

ANTECKNINGAR I  
MATEMATISK ANALYS I EN VARIABEL

AXEL MÅLQVIST

CHALMERS TEKNISKA HÖGSKOLA

*Utkast från 11 december 2017*

Copyright © 2017 Axel Målqvist

**Kopiering förbjuden**

Detta verk är skyddat av lagen om upphovsrätt. Ingen del av detta verk får reproduceras eller kopieras utan rättighetsinnehavarens skriftliga medgivande.

Art. No xxxxx

ISBN xxx-xx-xxxx-xx-x

Utgåva 0.0

Omslagsdesign av Anders Logg

Publicerad av Chalmers tekniska högskola  
Tryckt i Göteborg

# Innehåll

<b>1</b>	<b>System av ODE</b> .....	<b>2</b>
1.1	Introduktion till linjär algebra	2
1.2	System av första ordningens ODE	7
1.3	Högre ordningens ODE som system av första ordningen	10
1.4	Existens och entydighet av lösning	12
1.5	Alternativ definition av $e^x$ och $\ln(x)$	15
1.6	Alternativ definition av $\cos(x)$ och $\sin(x)$	18
<b>2</b>	<b>Numerisk lösning av ODE</b> .....	<b>27</b>
2.1	Numeriska metoder	27
2.2	Konvergens	30
2.3	Stabilitet	32
2.4	Algoritmer	34
2.5	Generalisering till system av ODE	37
2.6	Randvärdesproblem	40
<b>3</b>	<b>Laplacetransform</b> .....	<b>47</b>
3.1	Definition av Laplacetransform	47
3.2	Skalning	51
3.3	Laplacetransform av derivator och integraler	52
3.4	Faltning	54
3.5	Impulser och diskontinuerliga funktioner	56
3.6	Lösning av ODE med Laplacetransform	58
	<b>Litteratur</b> .....	<b>75</b>





# 1. System av ODE

1.1	Introduktion till linjär algebra	2
1.2	System av första ordningens ODE	7
1.3	Högre ordningens ODE som system av första ordningen	10
1.4	Existens och entydighet av lösning	12
1.5	Alternativ definition av $e^x$ och $\ln(x)$	15
1.6	Alternativ definition av $\cos(x)$ och $\sin(x)$	18

*System av ordinära differentialekvationer är vanligt förekommande inom mekanik, matematisk biologi, kemi och många andra discipliner. I detta kapitel introducerar vi matrisvektor notation för att enklare kunna hantera system av ODE. Vi visar att högre ordningens ODE kan skrivas om som system av första ordningen genom införandet av hjälpvariabler. Existens och entydighet av lösning för Lipschitzkontinuerliga högerled följer genom en generalisering av Picards sats. Vi använder existensresultatet för att ge alternativa definitioner till elementära funktioner som lösningar till differentialekvationer.*

## 1.1 Introduktion till linjär algebra

Vi börjar med en kort introduktion till linjär algebra för att kunna införa lämplig notation för system av ODE. Ett centralt begrepp inom linjär algebra är vektorer.

**Definition 1.1 (Vektornotation)** En vektor  $\vec{v} = [v_1 \ v_2 \ \dots \ v_n]$  är en uppsättning av  $n \in \mathbb{N}$  element  $v_i \in \mathbb{R}$ ,  $i = 1, \dots, n$ .

Vi definierar vektorer genom deras algebraiska egenskaper. Vektorer kan adderas och multipliceras med tal (skalärer). Givet två vektorer  $\vec{v} = [v_1 \ v_2 \ \dots \ v_n]$  och  $\vec{w} = [w_1 \ w_2 \ \dots \ w_n]$  med lika många element fås summan genom att summera komponentvis,

$$\vec{v} + \vec{w} = [v_1 + w_1 \ v_2 + w_2 \ \dots \ v_n + w_n].$$

Multiplikation mellan en skalär  $\alpha \in \mathbb{R}$  och en vektor  $\vec{v}$  fås genom att multiplicera varje komponent med skalären,

$$\alpha \vec{v} = [\alpha v_1 \ \alpha v_2 \ \dots \ \alpha v_n].$$

Vektorer kan även multipliceras med varandra förutsatt att de har lika många element. Skalärprodukten tar två vektorer som indata och ger ut ett tal, en skalär, som utdata.

**Definition 1.2 (Skalärprodukt och längd)** Låt  $\vec{v} = [v_1 \ v_2 \ \dots \ v_n]$  och  $\vec{w} = [w_1 \ w_2 \ \dots \ w_n]$  vara två vektorer. Skalärprodukten ges av  $\vec{v} \cdot \vec{w} = \sum_{i=1}^n v_i w_i$ . Längden ges av  $|\vec{v}| = (\vec{v} \cdot \vec{v})^{1/2} = \left(\sum_{i=1}^n v_i^2\right)^{1/2}$ .

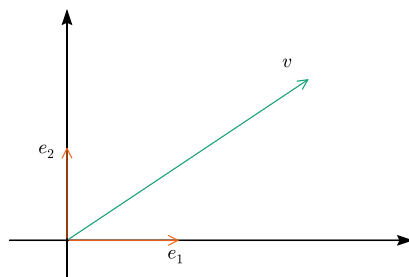
**Exempel 1.1** Beräkna skalärprodukten mellan vektorerna  $\vec{v} = [2 \ 1 \ -2]$  och  $\vec{w} = [0 \ -2 \ -2]$  samt längden av  $\vec{v}$ . Vi får  $\vec{v} \cdot \vec{w} = 2 \cdot 0 + 1 \cdot (-2) + (-2) \cdot (-2) = 2$ . Längden ges av  $|\vec{v}| = \sqrt{2^2 + 1^2 + (-2)^2} = 3$ .

Vi inför enhetsvektorerna  $\vec{e}_1 = [1 \ 0 \ 0 \ \dots \ 0]$ ,  $\vec{e}_2 = [0 \ 1 \ 0 \ 0 \ \dots \ 0]$  till och med  $\vec{e}_n = [0 \ 0 \ 0 \ \dots \ 0 \ 1]$  innehållande  $n$  element. Varje vektor  $\vec{v} = [v_1 \ v_2 \ \dots \ v_n]$  kan skrivas som en kombination av enhetsvektorerna multiplicerade med vektorns element på följande vis,

$$\vec{v} = \sum_{i=1}^n v_i \vec{e}_i.$$

Enhetsvektorerna har längd  $|\vec{e}_i| = 1$ . Skalärprodukten mellan enhetsvektorerna ges av  $\vec{e}_i \cdot \vec{e}_i = 1$  och  $\vec{e}_i \cdot \vec{e}_j = 0$  om  $i \neq j$ .

Vektorer har även en geometrisk tolkning som en riktad sträcka i  $\mathbb{R}^n$ . De brukar representeras som pilar. I Figur 1.1 ser vi en vektor  $\vec{v} \in \mathbb{R}^2$  tillsammans med enhetsvektorerna  $\vec{e}_1$  och  $\vec{e}_2$ . Vi noterar att vektorn  $\vec{v}$  representerar den riktade streckan mellan punkten  $(v_1, v_2, \dots, v_n) \in \mathbb{R}^n$  och origo. De geometriska tolkningarna av addition av vektorer och



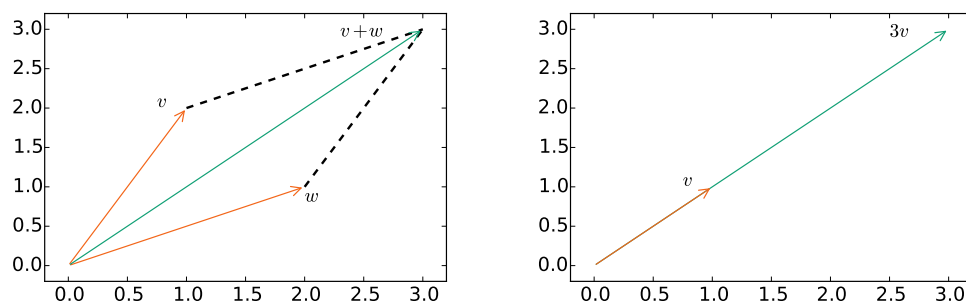
**Figur 1.1:** En vektor  $\vec{v} \in \mathbb{R}^2$  tillsammans med enhetsvektorerna  $\vec{e}_1$  och  $\vec{e}_2$ .

multiplikation med skalär illustreras i Figur 1.2

Även skalärprodukten har en geometrisk tolkning. Den geometriska skalärprodukten definieras som

$$v \cdot w = |v||w|\cos(\phi).$$

Skalärprodukten är som mest produkten av längderna och som minst minus produkten av längderna eftersom  $\cos(x)$  tar värden mellan  $-1$  och  $1$ . Om den mellanliggande vinkeln är  $\pi/2$  blir produkten noll. Vi säger då att vektorerna är vinkelräta eller ortogonala mot varandra. Detta gäller till exempel enhetsvektorerna. Om vinkeln är  $0$  blir skalärprodukten lika med produkten av vektorernas längder.

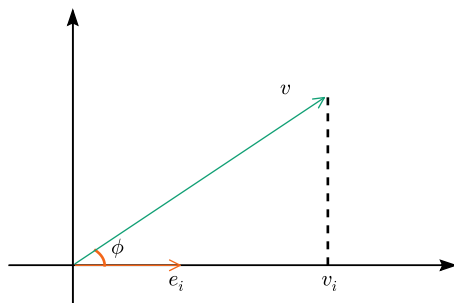


**Figur 1.2:** Geometrisk tolkning av summan av två vektorer  $[2 \ 1] + [1 \ 2] = [3 \ 3]$  (vänster) och multiplikation med skalär  $3[1 \ 1]=[3 \ 3]$ .

De två definitionerna av skalärprodukt är ekvivalenta. Vi kan se det med följande argument. Vi bildar den geometriska skalärprodukten mellan en vektor  $\vec{v} = [v_1 \ v_2 \ \dots \ v_n]$  och enhetsvektorn  $\vec{e}_i$ . Vi får,

$$\vec{v} \cdot \vec{e}_i = |\vec{v}| \cos(\phi) = v_i,$$

där  $\phi$  är vinkeln mellan  $\vec{v}$  och  $\vec{e}_i$ . I Figur 1.3 illustrerar vi detta resultat genom att välja koordinatsystemet så att den horisontella axeln går i riktningen av  $\vec{e}_i$ . Detta innebär att



**Figur 1.3:** En vektor  $\vec{v}$  i ett koordinatsystem där  $x$ -axeln väljs i linje med enhetsvektorn  $\vec{e}_i$ . Vinkeln mellan vektorerna ges av  $\phi$  och längden av  $\vec{v}$  vektorns komponent i riktning  $\vec{e}_i$  ges av talet  $v_i$ .

definitionerna är ekvivalenta för skalärprodukt mellan en godtycklig vektor och en enhetsvektor. Men en godtycklig vektor kan skrivas som en linjärkombination av enhetsvektorer  $\vec{w} = \sum_{i=1}^n w_i \vec{e}_i$  så därför gäller för att den geometriska definitionen leder till den algebraiska,

$$\vec{v} \cdot \vec{w} = \vec{v} \cdot \sum_{i=1}^n w_i \vec{e}_i = \sum_{i=1}^n w_i (\vec{v} \cdot \vec{e}_i) = \sum_{i=1}^n w_i v_i.$$

Vi går nu vidare till matriser.

**Definition 1.3 (Matrisnotation)** En  $n \times m$  matris  $A$ , med  $n$  rader och  $m$  kolumner, är ett rektangulärt talschema. Matrisen  $A$  med element  $a_{ij}$ ,  $1 \leq i \leq n$  och  $1 \leq j \leq m$



skrivs,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}. \quad (1.1)$$

Summation av två matriser och multiplikation med skalär görs elementvis på samma sätt som för vektorer,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2m} + b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{nm} + b_{nm} \end{bmatrix}$$

och

$$\alpha \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1m} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{n1} & \alpha a_{n2} & \dots & \alpha a_{nm} \end{bmatrix}$$

En vektor kan ses som ett specialfall av en matris med antingen endast en rad eller en kolumn.

**Definition 1.4 (Rad- och kolumnvektor)** En  $n \times 1$  matris kallas en kolumnvektor. En  $1 \times m$  matris kallas radvektor.

En kolumnvektor kan skrivas med följande notation  $\vec{v} = [v_1 \ v_2 \ \dots \ v_n]^T$ . Beteckningen  $T$  står för transponat och kan även appliceras på matriser enligt följande,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix}.$$

En matris kan multipliceras med en vektor om antalet kolumner i matrisen är samma som antal rader i vektorn.

