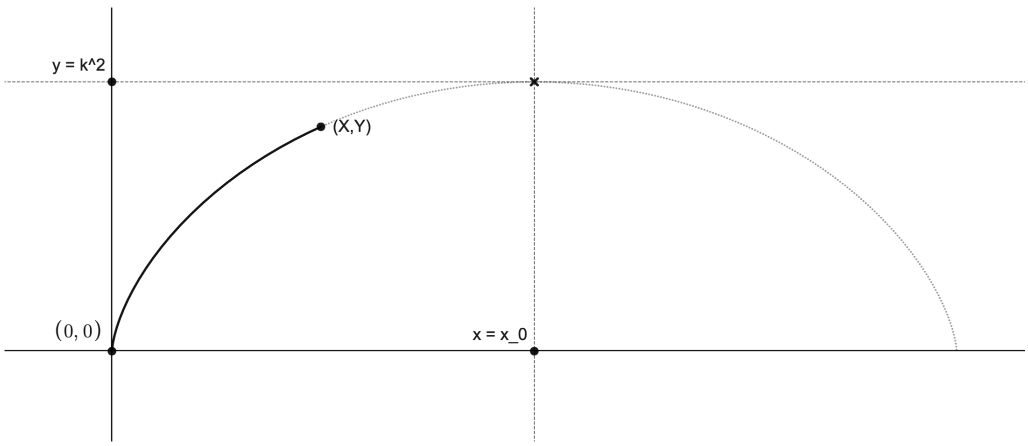
Denna integral F är ändlig eftersom $y(x) \leqslant k^2$ och då har vi

$$|F(y)| \leqslant \int_y^{k^2} \sqrt{\frac{k^2}{k^2 - s}} ds = |k| \int_y^{k^2} \frac{1}{\sqrt{k^2 - s}} ds = |k| \cdot \left[-2\sqrt{k^2 - s}\right]_y^{k^2} = 2|k|\sqrt{k^2 - y}.$$

Enligt hjälpsats 3.4.5 har vi en unik lösning eftersom en primitiv funktion F kommer vara strikt växande. $\square$



Figur 8.2: En visualisering av hur $y(x)$ har en avtagande derivata och är symmetrisk kring punkten där $y = k^2$.



Figur 8.3: En visualisering av hur en lösning $y(x)$ ibland inte 'hinner fram' till symmetripunkten innan den når slutet X, Y på sin bana.

Vi kan nu alltså nöja oss med att lösa ekvationen på ett intervall $[0, x_0]$ där vi vet att $\sqrt{\frac{y}{k^2-y}}y' = 1$ kommer att hålla.

Sats 8.1.5. *Lösningen till Brachistochronproblemet ges av en cykloid. Det betyder att den är en omparameterisering av en slät kurva i $\mathcal{S}^2_\Phi$ på formen*

$$\begin{cases} x = r(\varphi - \sin(\varphi)) \\ y = r(1 - \cos(\varphi)). \end{cases}$$

för något $r = \frac{k^2}{2}$ och $\Phi < 2\pi$.

Anmärkning 8.1.6. I övning 8.3 beskriver visar vi att omparameteriseringen är $\varphi = \sqrt{\frac{g}{r}}t$ som en funktion av tiden t och inser att färdtiden $\mathcal{T}$ är lösningen till ekvationsystemet

$$\begin{cases} X = r(\frac{g}{r}\mathcal{T} - \sin(\frac{g}{r}\mathcal{T})) \\ Y = r(1 - \cos(\frac{g}{r}\mathcal{T})). \end{cases}$$

Bevis. Vi utgår ifrån alla tidigare hjälpsatser i detta kapitel för att dra slutsatsen att det räcker att lösa differentialekvationen,

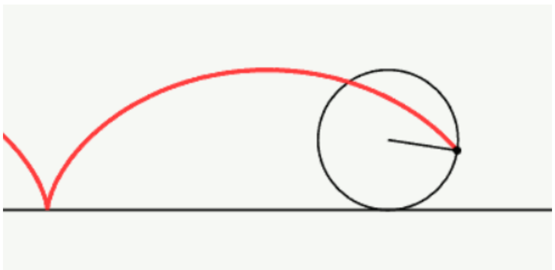
$$\sqrt{\frac{y}{k^2 - y}}y' = 1.$$

Detta ger oss en lösning $y(x)$ som är växande fram tills $y(x) = k^2$. För större x så använder vi symmetriargumentet från hjälpsats 8.1.4 för att förstå lösningen. Vi tar en närmare titt på symmetrin i övning 8.1. Vi ser att x är en primitiv funktion till högerledet i differentialekvationen ovan. Om vi kan hitta en primitiv funktion till vänsterledet som också är 0 då $y = 0$ kommer vi få ett samband mellan x och y, $F(y(x)) = x$ som beskriver lösningen till Brachistochronproblemet. Vi gör detta med hjälp av ett väldigt knepigt variabelbyte.

$$F(y(x)) = \int_0^x \sqrt{\frac{y(t)}{k^2 - y(t)}}y'(t)dt = \left\{ \begin{smallmatrix} y=y(t) & dy=y'(t)dt \\ y=0 \implies t=0 & y=y(x) \implies t=x \end{smallmatrix} \right\} =$$

$$\int_0^{y(x)} \sqrt{\frac{y}{k^2 - y}}dy = \left\{ \begin{smallmatrix} y=k^2\cdot\sin^2(\varphi/2) & dy=k^2\sin(\varphi/2)\cos(\varphi/2)d\varphi \\ \varphi=0 \implies y=0 & \varphi=\varphi(x) \implies y=y(x) \\ & \varphi\in[0,\pi] \end{smallmatrix} \right\}$$

här har vi låtit $\varphi(x) = 2\text{arcsin}\left(\sqrt{\frac{y(x)}{k^2}}\right)$. Notera att vi inte tillåter $\varphi > \pi$ för tillfället, därför att då går vi över symmetripunkten $y = k^2$ i hjälpsats 8.1.4 och differentialekvationen 'byter tecken'.