**Definition 1.5 (Matris-vektor multiplikation)** Låt  $A$  vara  $n \times m$  matrisen,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}, \quad (1.2)$$

och låt  $\vec{x}$  vara  $m \times 1$  kolumnvektorn,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}, \quad (1.3)$$

Då ges produkten av  $A$  och  $\vec{x}$  som en  $n \times 1$  vektor där varje rad är skalärprodukten mellan motsvarande rad i  $A$  och vektorn  $\vec{x}$ , alltså,

$$A\vec{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m \end{bmatrix}. \quad (1.4)$$

**Exempel 1.2 (Matris-vektor multiplikation)** Beräkna följande matris-vektor produkt,

$$\begin{bmatrix} 1 & 4 & -2 \\ 0.5 & 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \cdot 3 + 4 \cdot 1 + (-2) \cdot 0 \\ 0.5 \cdot 3 + 3 \cdot 1 + 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 7 \\ 4.5 \end{bmatrix}. \quad (1.5)$$

Vi ser att produkten beräknas som skalärprodukt mellan rader i matrisen och kolumnvektorn.

Ett linjärt ekvationssystem är ett system av linjära ekvationer på formen

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \quad a_i \in \mathbb{R}, \quad i = 1, \dots, n.$$

Givet matrisnotationen kan vi uttrycka linjära ekvationssystem, och senare system av ordinära differentialekvationer, med en mer kompakt notation.

**Definition 1.6 (Linjära ekvationssystem)** Ett linjärt ekvationssystem med  $n$  ekvationer och  $n$  obekanta kan skrivas på formen,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \quad (1.6)$$

eller på matrisform som

$$A\vec{x} = \vec{b}, \quad (1.7)$$

där matrisen  $A$  har elementen  $a_{ij}$  och vektorerna  $\vec{x}$  och  $\vec{b}$  elementen  $x_i$  respektive  $b_i$   $1 \leq i, j \leq n$ .

**Exempel 1.3 (Linjärt ekvationssystem)** Skriv följande ekvationssystemet på matrisform:

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ x_1 - x_2 - x_3 = -1 \\ 2x_1 + x_2 - 3x_3 = 0 \end{cases} \quad (1.8)$$

Vi låter,

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 2 & 1 & -3 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}. \quad (1.9)$$

Det linjära ekvationssystemet kan då skrivas som

$$A\vec{x} = \vec{b}. \quad (1.10)$$

Linjära ekvationssystem kan lösas på ett systematiskt sätt med Gauss elimination. Har vi endast två eller tre obekanta kan vi lösa systemen för hand med substitution. Detta beskrivs enklast med ett exempel.

**Exempel 1.4 (Substitution)** Om vi vill lösa det linjära ekvationssystemet från föregående exempel kan vi till exempel bryta ut  $x_1 = -1 + x_2 + x_3$  ur den andra ekvationen och sätta in i den första och sista. Vi får då två ekvationer och två obekanta  $x_2$  och  $x_3$  nämligen,

$$-1 + x_2 + x_3 + x_2 + x_3 = 2x_2 + 2x_3 - 1 = 3, \quad (1.11)$$

$$-2 + 2x_2 + 2x_3 + x_2 - 3x_3 = 3x_2 - x_3 - 2 = 0, \quad (1.12)$$

Vi använder nu den andra ekvationen för att lösa ut  $x_3 = 3x_2 - 2$  och sätter in det i den första

$$2x_2 + 2x_3 = 8x_2 - 4 = 4$$

vilket ger  $x_2 = 1$ . Givet  $x_2 = 1$  har vi  $x_3 = 3x_2 - 2 = 1$  och  $x_1 = -1 + x_2 + x_3 = 1$ . Lösningen blir alltså  $x_1 = x_2 = x_3 = 1$ .

## 1.2 System av första ordningens ODE

Vi är nu redo att studera system av första ordningens ODE och vi börjar med de linjära fallet. För linjära system är matris-vektor notationen speciellt användbar. Både ekvation och begynnelsevillkoren kan uttryckas på ett kompakt sätt.

**Definition 1.7 (Linjärt system av första ordningens ODE)** Ett linjärt ekvationssystem med  $n$  första ordningens differentialekvationer och  $n$  obekanta funktioner  $y_i(x)$  kan

skrivs på formen,

$$\begin{cases} \frac{d}{dx}y_1(x) = a_{11}(x)y_1(x) + a_{12}(x)y_2(x) + \cdots + a_{1n}(x)y_n(x) + b_1(x) \\ \frac{d}{dx}y_2(x) = a_{21}(x)y_1(x) + a_{22}(x)y_2(x) + \cdots + a_{2n}(x)y_n(x) + b_2(x) \\ \vdots = \vdots \\ \frac{d}{dx}y_n(x) = a_{n1}(x)y_1(x) + a_{n2}(x)y_2(x) + \cdots + a_{nn}(x)y_n(x) + b_n(x) \end{cases} \quad (1.13)$$

med begynnelsevillkor  $y_j(0) = c_j$ ,  $j = 1, \dots, n$ . På matrisform får vi

$$\frac{d}{dx}\vec{y}(x) = A\vec{y}(x) + \vec{b}(x), \quad \vec{y}(0) = \vec{c} \quad (1.14)$$

där matrisen  $A$  har elementen  $a_{ij}(x)$  och vektorerna  $\vec{y}(x)$ ,  $\vec{b}(x)$  och  $\vec{c}$  har elementen  $y_i(x)$ ,  $b_i(x)$  respektive  $c_i$ , där  $1 \leq i, j \leq n$ .

Vi har  $n$  ekvationer och lika många obekanta  $y_i$ ,  $i = 1, \dots, n$ . Dessutom krävs  $n$  begynnelsevillkor för att bestämma en entydig lösning.

**Exempel 1.5** (Linjär ODE på matrisform) Skriv följande system av ordinära differentialekvationer på matrisform:

$$\begin{cases} \frac{d}{dx}y_1(x) = 2y_2(x), & y_1(0) = 0, \\ \frac{d}{dx}y_2(x) = -y_1(x), & y_2(0) = 1. \end{cases} \quad (1.15)$$

Vi låter,

$$A = \begin{bmatrix} 0 & 2 \\ -1 & 0 \end{bmatrix}, \quad \vec{y}(x) = \begin{bmatrix} y_1(x) \\ y_2(x) \end{bmatrix}, \quad \vec{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (1.16)$$

Det linjära ekvationssystemet kan då skrivas som

$$\frac{d}{dx}\vec{y}(x) = A\vec{y}(x), \quad \vec{y}(0) = \vec{c} \quad (1.17)$$

På samma sätt som för linjära ekvationssystem kan substitution användas för att lösa system av linjära ODE.

**Exempel 1.6** Lös systemet av ordinära differentialekvationer med substitution:

$$\begin{cases} \frac{d}{dx}y_1(x) = 2y_2(x), & y_1(0) = 0, \\ \frac{d}{dx}y_2(x) = -y_1(x), & y_2(0) = 1. \end{cases} \quad (1.18)$$

Vi låter  $y_2(x) = \frac{1}{2} \frac{d}{dx}y_1(x)$  i andra ekvationen och får,

$$\frac{d^2}{dx^2}y_1(x) + 2y_1(x) = 0. \quad (1.19)$$

Den karakteristiska ekvationen  $r^2 + 2 = 0$  har rötter  $r_{1,2} = \pm i\sqrt{2}$ . Lösningen ges av

$$y_1(x) = A \cos(x\sqrt{2}) + B \sin(x\sqrt{2}). \quad (1.20)$$

Begynnelsevillkoren ger  $y_1(0) = A = 0$  och  $y_2(0) = \frac{1}{2} \frac{d}{dx} y_1(0) = \frac{B}{\sqrt{2}} = 1$  vilket ger  $y_1(x) = \sqrt{2} \sin(x\sqrt{2})$  och  $y_2(x) = \cos(x\sqrt{2})$ .

Vi går nu till en mer generell beskrivning av system av första ordningens ODE. Vi använder vektornotationen för att uttrycka ett allmänt system på kompakt form.

**Definition 1.8** (Allmänt system av första ordningens ODE) Ett allmänt ekvationssystem med  $n$  första ordningens differentialekvationer och  $n$  obekanta funktioner  $y_i(t)$  kan skrivas på formen,

$$\begin{cases} \frac{d}{dx} y_1(x) = f_1(x, y_1(x), \dots, y_n(x)) \\ \frac{d}{dx} y_2(x) = f_2(x, y_1(x), \dots, y_n(x)) \\ \vdots = \vdots \\ \frac{d}{dx} y_n(x) = f_n(x, y_1(x), \dots, y_n(x)) \end{cases} \quad (1.21)$$

med begynnelsevillkor  $y_j(0) = c_j$ ,  $j = 1, \dots, n$ . På vektorform får vi

$$\frac{d}{dx} \vec{y}(x) = \vec{f}(x, \vec{y}), \quad \vec{y}(0) = \vec{c} \quad (1.22)$$

där vektorn  $\vec{f}$  har elementen  $f_i$  och vektorerna  $\vec{y}$  och  $\vec{c}$  har elementen  $y_i$  respektive  $c_i$ , där  $1 \leq i \leq n$ .

Vi har fortfarande  $n$  ekvationer,  $n$  obekanta samt  $n$  begynnelsevillkor. Vektornotationen kan alltså även användas för icke-linjära system av ODE. Vi fortsätter med ett klassiskt exempel på ett icke-linjärt system av ODE som modellerar dynamiken i ett slutet system av predatorer och bytesdjur.

**Exempel 1.7** (Rävar och kaniner) En enkel modell för dynamiken i djurpopulationer är Volterra-Lotkas ekvation som beskriver populationen av predatorer och bytesdjur i ett slutet system. Rävarna (predatorerna) äter endast kaninerna (bytesdjuren) medan kaninerna har gott om mat. Vi låter antal rävar som funktion av tiden ges av  $y(t)$  och antal kaniner av  $x(t)$ . Lotka-Volterras modell ger följande icke-linjära första ordningens system av ODE,

$$\frac{dx(t)}{dt} = x(t)(a - by(t)), \quad x(0) = x_0 \quad (1.23)$$

$$\frac{dy(t)}{dt} = -y(t)(c - dx(t)), \quad y(0) = y_0 \quad (1.24)$$

där termen  $ax(t)$  anger att populationen kaniner ökar exponentiellt i avsaknad av rävar. Det finns alltså obegränsat med mat för kaninerna. Termen  $-bx(t)y(t)$  påverkar kaninbeståndet negativt och mäter effekten av att rävar och kaniner träffar på varandra. Motsvarande term i rävarnas ekvation  $dx(t)y(t)$  verkar på motsatt sätt positivt på rävbekäftet. Slutligen avtar rävbekäftet exponentiellt i avsaknad av

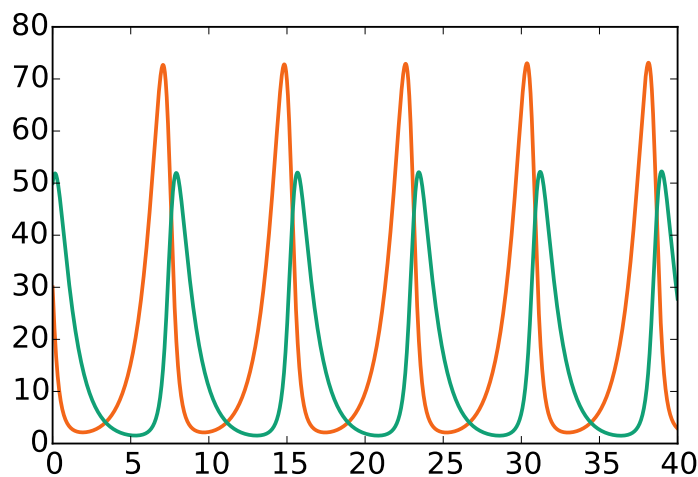
kaniner  $-cy(t)$ . Vi kan skriva systemet med vektornotation genom att införa

$$\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \vec{f} = \begin{bmatrix} x(a-by) \\ -y(c-dx) \end{bmatrix}$$

och  $\vec{v}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ . På vektorform får vi

$$\frac{d}{dt}\vec{v}(t) = \vec{f}(t, \vec{v}(t)), \quad \vec{v}(0) = \vec{v}_0.$$

I Figur 1.4 visar vi ett exempel på hur lösningen kan se ut. Vi börjar med 30 kaniner (orange) och 50 rävar (grön) och låter  $a = 1$ ,  $b = 0.07$ ,  $c = 1$  och  $d = 0.05$ . Sluttiden sätts till  $T = 40$ . Vi ser att när det finns ont om kaniner avtar rävarna i antal men då får kaninerna möjlighet att öka igen vilket gör att rävarna får mer mat och börjar öka igen. Lösningarna är periodiska. Det finns två jämviktspunkter när populationerna inte ändras mer. Det kan fås fram genom att sätta  $\frac{dx}{dt} = \frac{dy}{dt} = 0$ , alltså  $x(a-by) = 0$  och  $-y(c-dx) = 0$  med lösningar  $x = 0$  och  $y = 0$  eller  $x = \frac{c}{d} = 20$  och  $y = \frac{a}{b} = \frac{100}{7}$ . Notera att vi tillåter djurantalet att ej vara heltal.



Figur 1.4: Antal rävar (grön) och kaniner (orange) som funktion av tiden.

### 1.3 Högre ordningens ODE som system av första ordningen

I tillämpningar är systemen av ODE inte alltid av första ordningen. Planeterna i vårt solsystems positioner i rummet som funktion av tiden är ett exempel på system av andra ordningens ODE. Vi kan dock enkelt skriva om ett högre ordningens system till ett först ordningens system. Vi formulerar denna teknik i följande sats.

**Sats 1.1 (Högre ordningens ODE som system av första ordningen)** Låt  $F : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$  vara en given funktion och  $\{c_i\}_{i=0}^{n-1}$  en uppsättning reella tal. En  $n$ :te ordningens ODE,

$$y^{(n)} = F(x, y, y^{(1)}, y^{(2)}, \dots, y^{(n-1)}), \quad y^i(0) = c_i, \quad i = 0, \dots, n-1, \quad (1.25)$$

kan skrivas som ett system av första ordningens ODE.

*Bevis.* Låt  $\vec{y}$  vara en  $n$  kolumnvektor med komponenter  $[y_1 \ y_2 \ \dots \ y_n]$ . Vidare låt  $y_i = y^{(i-1)}$ ,  $i = 1, \dots, n$ , där  $y^{(0)} = y$ . Då kan ekvation (1.25) skrivas på formen,

$$\begin{cases} \frac{d}{dx} y_1(x) = & y_2(x) \\ \frac{d}{dx} y_2(x) = & y_3(x) \\ \vdots & = \quad \quad \quad \vdots \\ \frac{d}{dx} y_n(x) = & F(x, y_1(x), \dots, y_n(x)), \end{cases} \quad (1.26)$$

vilket är ett system av första ordningen. Bergynnevillkoren ges av  $y_i(0) = c_{i+1}$ ,  $i = 1, \dots, n$ .  $\square$

**Exempel 1.8 (Vågutbredning)** Vi studerar longitudinell vågutbredning i ett system där två massor sammanlänkade i fasta ytterväggar med tre fjädrar. Vi låter fjädrarna ha fjäderkonstant  $k$  och massorna massa  $m$ . Vi ställer upp ett system av första ordningens ODE vars lösning är avvikelserna från jämviktslägena  $x_1(t)$  och  $x_2(t)$  för de två massorna.

Newtons andra lag ( $F = m\ddot{x}$ , kraften är lika med massan gånger accelerationen) och Hookes lag ( $F = kx$ , kraften är proportionell mot förskjutning från jämviktsläget  $x$ ) ger två ekvationer som beskriver kraftjämvikt för de två massorna

$$m\ddot{x}_1(t) = -kx_1(t) + k(x_2(t) - x_1(t)), \quad (1.27)$$

$$m\ddot{x}_2(t) = -k(x_2(t) - x_1(t)) - kx_2(t). \quad (1.28)$$

Vi inför nu två nya variabler  $x_3(t) = \dot{x}_1(t)$  och  $x_4(t) = \dot{x}_2(t)$  och får följande system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2k/m & k/m & 0 & 0 \\ k/m & -2k/m & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}. \quad (1.29)$$

med begynnelsevärden  $x_i(t) = c_i$  för  $i = 1, 2, 3, 4$ . Vi kan alltså skriva om problemet som ett system av första ordningens ODE.

En stor mängd matematiska problem kan beskrivas som system av första ordningens ordinära differentialekvationer, inte minst på grund av att även högre ordningens ODE kan skrivas som system av första ordningen. En viktig fråga som då behöver besvaras är under vilka förutsättningar ett system av första ordningens ODE har entydig lösning. I det skalära fallet ger Picards sats svaret på den frågan. Vi ska nu studera hur Picards sats kan generaliseras till system av ODE.

## 1.4 Existens och entydighet av lösning

På samma sätt som för skalära ODE av första ordningen kan vi garantera existens och entydighet av lösning under vissa antaganden på funktionen  $\vec{f}$ . Vi börjar med att definiera funktionsrummet  $C([a, b])^n$  av funktioner som tar värden i  $\mathbb{R}^n$ , så kallade vektorvärda funktioner.

**Definition 1.9 (Kontinuerliga funktioner i  $\mathbb{R}^n$ )** En vektorvärd funktion av en variabel  $\vec{y} : \mathbb{R} \rightarrow \mathbb{R}^n$  är kontinuerlig på intervallet  $[a, b]$  om samtliga komponenter  $y_i \in C([a, b])$ ,  $i = 1, \dots, n$ , är kontinuerliga. Vi skriver  $\vec{y} \in C([a, b])^n$ .

Vi behöver nu definiera normen (storleken) av en funktion i  $C([a, b])^n$ . Skillnaden från det skalära fallet är att längden på vektorn ersätter absolutbeloppet. Notationen är dock den samma.

**Definition 1.10 (Norm av vektorvärdfunktion)** Vi definierar normen (storleken) på en kontinuerlig vektorvärd funktion  $\vec{f} \in C([a, b])^n$  som  $\|\vec{f}\| = \max_{x \in [a, b]} |\vec{f}(x)|$ .

Banachs fixpunktsats går att generalisera till vektorvärda funktioner.

**Sats 1.2** ★ [Banach's fixpunktssats för vektorvärda funktioner] Givet  $\vec{y}_0 \in C([a, b])^n$  och  $\alpha > 0$  låt  $S = \{\vec{y} \in C([a, b])^n : \|\vec{y} - \vec{y}_0\| \leq r\}$ . Låt vidare  $G$  vara en avbildning från  $S$  in i sig själv,  $G : S \rightarrow S$ . Om  $G$  är en kontraktion, alltså om för något  $0 \leq \gamma < 1$  gäller

$$\|G(\vec{y}) - G(\vec{z})\| \leq \gamma \|\vec{y} - \vec{z}\|, \quad \forall \vec{y}, \vec{z} \in S,$$

så har  $G$  en unik fixpunkt  $\vec{x} = G(\vec{x})$ ,  $\vec{x} \in S$ . Dessutom konvergerar fixpunktsiterationen  $\vec{y}_i = G(\vec{y}_{i-1})$  givet  $\vec{y}_0 \in S$ , det vill säga det finns  $\vec{y} \in S$  sådant att  $\|\vec{y} - \vec{y}_i\| \rightarrow 0$  då  $i \rightarrow \infty$ .

*Bevis.* Beviset liknar Banachs fixpunktsats och bygger på fullständigheten hos  $C([a, b])^n$  istället för hos de reella talen.  $\square$

Figur 1.5 illustrerar mängden  $S$  i fallet  $n = 1$ . Kurvan är  $\vec{y}_0$  och streckade linjerna motsvarar det band med avstånd  $r$  från kurvan där funktionerna i  $S$  får ligga om de ska uppfylla  $\|\vec{y} - \vec{y}_0\| \leq r$ .

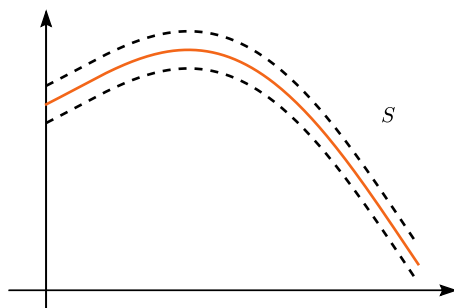
Vi formulerar och bevisar nu satsen om existens och entydighet av lösning som är en generalisering av Picards sats. Satsen ger existens och entydighet i en omgivning till begynnelsevillkoret  $x_0 - \delta \leq x \leq x_0 + \delta$  precis som Picards sats.

**Sats 1.3 (Existens och entydighet)** Vi studerar begynnelsevärdesproblemet

$$\begin{cases} \frac{d\vec{y}}{dx} = \vec{f}(x, \vec{y}), \\ \vec{y}(x_0) = \vec{y}_0, \end{cases} \quad (1.30)$$

där  $\vec{y}(x), \vec{y}_0 \in \mathbb{R}^n$  för alla  $x \in \mathbb{R}$  och  $\vec{f} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ . Låt vidare  $R = \{\vec{v} \in \mathbb{R}^n : a_i \leq v_i \leq b_i, \quad i = 1, \dots, n\}$  vara en generaliserad rektangel i  $\mathbb{R}^n$  sådan att  $\vec{y}_0 \in R$ . Om  $\vec{f}$  är kontinuerlig i  $x$ -variabeln på intervallet  $[a_0, b_0]$  innehållande  $x_0$  och lokalt Lipschitz





**Figur 1.5:** De streckade linjerna illustrerar det avstånd från en given kurva  $\vec{y}_0$  som funktionerna i  $S$  ligger inom.

kontinuerlig i  $\vec{y}$ -variabeln, det vill säga

$$\|\vec{f}(x, \vec{y}_1) - \vec{f}(x, \vec{y}_2)\| \leq K\|\vec{y}_1 - \vec{y}_2\|, \quad \vec{y}_1, \vec{y}_2 \in R, \quad a_0 \leq x \leq b_0,$$

så existerar entydig lösning  $\vec{y}(x)$  till begynnelsevärdesproblemet i intervallet  $x_0 - \delta \leq x \leq x_0 + \delta$  för något  $\delta > 0$ .

*Bevis.* ★ Beviset är analogt med Picards sats för skalära ODE. Vi repeterar beviset i kortet. Vi integrerar (1.32) från  $x_0$  till  $x$ ,

$$\vec{y}(x) = \vec{y}_0 + \int_{x_0}^x \vec{f}(t, \vec{y}(t)) dt. \quad (1.31)$$

Om ekvation (1.31) har en lösning  $\vec{y}(x)$  så är den deriverbar och löser (1.32). Fixpunktsiteration ger en avbildning  $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,

$$\vec{\phi}_i := G(\vec{\phi}_{i-1}) = \vec{y}_0 + \int_{x_0}^x \vec{f}(t, \vec{\phi}_{i-1}(t)) dt, \quad i = 1, \dots$$

Eftersom  $\vec{f}$  är kontinuerlig på  $R$  gäller att  $|\vec{f}(x, \vec{y})| \leq M$ , för  $x \in [a_0, b_0]$  och  $\vec{y} \in R$ . Vi låter  $I = [x_0 - \delta, x_0 + \delta] \subset [a_0, b_0]$  för något  $\delta > 0$  och bildar funktionsmängden  $S = \{\vec{\phi} \in C(I)^n : \|\vec{\phi}(x) - \vec{y}_0\| \leq r\}$ , med  $r = M\delta$ . Vi ser att för  $\vec{v} \in S$  gäller att,

$$\|G(\vec{v}) - \vec{y}_0\| \leq \max_{x \in I} \left| \int_{x_0}^x \vec{f}(t, \vec{v}(t)) dt \right| \leq M\delta.$$

Alltså har vi att  $G: S \rightarrow S$ . Vidare gäller för  $\vec{v}, \vec{w} \in S$  att,

$$\|G(\vec{v}) - G(\vec{w})\| \leq \max_{x \in I} \left| \int_{x_0}^x (\vec{f}(t, \vec{v}(t)) - \vec{f}(t, \vec{w}(t))) dt \right| \leq \delta K \|\vec{v} - \vec{w}\|.$$

Om  $\delta K < 1$  är  $G$  en kontraktion och Banachs fixpunktsats 1.2 garanterar en unik lösning  $\vec{y} \in S$  i intervallet  $x_0 - \delta \leq x \leq x_0 + \delta$ .  $\square$

Under starkare antaganden, nämligen global kontinuitet i  $x$  och Lipschitz kontinuitet i  $\vec{y}$ , visar vi även existens och entydighet för alla  $x \in \mathbb{R}$ . Resultatet håller naturligtvis även i specialfallet när  $n = 1$ .

**Sats 1.4 (Existens och entydighet av global lösning)** Vi studerar begynnelsevärdesproblemet

$$\begin{cases} \frac{d\vec{y}}{dx} = \vec{f}(x, \vec{y}), \\ \vec{y}(x_0) = \vec{y}_0, \end{cases} \quad (1.32)$$

där  $\vec{y}(x), \vec{y}_0 \in \mathbb{R}^n$  för alla  $x \in \mathbb{R}$  och  $\vec{f} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ . Om  $\vec{f}$  är kontinuerlig i  $x$  variabeln och globalt Lipschitzkontinuerlig i  $\vec{y}$ ,

$$|\vec{f}(x, \vec{y}_1) - \vec{f}(x, \vec{y}_2)| \leq K|\vec{y}_1 - \vec{y}_2|, \quad (x, \vec{y}_1), (x, \vec{y}_2) \in \mathbb{R}^{n+1},$$

existerar det en entydig lösning  $\vec{y}(x)$  till begynnelsevärdesproblemet för alla  $x \in \mathbb{R}$ .