Figur 8.4: Bild på cykloid från Wikipedia. Vi löser bara brachistochronproblemet fram tills dess att vi når den högsta punkten på den röda banan, resten får vi av ett symmetriargument.

$$= \int_0^{\varphi(x)} \tan(\varphi/2) \cdot k^2\sin(\varphi/2)\cos(\varphi/2)d\varphi = k^2 \int_0^{\varphi(x)} \sin^2(\varphi/2)d\varphi.$$

Nu använder vi lite trigonometri

$$\cos(2\varphi) = \cos^2(\varphi) - \sin^2(\varphi) = \cos^2(\varphi) - \sin^2(\varphi) - \sin^2(\varphi) - \cos^2(\varphi) + 1$$
$$= 1 - 2\sin^2(\varphi) \implies \sin^2(\varphi) = \frac{1}{2}\Big(1 - \cos(2\varphi)\Big).$$

Vilket ger oss

$$F(y(x)) = k^2 \int_0^{\varphi(x)} \sin^2(\varphi/2) d\varphi = \frac{k^2}{2} \int_0^{\varphi(x)} 1 - \cos(\varphi) d\varphi$$
$$= \frac{k^2}{2} \Big[\varphi - \sin(\varphi)\Big]_0^{\varphi(x)} \implies F(y(x)) = \frac{k^2}{2}\Big(\varphi(x) - \sin(\varphi(x))\Big).$$

Vi drar slutsatsen att, om vi låter $r = k^2/2$, och manipulerar formeln $\varphi(x) = 2\mathrm{arcsin}\left(\sqrt{\frac{y(x)}{k^2}}\right)$ så gäller

$$\begin{cases} x = F(y(x)) = r(\varphi(x) - \sin(\varphi(x))) \\ y = k^2 \sin^2(\varphi(x)) = r(1 - \cos(\varphi(x))) \end{cases}.$$

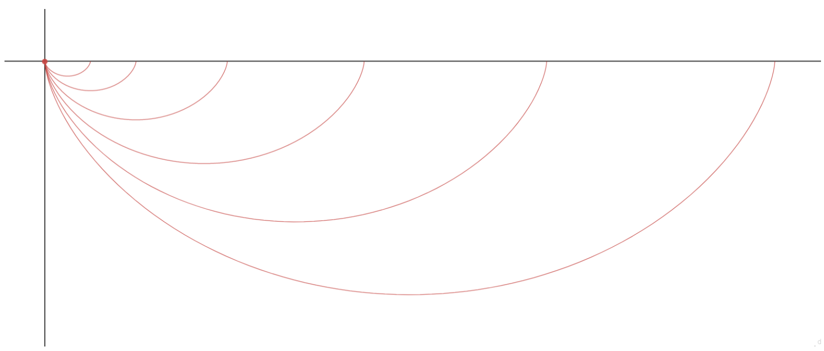
Vi avslutar beviset med att visa att $\varphi(x)$ är en omparametersiering av $[0, r\pi]$ till $[0, \pi]$. Vi har att

$$\varphi'(x) = 2\left(\mathrm{arcsin}(\sqrt{\frac{y(x)}{k^2}})\right)' = \frac{2}{\sqrt{1 - \frac{y(x)}{k^2}}}\Big(\sqrt{\frac{y(x)}{k^2}}\Big)' =$$
$$\frac{2}{\sqrt{1 - \frac{y(x)}{k^2}}}\Big(\frac{1}{2\sqrt{\frac{y(x)}{k^2}}}\frac{y'(x)}{k^2}\Big) \overset{y'>0}{>} 0$$

och $x = 0 \iff \varphi = 0, \varphi = \pi \iff x = r\pi$. Att verifiera att denna lösning som kan utvidgas till $\varphi > \pi$ blir symmetrisk i enlighet med hjälpsats 8.1.4 lämnas som övning 8.1. Då har vi vår lösning. $\square$

Anmärkning 8.1.7. Att verifiera och studera lösningen gör vi i form av ett flertal övningar i detta kapitel.

Anmärkning 8.1.8. Givet X, Y finns det ett unikt par $r \in \mathbb{R}_{>0}, \Phi \in [0, 2\pi]$ sådan att $(x(\Phi), y(\Phi)) = (X, Y)$. Detta följer från att $x - \sin(x)$ är inverterbar, men inversen går inte att uttrycka med vanliga funktioner så att skriva ned lösningen ger inga särskilda insikter såvitt vi vet. Vi nöjer oss med exempel på cykloider.



Figur 8.5: Cykloider för olika värden på r för $\varphi \in [0, 2\pi]$. Bild gjord i Desmos.

Övningar

Övning 8.1. Verifiera att cykloiden som parameteriseras av

$$\begin{cases} x = r(\varphi - \sin(\varphi)) \\ y = r(1 - \cos(\varphi)). \end{cases}$$

är symmetrisk kring $\varphi = \pi$, dvs att $x(2\pi - \varphi) + x(\pi) = 2\pi r$ och att $y(2\pi - \varphi) = y(\pi)$.

Övning 8.2. Vilka värden på r och $\varphi \in [0, 2\pi]$ får den motsvarande cykoliden som startar i origo att passera punkten $(3.524, -1.982)$?

(a) $r = \pi$, $\varphi = \pi$.

(b) $r = 1$, $\varphi = \frac{4\pi}{3}$.

(c) $r = \frac{13}{10}$, $\varphi = \frac{10}{3}$.

Övning 8.3 ($\star$)**.** Bevisa att följande förhållande gäller för kurvan som löser Brachistochronproblemet.

$$\varphi(t) = \sqrt{\frac{g}{r}}t.$$

(a) Ställ upp färdtidsintegralen $t = \frac{1}{\sqrt{2g}} \int_0^{x(t)} \sqrt{\frac{1+y'(x)^2}{y(x)}} dx$.

(b) Använd hjälpsatsen 8.1.2 och förenkla den till $t = \sqrt{\frac{r}{g}} \int_0^{x(t)} \frac{1}{y(x)} dx$

(c) Gör variabelbytet $x = x(\varphi)$ och dra slutsatsen utifrån $t = \sqrt{\frac{r}{g}} \int_0^{\varphi(t)} 1 d\varphi$.

Övning 8.4 ($\star$)**.** Cykloiden som parameteriseras av

$$s(\varphi) = r \cdot \Big(\varphi - \sin(\varphi), -(1 - \cos(\varphi)\Big).$$

kan också beskrivas som grafen av en funktion $-y(x)$ som i figuren 8.5 (enligt sats 6.3.4 och definition 6.3.2). Visa att $y'(\pi r) = 0$.