*Bevis.* ★ Beviset är likt beviset av Sats 1.3 men innehåller några skillnader. Vi låter igen  $I = [x_0 - \delta, x_0 + \delta]$  och bildar  $S = \{\vec{y} \in C(I)^n : \|\vec{y} - \vec{y}_0\| \leq r\}$  för något  $\delta > 0$  och  $r > 0$ . Fixpunktsiteration ger upphov till avbildningen  $G$ ,

$$G(\vec{v}) := \vec{y}_0 + \int_{x_0}^x \vec{f}(t, \vec{v}(t)) dt.$$

Vi vill välja  $r$  och  $\delta$  så att  $G : S \rightarrow S$  och att  $G$  är en kontraktion. Vi noterar att eftersom  $\vec{f}$  är kontinuerlig på  $I$  och därmed bergänsad (av något tal  $M > 0$ ) i  $x$ -variabeln och (globalt) Lipschitz kontinuerlig i  $\vec{y}$ -variabeln gäller att för  $x \in I$  och  $\vec{y} \in S$

$$|\vec{f}(x, \vec{y})| \leq |\vec{f}(x, \vec{y}_0)| + |\vec{f}(x, \vec{y}) - \vec{f}(x, \vec{y}_0)| \leq M + K|\vec{y} - \vec{y}_0| \leq M + K \cdot r.$$

Vi får därför att,

$$\|G(\vec{y}) - \vec{y}_0\| = \max_{x \in I} |G(\vec{y})(x) - \vec{y}_0| = \max_{x \in I} \left| \int_{x_0}^x \vec{f}(t, \vec{y}(t)) dt \right| \leq \max_{x \in I} |x - x_0| (M + Kr) \leq \delta(M + Kr).$$

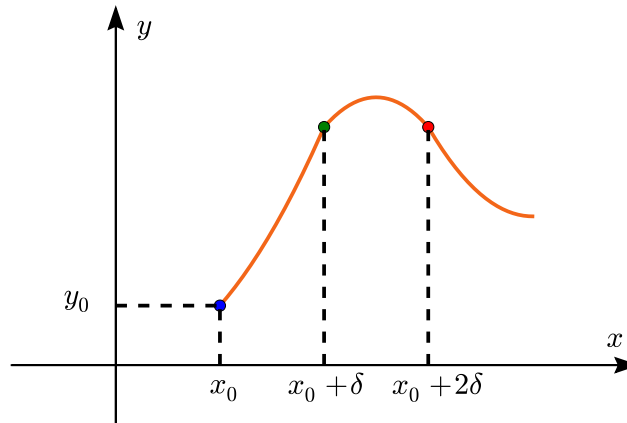
Om  $G(\vec{y}) \in S$  måste därför  $\delta(M + Kr) \leq r$ . Vi låter  $\delta = \frac{1}{2K}$  och  $r = \frac{M}{K}$ . Med dessa val gäller att  $G : S \rightarrow S$ . Dessutom gäller att,

$$\|G(\vec{v}) - G(\vec{w})\| \leq \max_{x \in I} \left| \int_{x_0}^x \vec{f}(t, \vec{v}(t)) dt - \int_{x_0}^x \vec{f}(t, \vec{w}(t)) dt \right| \leq \max_{t \in I} \delta K |\vec{v}(t) - \vec{w}(t)| = \frac{1}{2} \|\vec{v} - \vec{w}\|.$$

Banachs fixpunktssats ger att det finns en unik lösning  $\vec{y} \in S$  till begynnelsevärdesproblemet i intervallet  $x_0 - \delta \leq x \leq x_0 + \delta$ . Eftersom  $\delta$  bara beror på  $K$  som är en global Lipschitz konstant kan vi göra om samma argument med begynnelsevillkor i  $(x_0 + \delta, y(x_0 + \delta))$ . Samma resonemang ger då unik lösning i intervallet  $[x_0, x_0 + 2\delta]$ . Detta argument kan sedan upprepas till vi har existens av lösning för alla  $x \in \mathbb{R}$ .  $\square$

Sista steget i beviset där vi går från lokal lösning i  $[x_0 - \delta, x_0 + \delta]$  till global lösning i hela  $\mathbb{R}$  är intressant. När Lipschitzkonstanten är global ser vi att  $\delta$  kan väljas oberoende av

begynnelsevillkor till  $\delta = \frac{1}{2K}$ . Det innebär att vi kan börja om på  $x_0 + \delta$  med  $y(x_0 + \delta)$ , som vi vet existerar, som begynnelsevillkor. Detta ger existens och entydighet även i intervallet  $[x_0, x_0 + 2\delta]$  och så vidare. I Figur 1.6 ser vi hur lösningarna fogas samman till en global lösning. Om Lipschitz konstanten inte var global utan till exempel ökade med ökande  $x$  så skulle intervallens längd  $\delta$  kunna avta och till slut konvergera mot ett ändligt värde. Det är precis det som händer i exemplet med ekvationen  $y' = y^2$  med begynnelsevillkor  $y(0) = 1$ . För det problemet existerar bara lösningen på intervallet  $[0, 1)$ .



**Figur 1.6:** En global lösning sammanfogas av lokala lösningar på delintervall.

Det gäller att linjära ODE med kontinuerliga koefficienter är ett specialfall av globalt Lipschitzkontinuerliga högerled som omfattas av satsen. Sats 1.4 ger alltså existens och entydighet av lösning för linjära system av ODE för alla  $x \in \mathbb{R}$ . Eftersom högre ordningens ODE kan skrivas om som system av första ordningen är resultaten mycket användbara. Sats 1.4 kan till exempel även användas för att definiera elementära funktioner som  $e^x$ ,  $\cos(x)$ ,  $\sin(x)$  och mer exotiska funktioner. Vi studerar nu detta närmare.

## 1.5 Alternativ definition av $e^x$ och $\ln(x)$

I och med att vi har visat att ett system av första ordningens ODE med högerled som är globalt Lipschitz kontinuerligt har unik lösning kan vi presentera alternativa definitioner av de elementära funktionerna som lösningar till differentialekvationer. Om en elementär funktion löser ett system av ODE med Lipschitzkontinuerligt högerled kan vi använda detta som definition för funktionen. Sedan återstår att visa funktionernas egenskaper utifrån den nya definitionen. Eftersom de elementära funktionerna beskriver fysikaliska fenomen som populationstillväxt och svängningar visar det sig att de löser väldigt enkla och fundamentala differentialekvationer. Vi börjar med  $e^x$  och dess invers  $\ln(x)$ .

Ett sätt att definiera talet  $e$  är följande gränsvärde,

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

Logaritmen kan med hjälp av integralen definieras som,

$$\ln(x) = \int_1^x \frac{1}{y} dy.$$

Om vi istället vill definiera  $e^x$  som lösning till en differentialekvation får vi följande definition.

**Definition 1.11 (Alternativ definition av exponentialfunktionen)** Vi definierar  $y(x) = e^x$  som den entydiga lösningen till ekvationen,

$$y'(x) = y(x),$$

med begynnelsevillkor  $y(0) = 1$ .

Existens och entydighet följer direkt av Sats 1.4 med  $f(x, y) = y$  som är globalt Lipschitzkontinuerlig. Ekvationen som definierar exponentialfunktionen modellerar ett system där förändringen (derivatan) är proportionell mot populationens storlek.

Vi kan utgående från denna definition härleda exponentialfunktionens egenskaper och gränsvärdesdefinitionen av talet  $e$ . Vi har direkt att exponentialfunktionen är sin egen derivata eftersom  $y'(x) = y(x)$  per definition. I följande exempel härleder vi två av exponentialfunktionens egenskaper.

**Exempel 1.9** Vi vill bestämma vilken ekvation som  $y(kx) = e^{kx}$  löser, där  $k \in \mathbb{R}$ . Vi deriverar  $y(kx)$  och får

$$y(kx)' = y'(kx) \cdot \frac{dkx}{dx} = ky(kx), \quad x \in \mathbb{R}, \quad y(k \cdot 0) = y(0) = 1.$$

Alltså är  $e^{kx}$  den entydiga lösningen till ekvationen  $y'(x) = ky(x)$  med begynnelsevillkoret  $y(0) = 1$ .

**Exempel 1.10 (Produkt av exponentialfunktioner)** Låt  $y_1(x)$  lösa  $y_1'(x) = k_1 y_1(x)$  med begynnelsevillkoret  $y_1(0) = 1$  och  $y_2(x)$  lösa  $y_2'(x) = k_2 y_2(x)$  med begynnelsevillkor  $y_2(0) = 1$ . Per definition är alltså  $y_1(x) = e^{k_1 x}$  och  $y_2(x) = e^{k_2 x}$ . Då gäller att  $y_3 = y_1 \cdot y_2$  löser

$$y_3'(x) = y_1'(x)y_2(x) + y_1(x)y_2'(x) = (k_1 + k_2)y_1(x) \cdot y_2(x) = (k_1 + k_2)y_3(x),$$

med begynnelsevillkoret  $y_3(0) = y_1(0) \cdot y_2(0) = 1$ . Denna ekvation har per definition lösningen  $y_3(x) = e^{(k_1+k_2)x}$ . Alltså har vi härlett egenskapen  $e^{k_1 x} \cdot e^{k_2 x} = e^{(k_1+k_2)x}$ .

Vi kan även härleda gränsvärdes definitionen för talet  $e$  genom att använda en numerisk metod, som vi återkommer till i kapitel 2, för att approximera lösningen.

**Exempel 1.11 (Härledning av talet  $e$ )** Ekvationen,

$$y'(x) = y(x), \quad y(0) = 1,$$

Vi vill beräkna en approximation till  $y(1)$  genom att dela in intervallet  $[0, 1]$  i  $n$  delin-

tervall av längd  $\frac{1}{n}$ . För att hitta en approximativ lösning byter vi ut derivatan mot en differenskvot  $y'(x) \approx \frac{y_{i+1} - y_i}{1/n}$  som konvergerar mot derivatan när  $n \rightarrow \infty$  och så låter vi högerledet  $y(x)$  approximeras av  $y_i$  på intervallet  $[x_i, x_{i+1}]$ , alltså

$$\frac{y_{i+1} - y_i}{1/n} = y_i,$$

eller

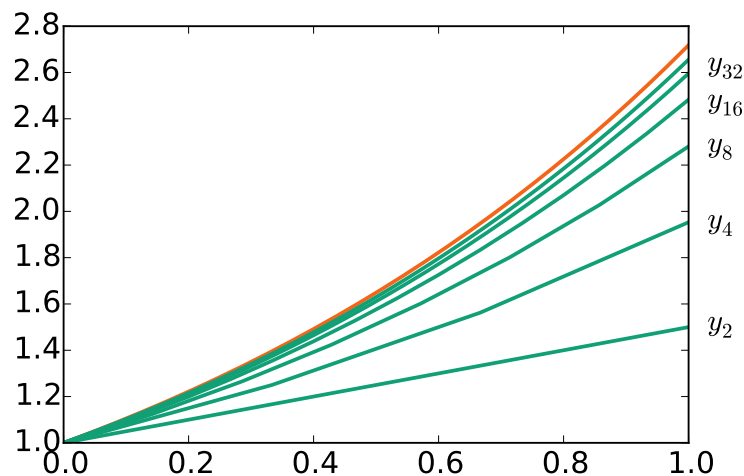
$$y_n = y_{n-1} + \frac{1}{n}y_{n-1} = \left(1 + \frac{1}{n}\right)y_{n-1} = \cdots = \left(1 + \frac{1}{n}\right)^n y_0 = \left(1 + \frac{1}{n}\right)^n.$$

När vi låter  $n \rightarrow \infty$  får vi  $y_\infty = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$ .

Den numeriska metod vi använder här kallas Eulers metod. Vi visar i Sats 2.1 att den beräknade lösningen konvergerar mot den exakta. Därför gäller att

$$e = y(1) = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

I Figur 1.7 ser vi hur lösningarna konvergerar mot den exakta lösningen då antal delintervall  $n$  ökar  $n = 2, 4, 8, 16, 32, 64$ .



**Figur 1.7:** Konvergens av numeriska approximationer av exponentialfunktionen för  $n = 1, 4, 8, 16, 32, 64$ .

I nästa exempel använder vi definitionen av exponentialfunktionen för att härleda definitionen av logaritmen.

**Exempel 1.12 (Logaritmen)** Vi vill definiera den naturliga logaritmen som inversen till exponentialfunktionen. Vi har  $y'(x) = y(x)$  samt  $y(0) = 1$  och vill bestämma  $v(y)$  så att  $x = v(y(x))$ . Vi deriverar uttrycket med avseende på  $x$  och använder kedjeregeln,

$$1 = \frac{dx}{dx} = \frac{dv(y(x))}{dx} = \frac{dv(y)}{dy} \frac{dy}{dx} = v'(y) \cdot y.$$

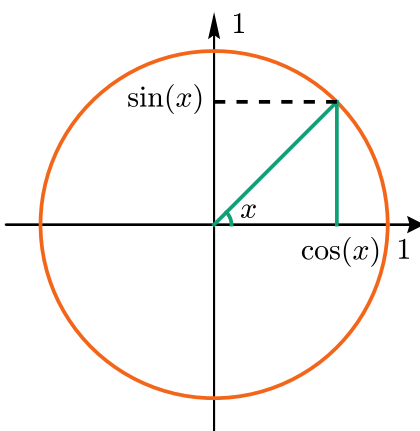
Alltså gäller  $v'(y) = \frac{1}{y}$  samt  $v(1) = v(y(0)) = 0$ . Integration från 1 till  $x$  ger,

$$v(x) - v(1) = \int_1^x \frac{1}{y} dy.$$

Inversen till exponentialfunktionen ges därmed av  $\ln(x) = \int_1^x \frac{1}{y} dy$ .

## 1.6 Alternativ definition av $\cos(x)$ och $\sin(x)$

Vi går nu över till att definiera  $\cos(x)$  och  $\sin(x)$ . Traditionellt definieras  $\cos(x)$  och  $\sin(x)$  med hjälp av enhetscirkeln, som i Figur 1.8. För varje punkt på cirkeln ges av koordinater  $(\cos(x), \sin(x))$ , där  $x$  är vinkeln mellan den horisontella axeln och den linje som passerar genom  $(\cos(x), \sin(x))$  och origo.



**Figur 1.8:** Definition av  $\cos(x)$  och  $\sin(x)$  som närliggande respektive motstående katet genom hypotenusan (lika med ett) i en rätvinklig triangel.

På samma sätt som för  $e^x$  kan vi istället definiera  $\cos(x)$  och  $\sin(x)$  som lösningar till en differentialekvation.

**Definition 1.12** (Alternativ definition av cosinus) Vi definierar  $y(x) = \cos(x)$  som den entydiga lösningen till ekvationen,

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

med begynnelsevillkor  $y(0) = 1$  och  $y'(0) = 0$ .

**Definition 1.13** (Alternativ definition av sinus) Vi definierar  $z(x) = \sin(x)$  som den

entydiga lösningen till ekvationen,

$$z''(x) + z(x) = 0,$$

med begynnelsevillkor  $z(0) = 0$  och  $z'(0) = 1$ .

Entydig lösning existerar till båda ekvationerna på grund av Sats 1.4 eftersom vi kan skriva om ekvationerna som system av första ordningen med ett Lipschitzkontinuerligt högerled. Låt  $u_1 = y$  och  $u_2 = \frac{dy}{dx}$ . Vi får,

$$\frac{d}{dx} \vec{u}(x) = \begin{bmatrix} u_1'(x) \\ u_2'(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix} = A\vec{u}. \quad (1.33)$$

Högerledet är linjärt i  $\vec{u}$ . Vi har att

$$|A\vec{v} - A\vec{w}| = \left| \begin{bmatrix} v_2 - w_2 \\ w_1 - v_1 \end{bmatrix} \right| = \sqrt{(v_2 - w_2)^2 + (v_1 - w_1)^2} = |\vec{v} - \vec{w}|.$$

Alltså är högerledet (globalt) Lipschitzkontinuerligt med konstant  $L = 1$ . Sats 1.4 ger existens och entydighet av lösning för alla  $x \in \mathbb{R}$ . Utifrån den alternativa definitionen av  $\cos(x)$  och  $\sin(x)$  kan vi härleda alla trigonometriska räkneregler.

**Exempel 1.13** För att visa att  $\frac{d}{dx} \sin(x) = \cos(x)$  gör vi följande kalkyl. Låt  $z(x) = \sin(x)$  enligt Definition 1.13. Vi deriverar ekvationen och får,

$$z'''(x) + z'(x) = 0, \quad z'(0) = 1, \quad z''(0) = -z(0) = 0.$$

Men eftersom  $z'(x)$  är unik lösning till samma ekvationen som  $y(x) = \cos(x)$ , enligt Definition 1.12, och med samma begynnelsevillkor gäller  $z'(x) = y(x) = \cos(x)$ .

**Exempel 1.14** För att härleda trigonometriska ettan noterar vi att med  $y(x) = \cos(x)$  och  $z(x) = \sin(x)$  enligt definitionerna ovan gäller

$$\frac{d}{dx}(y(x)^2 + z(x)^2) = 2y(x)y'(x) + 2z(x)z'(x) = 2y(x)(z''(x) + z(x)) = 0.$$

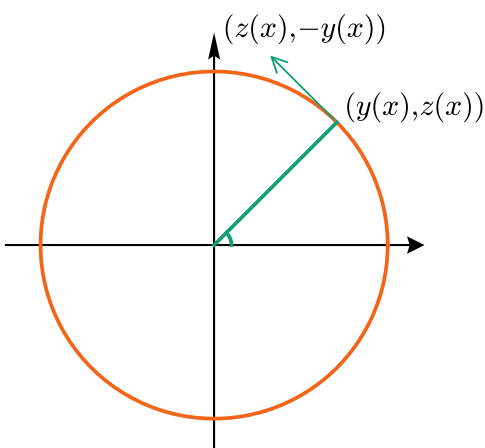
Alltså är  $y(x)^2 + z(x)^2 = C$  för någon konstant  $C$ . Men vi har  $y(0)^2 + z(0)^2 = 1 = C$  alltså  $y(x)^2 + z(x)^2 = 1$  det vill säga  $\cos^2(x) + \sin^2(x) = 1$ .

För att härleda den traditionella definitionen av  $\sin(x)$  med enhetscirkeln låter vi  $x$  parametrera kurvan  $C = (y(x), z(x)) = (y(x), y'(x))$ ,  $x \in \mathbb{R}$ . På grund av trigonometriska ettan som vi just härlett ligger alla punkter i  $C$  på enhetscirkeln, se Figur 1.9, och punkten  $(y(x), y'(x))$  rör sig kontinuerligt längs kurvan. Hastigheten som punkten  $(y(x), y'(x))$  rör sig med runt enhetscirkeln ges direkt av derivatan av positionen

$$\frac{d}{dx}(y(x), y'(x)) = (y'(x), y''(x)) = (y'(x), -y(x)) = (z(x), -y(x)).$$

Hastighetens riktning är vinkelrät mot linjen som går genom  $(y(x), z(x))$  och origo. Det följer av att skalärprodukten mellan dem är noll

$$(y(x), z(x)) \cdot (z(x), -y(x)) = y(x)z(x) - z(x)y(x) = 0.$$



**Figur 1.9:** Parametrisering av kurvan  $C = (y(x), y'(x))$ .

Farten (som är beloppet av hastigheten) är konstant lika med 1 eftersom

$$\sqrt{(y')^2 + (-y)^2} = 1, \quad \forall x \in \mathbb{R}.$$

Punkten  $(y(x), z(x))$  rör sig med konstant fart 1 längs enhetscirkeln moturs eftersom begynnelsehastigheten är  $(0, 1)$ . På ett varv går alltså  $x$  från 0 till  $2\pi$  eftersom enhetscirkelns omkrets är  $2\pi$ ,  $y(x) = y(x + 2\pi \cdot n)$  för alla  $n \in \mathbb{Z}$ . Detta innebär att  $y(x)$  exakt uppfyller den traditionella definitionen och  $y(x) = \cos(x)$ .

Vi kan även härleda övriga trigonometriska räkneregler. I nästa exempel visar vi att  $\sin(x + w) = \sin(w)\cos(x) + \cos(w)\sin(x)$  och motsvarande resultat för cosinus.

**Exempel 1.15** Vi låter  $z(x)$  lösa ekvationen,

$$z''(x) + z(x) = 0, \quad z(0) = \alpha, \quad z'(0) = \beta.$$

Vi ser att  $z(x) = \alpha \cos(x) + \beta \sin(x)$  eftersom det är en linjärkombination av lösningar till den homogena ekvationen samt att begynnelsevillkoren är uppfyllda. Vi låter nu  $\alpha = \sin(w)$  och  $\beta = \cos(w)$ . Detta motsvarar en punkt på enhetscirkeln med vinkeln  $w$ , se Figur 1.10.

Definitionen av sinus ger att  $z(x) = \sin(x + w)$  löser ekvationen

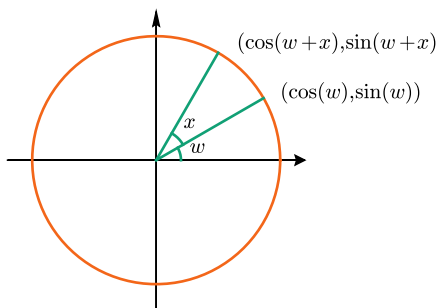
$$z''(x) + z(x) = 0, \quad z(0) = \sin(w), \quad z'(0) = \sin'(w) = \cos(w).$$

Men eftersom lösningarna är entydiga gäller

$$\sin(x + w) = \alpha \cos(x) + \beta \sin(x) = \sin(w)\cos(x) + \cos(w)\sin(x).$$

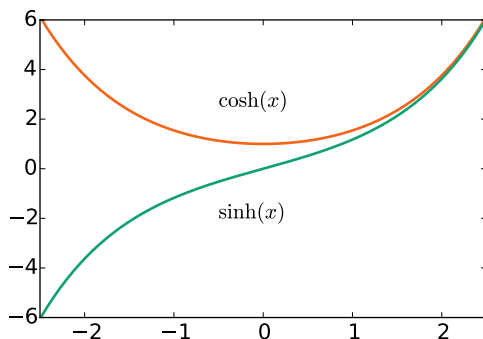
Genom derivering får vi också att  $\cos(x + z) = \frac{d}{dx} \sin(x + z) = \frac{d}{dx} \sin(w)\cos(x) + \cos(w)\sin(x) = \cos(w)\cos(x) - \sin(w)\sin(x)$ . Dessutom följer direkt att  $\sin(2x) = 2\cos(x)\sin(x)$  och  $\cos(2x) = \cos^2(x) - \sin^2(x)$ .





**Figur 1.10:** Funktionerna  $\cos(x+z)$  och  $\sin(x+z)$  illustrerade med hjälp av enhetscirkeln.

Andra elementära funktioner löser också differentialekvationer som är vanligt förekommande i modellering av fysikaliska förlopp. I Figur 1.11 ser vi funktionerna  $\sinh$  och  $\cosh$ . De kan



**Figur 1.11:** Funktionerna  $\cosh(x)$  och  $\sinh(x)$ .

även uttryckas med hjälp av exponentialfunktionen som,

$$\cosh(x) = \frac{e^x + e^{-x}}{2}, \quad (1.34)$$

$$\sinh(x) = \frac{e^x - e^{-x}}{2}. \quad (1.35)$$

De löser också andra ordningens ODE. Vi har att  $y(x) = \cosh(x)$  löser,

$$y''(x) - y(x) = 0, \quad y(0) = 1, \quad y'(0) = 0,$$

och  $z(x) = \sinh(x)$  löser,

$$z''(x) - z(x) = 0, \quad z(0) = 0, \quad z'(0) = 1.$$

Funktionen  $\cosh(x)$  beskriver en kedja som hänger fritt mellan två punkter under inverkan av gravitation.

## Övningar

### 1.1 Linjär algebra

**Övning 1.1** Givet vektorerna  $\vec{v} = [2 \ 1 \ -2]$  och  $\vec{w} = [0 \ 3 \ -1]$  bestäm följande uttryck.

- (a)  $\vec{v} + \vec{w}$  (b)  $3\vec{v} - \vec{w}$  (c)  $3\vec{w}$  (d)  $\vec{v} - 2\vec{w}$

**Övning 1.2** Beräkna längden av följande vektorer.

- (a)  $[2 \ 1 \ 1.5]$  (b)  $[2 \ 1]$  (c)  $[-2 \ -2 \ -2 \ -2]$  (d)  $[1 \ 1 \ \dots \ 1] \in \mathbb{R}^n$

**Övning 1.3** Beräkna skalärprodukten mellan följande vektorer.

- (a)  $[3 \ -1 \ 4] \cdot [0 \ 1 \ -1]$  (b)  $[2 \ 5] \cdot [1 \ 0]$  (c)  $[1 \ 1 \ 1] \cdot [0 \ 0 \ -1]$  (d)  $[1 \ 2 \ 3 \ 4] \cdot [1 \ 2 \ 3 \ 4]$

**Övning 1.4** Beräkna vinkeln mellan följande vektorer.

- (a)  $[\sqrt{3} \ 0 \ 0]$  och  $[3 \ \sqrt{3} \ 0]$  (b)  $[1 \ 1 \ 1]$  och  $[1 \ -2 \ -1]$  (c)  $[1 \ 0]$  och  $[1 \ \sqrt{3}]$  (d)  $[1 \ 0 \ \dots \ 0] \in \mathbb{R}^n$  och  $[1 \ 1 \ \dots \ 1] \in \mathbb{R}^n$ .