Övning 8.5 ($\star$)**.** Ge ett analogt påstående till Beltramis identitet 8.1.1 men där istället $\frac{d}{dy} f(x, y, z) = 0$

Övning 8.6 ($\star$)**.** I enlighet med uppgift 6.8, tillämpa Euler–Lagrange på funktionalen

$$\mathcal{L}(s) = \int_0^X \sqrt{1 + y'(x)^2} dx.$$

för att hitta den kurva $s(x) = (x, y(x))$ med $y(0) = 0$, $y(X) = Y$ som minimerar färdsträckan mellan $(0, 0)$ och (X, Y). *Använd övning 8.5.*

Lösningar till udda övningsuppgifter

Kapitel 1

Övning 1.1. (i) 0, 1, 2, 3, 4.

(ii) 1, 2, $\{2, 3\}$.

(iii) -3, -2, -1, 0, 1, 2, 3.

(iv) 1, 2.

(v) -3, -2, -1, 0, 1, 2, $\{2, 3\}$.

Övning 1.3. (i) B är en äkta delmängd av A.

(ii) A och B är lika.

(iii) Mängderna är disjunkta eftersom $B = \emptyset$.

(iv) Mängderna är varken disjunkta, lika eller äkta delmängder av varandra.

(v) Mängderna är disjunkta eftersom $A = \emptyset$.

Övning 1.5. Observera att dessa svar är förslag. Uppgifterna har flera korrekta svar.

(i) $\{n \in \mathbb{Z} \mid n = 2k \text{ för något heltal } k \geqslant 1\}$.

(ii) $\{p/2 \mid p \text{ är ett heltal }\}$.

(iii) $\{r \in \mathbb{R} \mid r \notin \mathbb{Q} \text{ och } |r| < 1\}$.

Övning 1.7. (i) Definitionsmängd: $\{1, 2, 3, \ldots\}$. Målmängd: $\mathbb{N}$.

(ii) Definitionsmängd: $\{T \mid T \text{ är en triangel }\}$. Målmängd: $\mathbb{R}$.

(iii) Definitionsmängd: $\{p(x) \mid p(x) \text{ är ett andragradspolynom}\}$. Målmängd: $\{p(x) \mid p(x) \text{ är ett förstagradspolynom}\}$.

Övning 1.9. (i) Detta är en funktion. Den är definierad för alla värden i definitionsmängden, den ger alltid samma värde för ett givet argument, och alla funktionsvärden ligger i målmängden.

(ii) Detta är inte en funktion, då funktionens värden inte ligger i målmängden ($f(2) = \sqrt{2} \notin \mathbb{Q}$).

(iii) Detta är inte en funktion, eftersom dess värden är slumpmässiga.

(iv) Detta är en funktion. Eftersom ordet Balkong börjar på B, så har funktionen ett definierat värde som ligger i målmängden.

Övning 1.11. (i) Funktionerna är lika. De har samma definitionsmängd, målmängd och $\sqrt{x^2} = |x|$ för alla x.

(ii) Funktionerna är inte lika, då deras definitionsmängder inte är samma.

(iii) Funktionerna är inte lika, då deras målmängder är olika.

Övning 1.13. Om n är jämnt så finns det per definition ett heltal k så att $n = 2k$. Då gäller att $n + 1 = 2k + 1$, det vill säga $n + 1$ är udda.

Övning 1.15. Antag motsatsen, det vill säga att det finns ett rationellt tal p/q och ett irrationellt tal r vars summa är rationell. Skriv summan som kvoten s/t. Då gäller

$$\frac{p}{q} + r = \frac{s}{t} \implies r = \frac{s}{t} - \frac{p}{q} = \frac{sq - pt}{tq}$$

genom att skriva vänsterledet på gemensamt bråkstreck. Men detta bevisar att r är rationellt, vilket är en motsägelse.

Övning 1.17. Basfallet är $n = 1$, och då gäller att $1 = 1^2$. För induktionssteget, antag att

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2$$

för något n. Då gäller att

$$1 + 3 + 5 + \cdots + (2n - 1) + (2n + 1) = n^2 + 2n + 1 = (n + 1)^2,$$

vilket avslutar induktionssteget och därmed beviset.

Övning 1.19. Antag motsatsen till satsen. Den kan delas in i två fall. I det ena fallet gäller att $ab = c$ och $a < \sqrt{c}$ och $b < \sqrt{c}$. Då gäller att

$$ab < a\sqrt{c} < \sqrt{c}\sqrt{c} = \sqrt{c}^2 = c.$$

Alltså gäller $ab < c$, vilket är en motsägelse.

I det andra fallet gäller att $ab = c$ och $a > \sqrt{c}$ och $b > \sqrt{c}$. Då gäller att

$$ab > a\sqrt{c} > \sqrt{c}\sqrt{c} = \sqrt{c}^2 = c.$$

Alltså gäller $ab > c$, vilket är en motsägelse.

Övning 1.21. (i) Per definition gäller $n = 0 + n = n + 0$, det vill säga $n \geqslant 0$ och $n \geqslant n$.

(ii) Enligt antagandet $n \geqslant m$ så finns det a så att $n = m + a$. Om dessutom $m \geqslant k$ så finns det b så att $m = k + b$. Då gäller

$$n = m + a = k + (b + a),$$

det vill säga $n \geqslant k$.

(iii) Om $n \geqslant m$ och $m \geqslant n$ så gäller att $n = m + k$ och $m = n + l$. Så

$$n = m + k = n + l + k \implies l + k = 0.$$

Det ger $l = k = 0$, det vill säga $n = m$.

(iv) Om $n \geqslant m$ så finns det ett a så att $n = m + a$. Då gäller

$$n + k = m + a + k = m + k + a,$$

alltså att $n + k \geqslant m + k$.

(v) Låt $n \geqslant m$, så att det finns ett tal a så att $n = m + a$. Vi ska använda induktion över k för att bevisa att $nk \geqslant mk$. Basfallet är $k = 0$, och då gäller att

$$nk = 0 \geqslant 0 = mk.$$

Antag att $nk \geqslant mk$ för något k. Det finns då b så att $nk = mk + b$. Då gäller att

$$n(k+1) = nk + n = mk + a + m + b = m(k+1) + a + b,$$

vilket bevisar att $n(k+1) \geqslant m(k+1)$. Detta avslutar induktionssteget och hela beviset.

Övning 1.23. Basfallet är $n = 2$, och då gäller att $1^2 + 2^2 = 5 < 2^3 = 8$.

Antag nu att $1^2 + 2^2 + \cdots + m^2 < m^3$ gäller för något m. Då har vi att

$$1^2 + 2^2 + \cdots + m^2 + (m+1)^2 < m^3 + m^2 + 2m + 1.$$

Om man multiplicerar ihop $(m+1)^3$ så ser vi att

$$(m+1)^3 = m^3 + 3m^2 + 3m + 1.$$

Differensen mellan högerleden blir

$$m^3 + 3m^2 + 3m + 1 - (m^3 + m^2 + 2m + 1) = 2m^2 + m$$

vilket är större än 0 när $m \geqslant 2$. Alltså gäller

$$1^2+2^2+\cdots+m^2+(m+1)^2 < m^3+m^2+2m+1 < m^3+3m^2+3m+1 = (m+1)^3.$$

Detta avslutar induktionssteget och därmed hela beviset.

Övning 1.25. Satsen att alla icke-tomma delmängder av $\mathbb{N}$ har ett minsta element kan omformulera som att alla mängder $S \subset \mathbb{N}$ som inte har ett minsta element är tom. Så om S inte har ett minsta element så gäller att $n \notin S$ för alla $n \in \mathbb{N}$. Detta är ekvivalent med att

$$0, 1, 2, \ldots, n \notin S$$

för alla n. Detta kan vi bevisa med induktion.

Låt S vara en mängd utan minsta element. Basfallet är att $0 \notin S$, vilket måste stämma eftersom $0 \leqslant n$ för alla $n \in \mathbb{N}$, så om $0 \in S$ har S ett minsta element, vilket är en motsägelse.

Anta nu att inget av talet 0, 1, $\ldots$, m ligger i S. Antag att $m + 1 \in S$. Då kommer $m + 1$ vara ett minsta element i S, eftersom S inte innehåller något mindre tal. Detta motsäger att S inte har ett minsta element. Alltså gäller att 0, 1, $\ldots$, $m + 1$ inte ligger i S Detta avslutar induktionen, och bevisar att S är tom.

Kommentar: *Detta resultat kallas för välordningsprincipen, och säger att* $\mathbb{N}$ *utgör en så kallad välordning. Dessa är användbara, då de tillåter oss att använda induktionsliknande resonemang.*

Vissa hävdar att induktion och välordningsprincipen är ekvivalenta. Detta stämmer inte, då induktion förutsätter att alla element i mängden kan nås genom ändligt många steg, vilket inte välordningsprincipen gör.

Övning 1.27. Mening (iv) är den vanliga definitionen av en rätvinklig triangel.

Mening (i) fungerar inte för att alla trianglar, även de som inte har en rät vinkel, har vinkelsumma 180 grader.

Mening (iii) är för snäv, då det finns rätvinkla trianglar som inte kan bildas genom att dra en diagonal i en kvadrat.

Mening (ii) är lurig. Eftersom en triangel har vinkelsumma 180 grader, kommer en triangel där två vinklar har summan 90 att vara rätvinklig, och vice versa. För trianglar är alltså påståendena ekvivalenta.

Men det finns många geometriska figurer med fler hörn som har två vinklar vars summa är 90 grader. Om (ii) var definitionen av en rätvinklig triangel, skulle vissa fyrhörningar rent formellt kunna vara rätvinkliga trianglar, vilket vore absurt.

Kapitel 2

Övning 2.1. (a)

$$|\mathbf{x}| = \sqrt{1^2 + 3^2 + 5^2 + 7^2} = \sqrt{84}$$

och

$$|\mathbf{y}| = \sqrt{(-1)^2 + 3^2 + (-5)^2 + 1^2} = \sqrt{36} = 6.$$

(b) $\mathbf{x} \cdot \mathbf{y} = -1 + 3^2 - 5^2 + 7 = -10.$

(c) Nej. Ett λ som uppfyller att $\lambda\mathbf{x} = \mathbf{y}$ Måste både uppfylla att $\lambda = -1$ och $3\lambda = 3$, vilket är omöjligt.

Övning 2.3. Basfallet $n = 2$ är den vanliga triangelolikheten.

För att bevisa induktionssteget gör vi induktionsantagandet att olikheten vi vill bevisa håller för n termer, och vill nu visa att detta implicerar att olikheten också håller för $n + 1$ termer.

Genom att använda triangelolikheten för 2 termer ser vi att

$$|x_1 + \ldots + x_{n+1}| = |(x_1 + \ldots + x_n) + x_{n+1}| \leqslant |x_1 + \ldots + x_n| + |x_{n+1}|.$$

Genom att nu använda induktionsantagandet på den första termen ser vi att

$$|x_1 + \ldots + x_n| + |x_{n+1}| \leqslant |x_1| + \ldots + |x_n| + |x_{n+1}|.$$

Tillsammans ger dessa två olikheter det önskade resultatet.

Övning 2.5. Låt $0 < a < 1$ vara en godtycklig punkt i $(0, 1)$. Vi vill visa att a är en inre punkt, det kommer vi göra genom att visa att det finns något tillräckligt litet δ så att $B_\delta(a) \subset (0, 1)$.

Låt nu

$$\delta = \frac{\min((a - 0), (1 - a))}{2},$$

det vill säga halva avståndet från a till den närmsta av punkterna 0 och 1 (samma bevis kommer såklart att fungera om vi väljer något ännu mindre δ, och i själva verket för vissa δ som är lite större).

Antag först att $\delta = a/2$, vilket då även betyder att $a \leqslant (1-a)$. Då är

$$\begin{aligned} B_\delta(a) = (a - a/2, a + a/2) \subset (a - a/2, a + (1-a)/2)) &= (a/2, 1/2 + a/2) \\ &\subset (a/2, 1) \subset (0,1). \end{aligned}$$

Notera att $(1+a)/2 < 1$ eftersom $a < 1$.