**Övning 1.5** Lös ekvationssystemen.

(a)

$$\begin{cases} 2x + 3y = 8 \\ x - y = -1 \end{cases}$$

(b)

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

(c)

$$\begin{cases} x + y + z = 6 \\ \quad + 2y + z = 7 \\ \quad - y + z = 1 \end{cases}$$

(d)

$$\begin{cases} x - y - 3z = 3 \\ \quad y + z = 2 \\ \quad - 3y + 2z = -1 \end{cases}$$

**Övning 1.6** Skriv problem 1.5.a-1.5d på matrisform,  $Ax = b$ .

**Övning 1.7** Beräkna produkten  $A\vec{x}$  mellan matrisen  $A = \begin{bmatrix} 1 & 0 & 3 \\ -1 & 2 & 0 \\ 0 & 2 & 1 \end{bmatrix}$  och följande vektorer.

$$(a) \vec{x} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \quad (b) \vec{x} = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} \quad (c) \vec{x} = \begin{bmatrix} 0.5 \\ 0 \\ 2 \end{bmatrix} \quad (d) \vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

### 1.2 System av första ordningen

**Övning 1.8** Skriv följande system av två första ordningens ODE'er som en andra ordningens ODE och bestäm lösningen.

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

**Övning 1.9** Skriv följande system av två första ordningens ODE'er som en andra ordningens ODE. Bestäm lösningen.

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} v \\ v^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

**Övning 1.10** Skriv om systemet:

$$\begin{cases} y_1'(t) = y_2(t) \\ y_2'(t) = ty_1(t)^2 \end{cases}$$

som en 2:a ordningens differentialekvation i variabeln  $y(t) = y_1(t)$ .

**Övning 1.11** Skriv om följande system som en andra ordningens ODE och beräkna lösningen.

$$\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \end{bmatrix} + \begin{bmatrix} u_2 \\ -u_1 \end{bmatrix} = \begin{bmatrix} \sin(2t) \\ \cos(2t) \end{bmatrix}, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

**Övning 1.12** Skriv följande system som en tredjeordningens ODE.

$$\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{u}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 0 & 1 \\ 2 & 1 & 0.5 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \\ u_3(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}.$$

### 1.3 Högre ordningens ODE som system av första ordningen

**Övning 1.13** Skriv följande andra ordningens ODE som ett system av två första ordningens ODE'er,

$$y'' - 2y' - 3y = x, \quad y(0) = 1, \quad y'(0) = 0.$$

Skriv systemet på matrisform. Bestäm lösningen.

**Övning 1.14** Skriv följande andra ordningens ODE som ett system av två första ordningens ODE'er,

$$y'' + \sin(y) = x \cos(x), \quad y(0) = 1, \quad y'(0) = 1.$$

**Övning 1.15** Skriv följande differentialekvation  $y'''(x) + x \sin(y'(x)) = x$ ,  $y(0) = 1$ ,  $y'(0) = 0$ ,  $y''(0) = 1$  som ett system av första ordningen.

**Övning 1.16** Skriv följande ODE som ett system av första ordningen,

$$y''(x) - 8y'(x) + 3y(x) = xe^x, \quad y(0) = 2, \quad y'(0) = 3.$$

**Övning 1.17** Skriv följande ODE som ett system av första ordningen,

$$y^{(4)}(x) - 2y''(x) + 3y'(x) + xy(x) = e^{2x} + \cos(x), \quad y(0) = 2, \quad y'(0) = 3, \quad y''(0) = 1, \quad y^{(3)}(0) = 2.$$

**Övning 1.18** Skriv följande ODE som ett system av första ordningen,

$$y''(x) + y^2 y'(x) = e^x, \quad y(0) = -1, \quad y'(0) = 2.$$

#### 1.4 Existens och entydighet av lösning

**Övning 1.19** Bestäm en Lipschitzkonstant för  $\vec{f}(t, \vec{y}) = \begin{bmatrix} -v \\ 4u \end{bmatrix}$  med avseende på  $\vec{y} = \begin{bmatrix} u \\ v \end{bmatrix}$  från Övning 1.8. Har systemet entydig lösning för alla  $t$ ?

**Övning 1.20** Är  $\vec{f}(t, \vec{v}) = \begin{bmatrix} v \\ v^2 \end{bmatrix}$  från Övning 1.9 Lipschitzkontinuerligt för alla  $\vec{v} = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2$ ?

**Övning 1.21** Ekvationen  $y'(x) = 4y^{3/4}(x)$  med begynnelsevillkor  $y(0) = 0$  har två lösningar  $y = 0$  och  $y = x^4$ . Vad i Sats 1.3 är inte uppfyllt (eftersom lösningen inte är entydig)?

#### 1.5 Alternativ definition av elementära funktioner

**Övning 1.22** Låt exponentialfunktionen  $y(x) = e^x$  vara definierad som lösning till ekvationen  $y'(x) = y(x)$ , med begynnelsevillkoret  $y(0) = 1$ . Visa utifrån denna definition att  $u(x) = y(x)^2$  löser ekvationen  $u'(x) = 2u(x)$ , där  $u(0) = 1$ .

**Övning 1.23** Härled logaritmlagen  $\ln(xy) = \ln(x) + \ln(y)$ ,  $x, y > 0$  utifrån relationen  $e^{xy} = e^x e^y$ .

**Övning 1.24** Härled logaritmlagen  $\ln(x^p) = p \ln(x)$ ,  $x, p > 0$ , utifrån relationen  $e^{xy} = (e^x)^y$ .

**Övning 1.25** Visa att  $\cosh(x) = \frac{e^x + e^{-x}}{2}$  och  $\sinh(x) = \frac{e^x - e^{-x}}{2}$  löser ekvationen  $y''(x) - y(x) = 0$ .

## Problem

### 1.1 Linjär algebra

**Problem 1.1** Givet två vektorer  $\vec{v}, \vec{w} \in \mathbb{R}^2$  men mellanliggande vinkel  $\theta$ , visa att  $\vec{v} \cdot \vec{w} = \cos(\theta) |\vec{v}| |\vec{w}|$  med hjälp av cosinussatsen  $|\vec{v} - \vec{w}|^2 = |\vec{v}|^2 + |\vec{w}|^2 - 2|\vec{v}| |\vec{w}| \cos(\theta)$ .

### 1.2 System a första ordningen

**Problem 1.2** Låt  $a = b = c = d = 1$  i Exempel 1.7. Avgör om högerledet i ODE'n är Lipschitz kontinuerligt för alla  $0 \leq x, y \leq 10$ . Beräkna jämviktspunkterna då  $\dot{x} = \dot{y} = 0$ .

### 1.3 Högre ordningens ODE som system av första ordningen

**Problem 1.3** Låt fjädrarna i exempel ha olika fjäderkonstanter  $k_1, k_2$  och  $k_3$ . Låt även massorna vara olika och ges av  $m_1$  och  $m_2$ . Härled motsvarande system av första ordningens ODE.

### 1.4 Existens och entydighet av lösning

**Problem 1.4** Vilken trigonometrisk funktion kan definieras av följande ODE,

$$y'(x) = 1 + y(x)^2, \quad y(0) = 0?$$

Är högerledet Lipschitzkontinuerligt för alla  $y$ ? Har vi existens av entydig lösning för alla  $x$ ?

### 1.5 Alternativ definition av exponentialfunktionen

**Problem 1.5** Studera ekvationen  $y'(x) = y(x)$ ,  $y(0) = 1$  som definierar exponentialfunktionen. Härled formeln  $e = \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n$  genom att använda följande approximation till  $\frac{y_{i+1} - y_i}{1/n} = y_{i+1}$  i Exempel 1.11.



## 2. Numerisk lösning av ODE

2.1	Numeriska metoder	27
2.2	Konvergens	30
2.3	Stabilitet	32
2.4	Algoritmer	34
2.5	Generalisering till system av ODE	37
2.6	Randvärdesproblem	40

Vi har presenterat flera analytiska metoder för att lösa ordinära differentialekvationer. För mer komplicerade modeller räcker oftast inte de analytiska teknikerna till. Numeriska metoder för att lösa ODE är ett generellt verktyg för att beräkna approximativa lösningar. Vi presenterar tre numeriska metoder med olika för- och nackdelar och därmed olika användningsområden. Vi studerar även metodernas konvergens och stabilitet. Vi generaliserar resultaten till system av ODE och diskuterar slutligen numerisk lösning av randvärdesproblem.

### 2.1 Numeriska metoder

Låt  $y(t)$  vara lösning till den ordinära differentialekvationen,

$$\begin{cases} \dot{y}(t) = f(t, y(t)), & t_0 \leq t \leq T, \\ y(t_0) = y_0. \end{cases} \quad (2.1)$$

Vi vill beräkna en numerisk approximation till  $y(t)$ , på tidsintervallet  $t \in [t_0, T]$ , där  $T$  är sluttiden. Vi börjar med att dela upp tidsintervallet  $[t_0, T]$  i en partition av  $n$  delintervall  $t_0 < t_1 < \dots < t_i < \dots < t_n = T$  med  $k_i = t_i - t_{i-1}$ ,  $i = 1, \dots, n$ , som steglängd. Idén med att dela upp intervallet i en partition känner vi igen från numerisk integration och Riemannsummor. I de numeriska metoder vi studerar approximeras tidsderivatan  $\dot{y}$  på intervallet  $[t_{i-1}, t_i]$  alltid med differenskvoten  $\dot{y} \approx \frac{y(t_i) - y(t_{i-1})}{k_i}$ . Evalueringen av högerledet  $f(t, y)$  kommer däremot göras på olika sätt. Detta ger de numeriska metoderna olika egenskaper och därmed olika användningsområden.

Eulers metod approximerar  $f(t, y(t))$  på intervallet  $t \in [t_i, t_{i-1}]$  med värdet i den vänstra punkten  $f(t_{i-1}, y(t_{i-1}))$ .

**Definition 2.1 (Eulers metod)** Låt  $t_0 < t_1 < \dots < t_n = T$  vara en partition av intervallet  $[t_0, T]$  i  $n$  delintervall av längd  $k_i = t_i - t_{i-1}$ ,  $i = 1, \dots, n$ . Eulers metod för att beräkna

en approximation  $y_n$  av lösningen  $y(t_n)$  till ekvation (2.1) ges av

$$\frac{y_i - y_{i-1}}{k_i} = f(t_{i-1}, y_{i-1}), \quad i = 1, \dots, n. \quad (2.2)$$

En god egenskap som Eulers metod har är att nästkommande värde  $y_i$  kan beräknas direkt genom insättning av föregående värde  $y_{i-1}$  i formeln  $y_{i-1} + kf(t_{i-1}, y_{i-1})$  utan att en algebraisk ekvation behöver lösas. Vi säger, på grund av detta, att Eulers metod är en explicit metod.

**Exempel 2.1 (Eulers metod)** Låt  $y(t)$  vara lösning till,

$$\begin{cases} \dot{y}(t) = \cos(y(t)), & t > 0 \\ y(0) = 1. \end{cases} \quad (2.3)$$

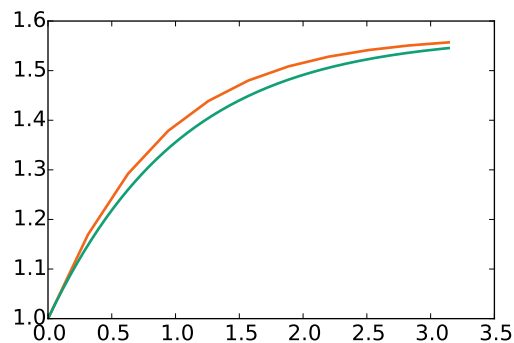
Använd Eulers metod för att beräkna en approximation till lösningen  $y(\pi)$  med konstant steglängd  $k = \pi/10$ . Vi noterar först att högerledet är Lipschitzkontinuerligt eftersom

$$|\cos(z) - \cos(y)| \leq \max_{x \in \mathbb{R}} |\sin(x)| |z - y| \leq |z - y|,$$

där vi använder att Lipschitzkonstanten är begränsad av derivatans ( $\frac{d}{dx} \cos(x) = -\sin(x)$ ) maximala värde till absolutbelopp. Picards sats ger därmed att det finns en unik lösning  $y(t)$  definierad för alla  $t \geq 0$ . Vi låter  $t_i = ik = i\pi/10$ ,  $i = 0, \dots, 10$ . I Eulers metod ges nästkommande värde av

$$y_i = y_{i-1} + k \cos(y_{i-1}). \quad (2.4)$$

Vi jämför Euler-approximationen med en referenslösning beräknad på ett mycket finare nät. Felet mellan Euler-approximationen med tio delintervall och en referenslösningen vid sluttiden är 0.011574855438. I Figur 2.1 ser vi hur väl den numeriska lösningen med Eulers metod (orange) approximerar den exakta lösningen (grön).