Antag nu att $\delta = (1-a)/2$, vilket då även betyder att $(1-a) \leqslant a$. Då är

$$\begin{aligned} B_\delta(a) = (a - (1-a)/2, a + (1-a)/2) &\subset (a - a/2, a + (1-a)/2)) \\ &= (a/2, 1/2 + a/2) \subset (a/2, 1) \subset (0,1). \end{aligned}$$

Notera återigen att $(1+a)/2 < 1$ eftersom $a < 1$.

Eftersom a var en godtycklig punkt i mängden är beviset nu klart.

Övning 2.7. Från en tidigare övningsuppgift vet vi att den öppna enhetsskivan

$$\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\},$$

är en öppen mängd, och således är det en delmängd till det inre av den slutna enhetsskivan.

Denna delmängd är i själva verket hela det inre, men för att visa detta behöver vi visa att ingen punkt som inte ligger i denna delmängd är en inre punkt.

Punkterna som ligger i den slutna enhetsskivan, men inte i den öppna enhetsskivan är precis punkterna på enhetscirkeln. Det vill säga $\{(x,y) : x^2+y^2 = 1\}$. Vi ska visa att alla punkter på enhetscirkeln är randpunkter. Detta gör vi genom att visa att om (x,y) ligger på enhetscirkeln så kommer varje liten öppen boll centrerad i (x,y) innehålla punkter utanför den slutna enhetsskivan.

Alltså, om (x,y) ligger på enhetscirkeln, det vill säga $x^2 + y^2 = 1$, så kommer $B_\varepsilon((x,y))$ innehålla alla punkter vars avstånd till (x,y) är strikt mindre än ε. Till exempel innehåller $B_\varepsilon((x,y))$ punkten $(1+\varepsilon/2)(x,y))$. Men $(1+\varepsilon/2)(x,y))$ ligger *inte* i den slutna enhetsskivan eftersom

$$|(1+\varepsilon/2)(x,y)|^2 = (1+\varepsilon/2)^2 x^2 + (1+\varepsilon/2)^2 y^2 = (1+\varepsilon/2)^2 > 1.$$

Eftersom $\varepsilon > 0$ var godtyckligt visar detta att (x,y) inte är en inre punkt, och eftersom (x,y) var en godtycklig punkt på enhetscirkeln visar detta att ingen punkt utanför den öppna enhetsskivan är en inre punkt, och således är det inre av den slutna enhetsskivan helt enkelt den öppna enhetsskivan.

Kapitel 3

Övning 3.1. Lösningen ges av analysens fundamentalsats

$$\begin{aligned} \int_1^2 x^3 - 2x^2 + \pi dx = \left[\frac{x^4}{4} - 2\frac{x^3}{3} + \pi x\right]_1^2 &= (\frac{2^4}{4} - 2\frac{2^3}{3} + 2\pi) - (\frac{1}{4} - 2\frac{1}{3} + \pi) \\ &= \frac{15}{4} - \frac{14}{3} + \pi = \frac{45-56}{12} + \pi = -\frac{11}{12} + \pi. \end{aligned}$$

Övning 3.3. Funktionen $-\ln(\cos(x))$ är primitiv till $\tan(x)$ på $(-\frac{\pi}{2}, \frac{\pi}{2})$. Detta får vi genom att beräkna

$$\int_0^x \tan(\mathrm{t})dt = \int_0^x \frac{\sin(\mathrm{t})}{\cos(\mathrm{t})}dt = \left\{ {}^{\;\;\;\;\;\;\;\;\;\;\;\;\;y=\cos(x) \quad dy=-\sin(t)dt}_{t=0 \implies y=\cos(0)=1 \quad t=x \implies y=\cos(x)} \right\} =$$

$$\int_1^{\cos(x)} \frac{-1}{y} dy = [-\ln(y)]_1^{\cos(x)} = -\ln(\cos(x)).$$

Funktionen $-\frac{2x}{(1+x^2)^2}$ är derivatan till $\frac{1}{1+x^2}$ på hela $\mathbb{R}$. Detta får vi genom att beräkna

$$(\frac{1}{1+x^2})' = \frac{0(1+x^2)-2x}{(1+x^2)^2} = \frac{-2x}{(1+x^2)^2}.$$

För $\arcsin(x)$ låter vi $f(x) = 1$ och $g(x) = \arcsin(x)$ i formeln för partiell intergation.

$$\int_0^x \arcsin(t)dt = \Big[t \cdot \arcsin(\mathrm{t})\Big]_0^x - \int_0^x \frac{t}{\sqrt{1-t^2}}dt = \{\text{varibelbyte } y = t^2\}$$

$$x \cdot \arcsin(\mathrm{x}) + \frac{1}{2}\sqrt{1-x^2}.$$

Övning 3.5.

$$\int_1^x \ln(t)dt = [t\ln(t)]_1^x - \int_1^x t\frac{1}{t}dt =$$

$$[t\ln(t)]_1^x - \int_1^x 1dt = x\ln(t) - 0 - (x-1) = x(\ln(x) - 1 + 1).$$

Övning 3.7. Antag att $x_1 \neq x_2$, då kan vi anta utan förlust av generalitet att $x_1 < x_2$. Per definition $f(x_1) < f(x_2) \implies f(x_1) \neq f(x_2)$. Alltså måste olika x ge olika funktionsvärden.

Övning 3.9. Eftersom y löser ekvationen ser vi att $(y'(\pi))^2 = \frac{1}{1/2} - 1 = 1 \implies (y'(\pi))^2 = 1 \implies y'(\pi) = \pm 1$.

Övning 3.11. Beräkning är som följer

$$\frac{d}{dt}\frac{1}{2}\left(x\sqrt{1+x^2} + \ln(x+\sqrt{1+x^2})\right)$$

$$= \frac{1}{2}\left((\sqrt{1+x^2} + \frac{x^2}{\sqrt{1+x^2}}) + \frac{1+\frac{x}{\sqrt{1+x^2}}}{(x+\sqrt{1+x^2})}\right)$$

$$= \frac{1}{2\sqrt{1+x^2}}\left((1+x^2+x^2) + \frac{\sqrt{1+x^2}+x}{(x+\sqrt{1+x^2})}\right)$$

$$= \frac{1}{2\sqrt{1+x^2}}\big((1+x^2+x^2)+1\big) = \frac{1+x^2}{\sqrt{1+x^2}} = \sqrt{1+x^2}.$$

Kapitel 4

Övning 4.1. För ett fixt värde på x, y är

$$f'_x(x, y) = \cos(\cos(xy))(-\sin(xy))y$$

och

$$f'_y(x, y) = \cos(\cos(xy))(-\sin(xy))x$$

Övning 4.3. Det följer omedelbart från definitionen av differentierbarhet att

$$\lim_{\mathbf{h}\to 0} f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \lim_{\mathbf{h}\to 0} A_1h_1 + \dots A_nh_n + |\mathbf{h}|g(\mathbf{h}) = 0,$$

vilket innebär att f är kontinuerlig i punkten $\mathbf{a}$.