**Figur 2.1:** Approximationen till  $y(t)$  uträknad med Eulers metod (orange) med 10 delintervall tillsammans med den exakta lösningen (grön).



Bakåt Euler är snarlik Eulers metod i formuleringen men har andra egenskaper och användningsområden. Istället för att evaluera högerledet i  $f(t_{i-1}, y_{i-1})$  använder vi den högra ändpunkten  $f(t_i, y_i)$ .

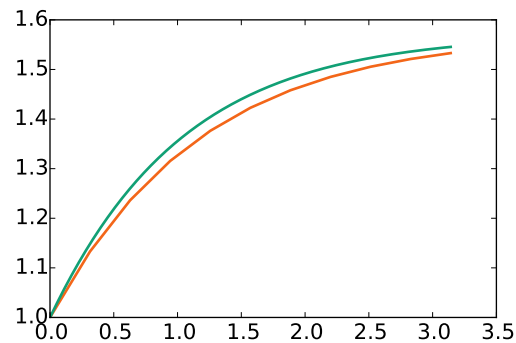
**Definition 2.2 (Bakåt Euler)** Låt  $t_0 < t_1 < \dots < t_n = T$  vara en partition av intervallet  $[t_0, T]$  i  $n$  delintervall av längd  $k_i = t_i - t_{i-1}$ ,  $i = 1, \dots, n$ . Bakåt Eulermetoden för att beräkna en approximation  $y_n$  av lösningen  $y(t_n)$  till ekvation (2.1) ges av

$$\frac{y_i - y_{i-1}}{k_i} = f(t_i, y_i), \quad i = 1, \dots, n. \quad (2.5)$$

**Exempel 2.2 (Bakåt Euler)** Vi använder samma exempel som för Eulers metod för att kunna jämföra approximationerna. Med bakåt Euler och  $t_i = i \cdot k = i\pi/10$ ,  $i = 0, \dots, 10$  får vi

$$y_i = y_{i-1} + k \cos(y_i). \quad (2.6)$$

Vi behöver lösa en algebraisk ekvation på formen  $x = g(x) := y_{i-1} + k \cos(x)$  givet  $y_{i-1}$  och  $k$ . Vi ser direkt att  $|g'(x)| \leq k$  för alla  $x \in \mathbb{R}$  vilket innebär att fixpunktsiteration kommer konvergera om  $k < 1$ . Vi jämför även bakåt Euler-approximationen med en referenslösning. Felet mellan bakåt Euler-approximationen med tio delintervall och den exakta lösningen i sluttiden är 0.0124759125631, alltså av samma storleksordning som Eulers metod. I Figur 2.2 ser vi den exakta lösningen (grön) tillsammans med bakåt Euler-approximationen (orange).



**Figur 2.2:** Approximationen till  $y(t)$  uträknad med bakåt Euler (orange) med 10 delintervall tillsammans med referenslösningen (grön).

Som vi ser i exemplet kräver bakåt Euler att en algebraisk ekvation löses i varje tidsteg för att beräkna nästa värde  $y_i$ . Metoden är därför inte explicit, som Eulers metod, utan implicit. I Eulers metod ges nästa värde explicit genom insättning av föregående medan i bakåt Euler ges den implicit genom ekvationslösning. Bakåt Euler går även under namnet implicit Euler.

Mittpunktsmetoden evaluerar, som namnet antyder, högerledet i mittpunkten på intervallet, det vill säga  $f(\frac{t_i+t_{i-1}}{2}, \frac{y_i+y_{i-1}}{2})$ . Detta innebär att mittpunktsmetoden är implicit precis som bakåt Euler, eftersom den okända  $y_i$  ingår i högerledet.

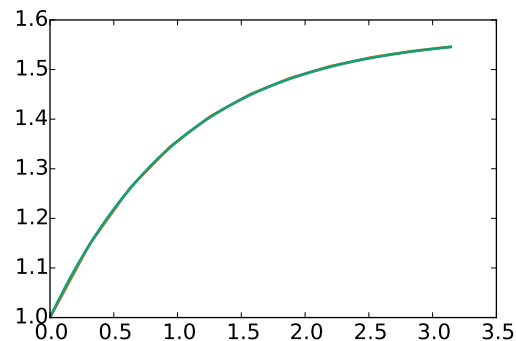
**Definition 2.3 (Mittpunktsmetoden)** Låt  $t_0 < t_1 < \dots < t_n = T$  vara en partition av intervallet  $[t_0, T]$  i  $n$  delintervall av längd  $k_i = t_i - t_{i-1}$ ,  $i = 1, \dots, n$ . Mittpunktsmetoden för att beräkna en approximation  $y_n$  av lösningen  $y(t_n)$  till ekvation (2.1) ges av

$$\frac{y_i - y_{i-1}}{k_i} = f\left(\frac{t_i + t_{i-1}}{2}, \frac{y_i + y_{i-1}}{2}\right), \quad i = 1, \dots, n. \quad (2.7)$$

**Exempel 2.3 (Mittpunktsmetoden)** Vi fortsätter med samma exempel för att kunna jämföra metoderna. Mittpunktsmetoden ges av,

$$y_i = y_{i-1} + k \cdot \cos\left(\frac{y_{i-1} + y_i}{2}\right). \quad (2.8)$$

Även här behöver vi lösa en algebraisk ekvation på formen  $x = g(x) := y_{i-1} + k \cos\left(\frac{y_{i-1} + x}{2}\right)$  givet  $y_{i-1}$  och  $k$ . Vi ser att  $|g'(x)| \leq \frac{k}{2}$  för alla  $x \in \mathbb{R}$  vilket innebär att fixpunktsiteration kommer konvergera för  $k < 2$ . Som startgissning kan vi använda  $y_{i-1}$ . I Figur 2.3 jämför mittpunktsmetoden (orange) med referenslösningen (grön). Felet är nu bara 0.000639866293157 i sluttiden, vilket är ungefär 18 – 19 gånger mindre än för de andra metoderna. Det är svårt att se skillnad på approximationen och den exakta lösningen i figuren.



**Figur 2.3:** Approximationen till  $y(t)$  uträknad med mittpunktsmetoden (orange) med 10 delintervall och referenslösningen (grön). Kurvorna går nästan helt i varandra.

Felet i approximationen avtar snabbare när antal intervall ökas för mittpunktsmetoden i jämförelse med både Eulers metod och bakåt Euler. Vi ska nu studera konvergenserna mer i detalj.

## 2.2 Konvergens

Vi har redan sett att felet i mittpunktsmetodens approximation var betydligt mindre än motsvarande fel för Eulers metod och bakåt Euler i exemplen med ekvationen  $y' = \cos(y)$ . Begreppet konvergensordning används för att beskriva noggrannheten i en numerisk metod.

**Definition 2.4 (Konvergensordning)** En numerisk metod för att lösa ett begynnelsevärdesproblem med konstant steglängd  $k$  är av ordning  $n$  om felet vid sluttiden är proportionellt mot  $k^n$ .

Vi formulerar satsen om konvergens för Eulers metod. Beviset bygger på Taylorutveckling av den exakta lösningen i varje tidssteg.

**Sats 2.1 (Konvergens av Eulers metod)** Låt  $y(t)$  lösa begynnelsevärdesproblemet

$$\frac{dy}{dt} = f(t, y), \quad y(t_0) = y_0, \quad (2.9)$$

där den kontinuerliga funktionen  $f(t, y)$  är Lipschitzkontinuerlig i  $y$  med Lipschitzkonstant  $L$  och lösningen uppfyller  $\max_{t \in [t_0, T]} \left| \frac{d^2 y}{dt^2} \right| \leq M$ . Låt vidare  $y_n$  vara Eulerapproximationen av  $y(t_n)$  med konstant steglängd  $k$ , där  $y_n = nk$ ,  $n = 0, 1, 2, \dots$ . Då gäller

$$|y(t_n) - y_n| \leq \frac{Mk}{2L} (e^{L(t_n - t_0)} - 1). \quad (2.10)$$

*Bevis.* Vi studerar först felet  $e_i = y(t_i) - y_i$  på intervallet  $[t_{i-1}, t_i]$ . Taylors formel ger att  $y(t_i)$  kan uttryckas med en Taylorserie kring punkten  $y(t_{i-1})$

$$y(t_i) = y(t_{i-1}) + ky'(t_{i-1}) + \frac{k^2}{2} y''(\eta_i),$$

för något  $\eta_i \in [t_{i-1}, t_i]$ . Av ekvationen  $y'(t) = f(t, y(t))$  följer att,

$$y(t_i) = y(t_{i-1}) + kf(t_{i-1}, y(t_{i-1})) + \frac{k^2}{2} y''(\eta_i).$$

Eulers metod ger att  $y_i = y_{i-1} + k \cdot f(t_{i-1}, y_{i-1})$ . Vi får,

$$e_i = e_{i-1} + k(f(t_{i-1}, y(t_{i-1})) - f(t_{i-1}, y_{i-1})) + \frac{k^2}{2} y''(\eta_i). \quad (2.11)$$

och eftersom  $f$  är Lipschitzkontinuerlig i  $y$  med konstant  $L$  och andra derivatan av  $y$  är begränsad av  $M$  får vi

$$|e_i| \leq |e_{i-1}| + kL|e_{i-1}| + \frac{k^2 M}{2} = (1 + kL)|e_{i-1}| + \frac{k^2 M}{2}.$$

Om vi upprepar denna räkning från  $i = n$  ner till  $i = 1$  får vi för det globala felet efter  $n$  steg

$$|e_n| \leq (1 + kL)|e_{n-1}| + \frac{k^2 M}{2} \quad (2.12)$$

$$\leq (1 + kL)^2 |e_{n-2}| + (1 + kL) \frac{k^2 M}{2} + \frac{k^2 M}{2} \quad (2.13)$$

$$\leq (1 + kL)^n |e_0| + \frac{k^2 M}{2} (1 + (1 + kL) + (1 + kL)^2 + \dots + (1 + kL)^{n-1}) \quad (2.14)$$

$$= \frac{k^2 M}{2} \sum_{i=0}^{n-1} (1 + kL)^i, \quad (2.15)$$

eftersom  $|e_0| = |y(t_0) - y_0| = 0$ . Summationsformeln för geometrisk summa ger

$$|y(t_n) - y_n| = |e_n| \leq \frac{k^2 M}{2} \frac{(1 + kL)^n - 1}{kL} \leq \frac{Mk}{2L} (e^{nkL} - 1) = \frac{Mk}{2L} (e^{(T-t_0)L} - 1),$$

där vi använder att  $1 + x \leq e^x$ , med  $x = kL$ . Detta följer eftersom vi har likhet för  $x = 0$  och högerledet har derivata strikt större än ett för  $x > 0$ .  $\square$

Bakåt Euler är också en första ordningens metod medan mittpunktsmetoden är av andra ordningen. Beviset för bakåt Euler är snarlikt det för Eulers metod och länas som övning. För mittpunktsmetoden görs Taylorutvecklingen av  $y(t_i)$  och  $y_{t_{i-1}}$  kring mittpunkten till och med tredjederivator, vilket leder till cancellation av andraderivatorna. I övrigt är idén densamma.

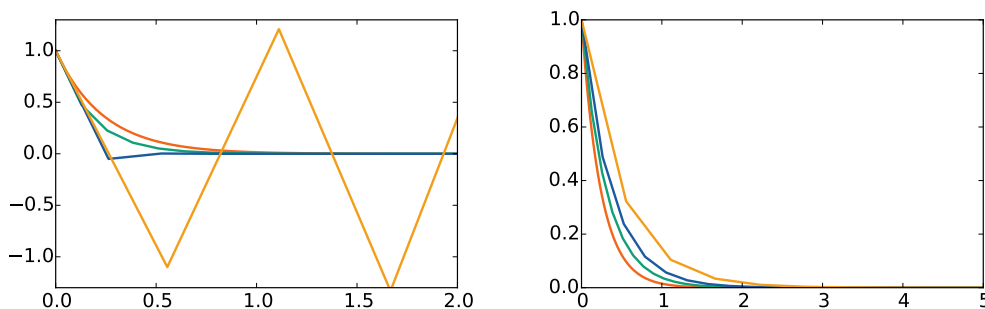
## 2.3 Stabilitet

Om vi använder Eulers metod för att lösa en ODE vars lösning varierar snabbt, uppträder ett fenomen som kallas instabilitet. Vi illustrerar det med ett exempel.