Övning 4.5. Sammansättningen ges av $f(\mathbf{g}(t)) = \sin(t^2 + t^3)$, och derivatan ges av $(2t + 3t^2)\cos(t^2 + t^3)$.

Om vi istället använder ekvation (4.1) ser vi att derivatan ges av

$$f'_1(\mathbf{g}(t))g'_1(t) + f'_2(\mathbf{g}(t))g'_2(t).$$

Genom att använda att $f'_1 = f'_2 = \cos(x + y), g'_1(t) = 2t$, och $g'_2(t) = 3t^2$ ger ekvationen ovan också att

$$\frac{d}{dt} f \circ \mathbf{g}(t) = (2t + 3t^2)\cos(t^2 + t^3).$$

Övning 4.7. Sammansättningen ges av $f(\mathbf{g}(t)) = t^3 + t^4 + t^3$, och derivatan ges av $6t^2 + 4t^3$.

Om vi istället använder ekvation (4.1) ser vi att derivatan ges av

$$f'_1(\mathbf{g}(t))g'_1(t) + f'_2(\mathbf{g}(t))g'_2(t) + f'_3(\mathbf{g}(t))g'_3(t).$$

Genom att använda att $f'_1 = 1, f'_2 = 2y, f'_3 = 3z^2, g'_1(t) = 3t^2, g'_2(t) = 2t$ och $g'_3(t) = 1$ ger ekvationen ovan också att

$$\frac{d}{dt} f \circ \mathbf{g}(t) = 3t^2 + (2t^2)(2t) + 3t^2 = 6t^2 + 4t^4.$$

Övning 4.9. Vi har att $f'_x = 2x\cos(x^2+y)$, så $f''_{xx} = 2\cos(x^2+y) - 4x^2\sin(x^2+y)$, och $f''_{xy} = -2x\sin(x^2+y)$. Vidare är $f'_y = \cos(x^2+y)$, så $f''_{yx} = -2x\sin(x^2+y)$ och $f''_{yy} = -\sin(x^2 + y)$. Notera att $f''_{xy} = f''_{yx}$, vilket är att vänta då funktionen är av klass C^2.

Kapitel 5

Övning 5.1. De partiella derivatorna ges av

$$f'_x(x, y) = 2xye^{x^2y} \text{ och } f'_y(x, y) = x^2e^{x^2y},$$

och gradienten i en punkt (x, y) är således vektorn

$$\left(2xye^{x^2y}, x^2e^{x^2y}\right).$$

Övning 5.3. Till att börja med är definitionsmängden D_f hela $\mathbb{R}^3$.

Eftersom var och en av termerna är icke-negativ följer det omedelbart att

$$f(x, y, z) \geqslant 0 = f(0, 0, 0).$$

Vidare är $x^4 > 0$ om $x \neq 0$, $y^2 > 0$ om $y \neq 0$, och $z^2 > 0$ om $z \neq 0$, vilket innebär att

$$f(x, y, z) > 0 = f(0, 0, 0) \text{ för alla } (x, y, z) \neq 0,$$

och f har således ett globalt minimum i orgio.

Övning 5.5. Till att börja med ges gradienten av

$$\text{grad}(f)(x, y) = (2xy, x^2).$$

Dock är den givna riktningen inte normerad, så vi måste först normera den till

$$\frac{v}{\|v\|} = \frac{(1, 2)}{\sqrt{1^2 + 2^2}} = \frac{(1, 2)}{\sqrt{5}}.$$

Riktningsderivatan ges nu av

$$\text{grad}(f)(x, y) \cdot \frac{v}{\|v\|} = \frac{2xy + 2x^2}{\sqrt{5}}.$$

Övning 5.7. Gradienten ges av

$$\text{grad}(f)(x, y) = (\cos(y)e^{x\cos(y)}, -x\sin(y)e^{x\cos(y)}),$$

så

$$\text{grad}(f)(0, 0) = (1, 0) \neq (0, 0),$$

så origo kan omöjligt vara en lokal extrempunkt.

Övning 5.9. Till att börja med ges gradienten till f av

$$\text{grad}(f)(x, y, z) = (2xy^2z, 2x^2yz, x^2y^2),$$

så

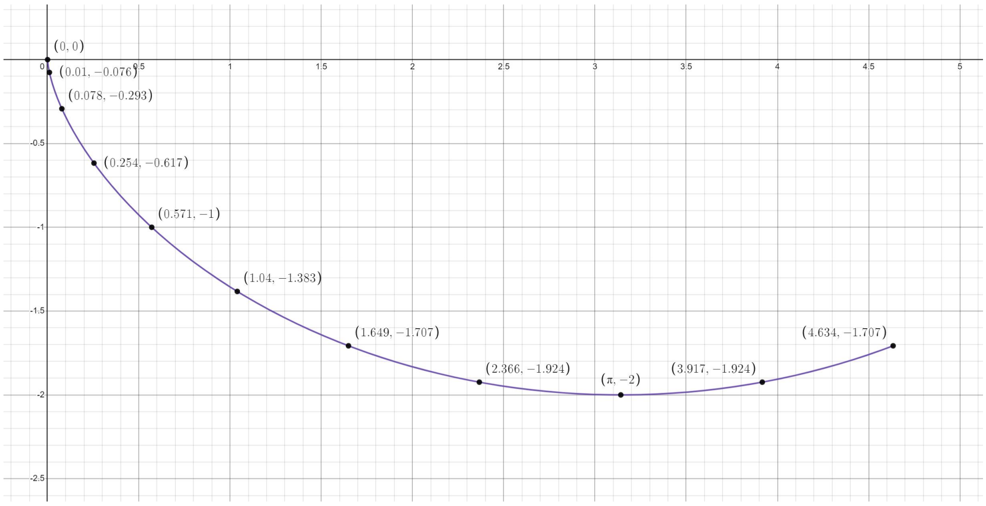
$$\text{grad}(f)(-1, -1, 3) = (-6, -6, 1).$$

Funktionen förändras snabbast i gradientens riktning, det vill säga i riktiningen $(-6, -6, 1)/\sqrt{73}$, och det maximala mätetalet på förändringen i den riktningen ges av

$$|\text{grad}(f)(-1, -1, 3)| = |(-6, -6, 1)| = \sqrt{73}.$$

Kapitel 6

Övning 6.1. Bilden bör se ut ungefär såhär,



Figur 6.6: Bild gjord i Desmos.

Övning 6.3.

(a) Ja.

(b) Ja

(c) Nej, I(f) ger en funktion, inte ett reellt tal.

(d) Nej, fisk är inte ett reellt tal.

(e) Ja, detta är samma funktional som (g) eftersom $\int_0^{\cos(x)} f(t)dt = e^{-f(\cos(x))\cdot\sin(x)}$ är alltid 1 när $x = 0$.

(f) Ja

(g) Ja

(h) Nej, e^x är en funktion och vi har inte fått x specificerat.

Övning 6.5. Lösningen går ut på att utföra variabelbytet $y = e^x$ och sedan använda övning 3.11.

$$L(s) = \int_0^1 |s'(t)|dt = \int_0^1 e^t\sqrt{1+e^{2t}}dt = \left\{ t{=}0 \overset{y=e^t}{\Longrightarrow} y{=}1 \quad t{=}1 \overset{dy=e^t dt}{\Longrightarrow} y{=}e \right\} =$$
$$\int_1^e \sqrt{1+y^2}dy = \frac{1}{2}\left[y\sqrt{1+y^2} + \ln(y+\sqrt{1+y^2})\right]_1^e =$$
$$\frac{1}{2}\left[e\sqrt{1+e^2} + \ln(e+\sqrt{1+e^2})\right] - \frac{1}{2}\left[\sqrt{2} + \ln(1+\sqrt{2})\right]$$

Övning 6.7. Lösningen är:

- Följ övningen 3.7.
- $x(0) = 0, x(T) = x(T)$ och $x'(t) > 0$.
- $\tilde{s}(x(t)) = (x(t), (y \circ x^{-1})(x(t))) = (x(t), y(t)) = s(t)$.

Övning 6.9. Idéen är att givet en kurva $s(t) = (x_1(t), \ldots, x_n(t)) \in \mathcal{S}_T^n$ så vill vi omparameterisera den så att $v(t) = f(t)$ för någon godtycklig funktion f. Vi då kan hitta en den önskade omparameteriseringen γ som lösning till den separabla differentialekvationen

$$(s \circ \gamma)'(t) = \gamma'(t) \cdot s'(\gamma(t)) = f(t)$$

Vi kan formulera den som

$$\left(\sqrt{\sum_{i=1}^{n} x_i'^2(\gamma(t))}\right) \gamma'(t) = f(t).$$

I enlighet med 3.4.5 så existerar en, strikt växande och inverterbar, funktion $L(y) = \int_0^y \sqrt{\sum_{i=1}^n x_i'^2(u)}du$ sådan att när vi utför operationen $\int_0^t$ på bägge sidor erhåller vi

$$L(\gamma(t)) = \int_0^t f(x)dx.$$

Alltså är den eftersökta omparameteriseringen $\gamma(t) = L^{-1}(\int_0^t f(x)dx)$.

Kapitel 7

Övning 7.1.

(a) Funktionalen ges av

$$I : y(x) \mapsto \int_0^\pi y(x)^2 - y'(x)^2 dx.$$

(b) Euler–Lagranges ekvation tillhörande funktionalen ovan ges av

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0,$$

där $f(x, y, y')$ är $y^2 - y'^2$.

Till att börja med är $\partial f/\partial y' = -2y'$, och därmed är

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{d}{dx}\left(-2y'\right) = -2y''.$$

Vidare är $\partial f/\partial y = 2y$. Euler–Lagranges ekvation ges sålunda av

$$-2y'' - 2y = 0.$$

(c) Vi vill hitta en funktion y som uppfyller att

$$y''(x) = -y(x),$$

och $y(0) = y(\pi) = 0$.

Det finns flera sett att göra detta på. I detta fall kan man nästan omedelbart se att både $\cos(x)$ och $\sin(x)$ är lösningar till differentialekvationen, och således är även linjärkombinationer av dessa två funktioner lösningar. Vidare kan man se att $y(x) = \sin(x)$ även uppfyller bivillkoren.

Övning 7.3.

(a) Funktionalen ges av

$$I : y(x) \mapsto \int_0^1 x^2 y(x) dx.$$

(b) Euler–Lagranges ekvation tillhörande funktionalen ovan ges av

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0,$$

där $f(x, y, y')$ är $x^2 y$.

Eftersom f ej beror av y' är $\partial f / \partial y' = 0$, och därmed är

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{d}{dx}(0) = 0.$$

Vidare är $\partial f / \partial y = x^2$. Euler–Lagranges ekvation ges sålunda av

$$-x^2 = 0.$$

(c) Eftersom $-x^2$ inte är identiskt 0 på intervallet $[a, b]$, och ekvationen $-x^2 = 0$ inte beror på $y(x)$ kan inte ekvationen uppfyllas för något val av y.

(d) Funktionalen har inget maximum. För att visa detta kommer vi att konstruera en följd av funktioner $y_n(x)$ så att $I(y_n) \to \infty$ då $n \to \infty$.

Det finns många sätt att göra detta på. Idén bakom följande konstruktion är att observera att integranden $x^2 y(x)$ består av två faktorer; x^2 och $y(x)$. Eftersom x^2 alltid är positiv på $[0, 1]$, kan vi helt enkelt välja $y_n(x)$ som alltid är positiva och antar större och större värden på $[0, 1]$, och på så sätt få större och större värden på $I(y)$.

Eftersom alla $y_n(x)$ måste uppfylla att $y_n(0) = y_n(1) = 0$ låter vi alla funktioner ha en faktor $x(1 - x)$. Notera att vi har $(1 - x)$ istället för $(x - 1)$ eftersom vi vill att funktionerna ska vara positiva.

För att få större och större värden på funktionen definierar vi slutligen

$$y_n(x) := nx(1 - x).$$

Direkt beräkning ger nu att

$$I(y_n) = \int_0^1 x^2 nx(1 - x) dx = n\left[\frac{x^4}{4} - \frac{x^5}{5}\right]_0^1 = \frac{n}{20},$$

varav det följer att I kan anta godtyckligt stora värden.

Övning 7.5. Funktionalen ges av

$$I : y(x) \mapsto \int_0^1 x + y(x) + y'(x)^2 dx.$$

Vidare är $\tilde{y}_\alpha(x) = -x + \alpha x^2$, och eftersom $\tilde{y}'_\alpha(x) = -1 + 2\alpha x$ är

$$\begin{aligned}
I_\eta(\alpha) &:= I(\tilde{y}_\alpha) = \int_0^1 x + \tilde{y}_\alpha(x) + \tilde{y}'_\alpha(x)^2 dx \\
&= \int_0^1 x + (-x + \alpha x^2) + (-1 + 2\alpha x)^2 dx = \int_0^1 \alpha x^2 + (1 - 4\alpha x + 4\alpha^2 x^2) dx \\
&= [(\alpha + 4\alpha^2)x^3/3 + x - 2\alpha x^2]_0^1 = (\alpha + 4\alpha^2)/3 + 1 - 2\alpha
\end{aligned}$$

Genom att derivera högerledet med avseende på α ser vi att

$$I'_\eta(\alpha) = \frac{1 + 8\alpha}{3} - 2,$$

och $I'_\eta(0) = -5/3$.

Övning 7.7. Enligt kedjeregeln är

$$\frac{\partial}{\partial \alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha) = \frac{\partial}{\partial \tilde{y}_\alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha) \frac{\partial \tilde{y}_\alpha}{\partial \alpha} + \frac{\partial}{\partial \tilde{y}'_\alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha) \frac{\partial \tilde{y}'_\alpha}{\partial \alpha}.$$

Eftersom

$$\frac{\partial}{\partial \alpha} \tilde{y}_\alpha(x) = \eta(x) y(x)$$

och

$$\frac{\partial}{\partial \alpha} \tilde{y}'_\alpha(x) = \frac{\partial}{\partial \alpha} (\alpha(\eta'(x)y(x) + \eta(x)y'(x))) = \eta'(x)y(x) + \eta(x)y'(x),$$

ger ekvationen ovan att

$$\frac{\partial}{\partial \alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha) = \eta(x)y(x) \frac{\partial}{\partial \tilde{y}_\alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha) + (\eta'(x)y(x) + \eta(x)y'(x)) \frac{\partial}{\partial \tilde{y}'_\alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha).$$

Om $\eta(x) = 1$ är $\eta'(x) = 0$, blir ekvationen ovan

$$\frac{\partial}{\partial \alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha) = y(x) \frac{\partial}{\partial \tilde{y}_\alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha) + y'(x) \frac{\partial}{\partial \tilde{y}'_\alpha} f(\tilde{y}_\alpha, \tilde{y}'_\alpha).$$

Kapitel 8

Övning 8.1.

$$\begin{aligned}
x(2\pi - \varphi) + x(\varphi) = r \cdot (2\pi - \varphi + \sin(\varphi)) + r \cdot (\varphi - \sin(\varphi)) = 2\pi r. \\
y(2\pi - \varphi) = r(1 - \cos(2\pi - \varphi)) = r(1 - \cos(\varphi)) = y(\varphi).
\end{aligned}$$

Övning 8.3.

(a) Insikten att denna integral är rätt är inte rätt, därför är del a) inte så mycket av en uppgift utan mer en ledtråd.

(b) Vi använder $1 + (y'(x))^2 = \frac{k^2}{y(x)} = \frac{2r}{y(x)}$ där $r = \frac{k^2}{2}$ enligt 8.1.5.

$$t = \frac{1}{\sqrt{2g}} \int_0^{x(t)} \sqrt{\frac{1 + y'(x)^2}{y(x)}} dx = \frac{1}{\sqrt{2g}} \int_0^{x(t)} \sqrt{\frac{2r}{y(x)^2}} dx = \sqrt{\frac{r}{g}} \int_0^{x(t)} \frac{1}{\sqrt{y(x)^2}} dx.$$

Svaret får vi genom att använda $y(x) \geqslant 0 \implies \sqrt{y(x)^2} = y(x)$.

(c) Vi börjar med att notera från 8.1.5 att $\frac{dx}{d\varphi} = r(1 - \cos(\varphi)) = y(\varphi)$. Då får vi att

$$t = \sqrt{\frac{r}{g}} \int_0^{x(t)} \frac{1}{\sqrt{y(x)^2}} dx = \left\{ \begin{smallmatrix} x = x(\varphi) & dx = \frac{dx}{d\varphi} d\varphi = y(\varphi) d\varphi \\ x = 0 \implies \varphi = 0 & x = x(t) \implies \varphi = \varphi(t) \end{smallmatrix} \right\} =$$

$$\sqrt{\frac{r}{g}} \int_0^{\varphi(t)} \frac{1}{y} y d\varphi = \sqrt{\frac{r}{g}} \int_0^{\varphi(t)} 1 d\varphi = \sqrt{\frac{r}{g}} \varphi(t).$$

Övning 8.5. Då vi vet att $\frac{\partial f}{\partial y} = 0$ förenklas Euler-Lagrange ekvationen till

$$\frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) = 0 \iff$$

$$\frac{\partial f}{\partial y'} = C$$

Referenser och förslag till vidare läsning

Arne Persson, Lars-Christer Böiers: *Analys i en variabel*
Studentlitteratur, 2010, Tredje upplagan

Arne Persson, Lars-Christer Böiers: *Analys i flera variabler*
Studentlitteratur, 2005, Tredje upplagan

Tomas Ekholm: *Envariabelanalys för Teknisk Fysik*
https://people.kth.se/ tomase/booklets/envariabelanalys.pdf

George F. Simmons, Steven G. Krantz: *Differential Equations: Theory, Technique, and Practice*
McGraw-Hill International Editions, 2006, Third Edition