**Exempel 2.4** Vi studerar följande ODE,

$$y'(t) = -4.2y(t), \quad y(0) = 1.$$

Den analytiska lösningen ges av  $y(t) = e^{-4.2t}$ . Vi beräknar lösningen fram till  $T = 5$  med Eulers metod. I Figur 2.4 (vänster) är den röda linjen den analytiska lösningen, grön motsvarar steglängden  $k = 5/40$ , blå  $k = 5/20$  och orange  $k = 5/10$ . För  $k = 5/10$  beter sig den numeriska approximationen märkligt. Lösningen börjar oscillera och växa trots att den analytiska lösningen avtar. Motsvarande fenomen finns inte för bakåt Euler, som syns i Figur 2.4 (höger). Vi säger att bakåt Euler är stabil för detta problem medan Eulers metod med steglängd  $k = 5/10$  är instabil.



**Figur 2.4:** Eulers metod (vänster) och bakåt Euler (höger) med steglängd  $k = 5/10, 5/20, 5/40$ . Vi ser att Eulers metod med steglängd  $k = 5/10$  ger en instabil, växande, lösning medan bakåt Euler är stabil för alla  $k$ .

Exemplet visar på en svaghet med feluppskattningarna i Sats 2.1. Nämligen att faktorn  $e^{L|t_n - t_0|}$  i uppskattningen kan bli väldigt stor. Denna term uppstår eftersom vi i varje iteration multiplicerar felet från föregående tidssteg med faktorn  $1 + kL$ , där  $L$  är högerledets

( $f$ 's) Lipschitzkonstant med avseende på  $y$ . Denna skattning är ofta för grov. Om vi går tillbaka till ekvation (2.11) i beviset, som beskriver hur felet i tidssteg  $i$  beror av felet i tidssteg  $i - 1$ , har vi för problemet i exemplet med  $f(t, y) = -4.2y$  att,

$$e_i = e_{i-1} + k(f(t_{i-1}, y(t_{i-1})) - f(t_{i-1}, y_{i-1})) + \frac{k^2}{2} y''(\eta_i) = (1 - 4.2k)e_{i-1} + \frac{k^2}{2} y''(\eta_i).$$

Om faktorn  $|1 - 4.2k| > 1$  ser vi att vi riskerar att få en exponentiell tillväxt i felet. Det är precis den effekten vi ser i Figur 2.4 (vänster) då  $k = 5/10$  vilket innebär att  $|1 - 4.2 \cdot 5/10| = 1.1 > 1$ .

Det finns en teori för att avgöra när en numerisk metod är stabil. Vi använder en testekvation på formen

$$y'(t) = \lambda y(t), \quad \lambda \in \mathbb{C},$$

där  $\lambda \in \mathbb{C}$ , som tillåts vara komplex. Lösningen är då  $y(t) = e^{\lambda t}$  som avtar exponentiellt om realdelen av  $\lambda$  är mindre än noll och växer exponentiellt om realdelen av  $\lambda$  är större än noll.

Vi definierar stabilitetsområdet för en numerisk metod på följande sätt.

**Definition 2.5 (Stabilitet av numerisk metod)** Låt  $y_n$  vara en approximation till testekvationen  $y'(t) = \lambda y(t)$  som definieras av relationen  $y_i = g(k\lambda)y_{i-1}$ , där  $k$  är metodens konstant steglängd och  $i = 1, \dots, n$ . Då säger vi att metodens stabilitetsområde ges av alla  $k\lambda$  sådana att  $|g(k\lambda)| < 1$ .

Vi noterar till exempel att  $g(k\lambda) = 1 + k\lambda$  för Eulers metod. Givet denna definition kan vi härleda stabilitetsområdet för våra tre numeriska metoder.

**Sats 2.2 (Stabilitetsområden)** Stabilitetsområdet för Eulers metod ges av  $|1 + k\lambda| < 1$ , för implicit Euler  $|1 - k\lambda| > 1$  och för mittpunktsmetoden  $|\frac{1+k\lambda/2}{1-k\lambda/2}| < 1$ .

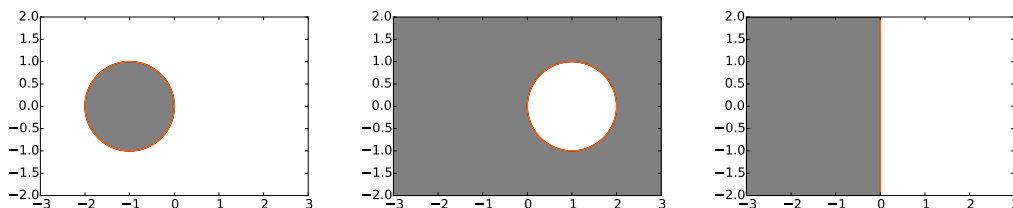
*Bevis.* Vi studerar testekvationen

$$y'(t) = \lambda y(t), \quad t > 0.$$

Eulers metod ger  $y_i = y_{i-1} + k\lambda y_{i-1} = (1 + k\lambda)y_{i-1}$ , bakåt Euler  $y_i = \frac{1}{1 - k\lambda} y_{i-1}$  och mittpunktsmetoden  $y_i = y_{i-1} + k\lambda \frac{y_{i-1} + y_i}{2}$  eller  $y_i = \frac{1 + \frac{k\lambda}{2}}{1 - \frac{k\lambda}{2}} y_{i-1}$ . Enligt Definition 2.5 gäller då att Eulers metod har stabilitetsområde  $|1 + k\lambda| < 1$ , implicit Euler  $|1 - k\lambda| > 1$  och mittpunktsmetoden  $|\frac{1 + \frac{k\lambda}{2}}{1 - \frac{k\lambda}{2}}| < 1$ .  $\square$

I Figur 2.5 illustrerar vi stabilitetsområdena med grå färg i det komplexa talplanet för Eulers metod, bakåt Euler och mittpunktsmetoden. Vi ser att olika värden på  $k\lambda$  leder antingen till en stabil eller instabil approximation.

Det är tydligt att bakåt Euler har ett stort stabilitetsområde. I vårt exempel var  $k = 5/10$  och  $\lambda = -4.2$  vilket innebär att  $k\lambda = -2.1$  som är inom bakåt Eulers stabilitetsområde. För Eulers metod däremot hamnar vi utanför stabilitetsområdet. Dock kan vi genom att minska  $k$  komma in i stabilitetsområdet igen. Det ser vi även i Figur 2.4 (vänster) då vi låter  $k = 5/20$  och  $k = 5/40$ .



**Figur 2.5:** De grå områdena är värden  $k\lambda$  i det komplexa talplanet där Eulers metod (vänster) bakåt Euler (mitten) och mittpunktsmetoden (höger) är stabila.

Sammanfattningsvis drar vi slutsatsen att både bakåt Euler och mittpunktsmetoden har ett stort stabilitetsområde. De ger därför tillförlitliga lösningar för en stor klass problem. Eulers metod har däremot ett litet stabilitetsområde vilket innebär att det i många fall krävs ett väldigt liten steglängd  $k$  för att lösningarna ska vara tillförlitliga över längre tidsintervall.

Våra tre numeriska metoder har alltså olika fördelar och nackdelar. Med Eulers metod slipper vi lösa algebraiska ekvationer vilket är en stor fördel eftersom de kan ta mycket datorkraft i anspråk. Bakåt Euler är den mest robusta metoden som nästan alltid fungerar. Mittpunktsmetoden har högre ordning vilket innebär att om metoden är stabil för ett givet problem kommer lösningen konvergera mycket snabbare än de övriga metoderna. Dock krävs att tredjederivatan av lösningen är begränsad för att en högre konvergensordningen ska uppnås. Olika problemtyper behöver alltså olika numeriska metoder. Det finns fler egenskaper som är viktiga för en numerisk metod. En sådan är metodens förmåga att bevara egenskaper hos den exakta lösningen, till exempel konservering av energi. Vi återkommer till detta längre fram i kapitlet.

## 2.4 Algoritmer

Vi är nu redo att formulera Eulers metod som en algoritm, se Algoritm 1. Sats 2.1 ger att felet i den beräknade lösningen kan begränsas i termer av steglängden  $k$ , Lipschitzkonstanten  $L$  och begränsningen av andraderivatan  $M$ . Givet  $f$  kan  $L$  beräknas. Konstanten  $M$  är däremot svår att uppskatta när lösningen inte är känd. Den kommer istället med som ett villkor i algoritmen. Vi låter för enkelhets skull steglängden vara konstant.

De implicita metoderna, bakåt Euler och mittpunktsmetoden, är lite svårare att implementera eftersom en algebraisk ekvation behöver lösas i varje tidssteg. Fixpunktsiteration är en naturlig metod att använda för att lösa dessa ekvationer eftersom vi vill lösa ekvationer på formen  $x = g(x)$ . För bakåt Euler har vi att

$$g(x) = y_{i-1} + k_i f(t_i, x),$$

och för mittpunktsmetoden att

$$g(x) = y_{i-1} + k_i f\left(\frac{t_{i-1} + t_i}{2}, \frac{y_{i-1} + x}{2}\right).$$

Vi har även en god startgissning i  $x_0 = y_{i-1}$ . Vi presenterar bakåt Euler i Algoritm 2 och mittpunktsmetoden i Algoritm 3.

**Algorithm 1** Eulers metod

**Indata:** Kontinuerlig funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , Lipschitz kontinuerlig ( $L$ ) i andra variabeln, startpunkt  $t_0, y_0 \in \mathbb{R}$ , steglängd  $k$ , slutpunkt  $t_n = T$

**Utdata:** Approximativ lösning  $\hat{y} \in \mathbb{R}$  sådan att  $|y(t_n) - \hat{y}| \leq \frac{kM}{2L}(e^{L(t_n-t_0)} - 1)$

**Villkor:**  $\max_{t \in [t_0, t_n]} |y''(t)| \leq M$

```

1:  $i \leftarrow 1$ 
2: while  $t_{i-1} < t_n$  do
3:    $t_i \leftarrow t_{i-1} + k$ 
4:    $y_i \leftarrow y_{i-1} + kf(t_{i-1}, y_{i-1})$ 
5:    $i \leftarrow i + 1$ 
6: end while
7:  $\hat{y} \leftarrow y_{i-1}$ 

```

**Algorithm 2** Bakåt Euler

**Indata:** Kontinuerlig funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , Lipschitzkontinuerlig ( $L$ ) i andra variabeln, startpunkt  $t_0, y_0 \in \mathbb{R}$ , steglängd  $k$ , slutpunkt  $t_n$

**Utdata:** Approximativ lösning  $\hat{y} \in \mathbb{R}$  sådan att  $|y(t_n) - \hat{y}| \leq Ck$

**Villkor:**  $\max_{t \in [t_0, t_n]} |y''(t)|$

```

1:  $i \leftarrow 1$ 
2: while  $t_{i-1} < t_n$  do
3:    $t_i \leftarrow t_{i-1} + k$ 
4:    $y_i \leftarrow y_{i-1} + kf(t_i, y_i) := g(y_i)$    Löses med fixpunktsiteration!
5:    $i \leftarrow i + 1$ 
6: end while
7:  $\hat{y} \leftarrow y_{i-1}$ 

```

Banachs fixpunktssats ger att vi har konvergens om  $|g'(x)| < 1$  för all  $x$  i ett intervall. För bakåt Euler innebär det att

$$|g'(x)| = k \left| \frac{df}{dy}(t_i, x) \right| < 1,$$

för alla värden  $x$  i iterationen. I exemplen ovan med  $f(t, y) = \cos(y)$  var derivatan till be-  
lopp  $|\sin(y)|$  begränsad av 1. Därför krävs att  $k < 1$  för alla  $i$ . Motsvarande resultat för  
mittpunktsmetoden blir  $|\frac{1}{2} \sin(y)| \leq 1/2$  eftersom den inre derivatan är  $\frac{1}{2}$ . Alltså konver-  
gerar fixpunktsiteration för  $k < 2$ . Om  $y$ -derivatan av  $f$  är stor för de aktuella värdena på  
 $x$  kan det vara bättre att skriva om fixpunktsiterationen på en annan form eller använda  
en annan metod som bisektion eller Newtons metod för ekvationen  $x - g(x) = 0$ . Vi vill  
nämligen undvika ett för litet tidssteg som kan leda till långsamma beräkningar.

Det är inte självklart att de algebraiska ekvationer som uppkommer har reella lösningar  
för alla steglängder. Vi illustrerar detta i ett exempel.

**Exempel 2.5 (Lösbarhet)** Vi studerar

$$y'(x) = y(x)^2, \quad y(0) = 1,$$

vars lösning, som ges av  $y(x) = \frac{1}{1-x}$ , bara existerar på intervallet  $[0, 1)$ . Vi vill nu

**Algorithm 3** Mittpunktsmetoden

**Indata:** Kontinuerlig funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , Lipschitzkontinuerlig ( $L$ ) i andra variabeln, startpunkt  $t_0, y_0 \in \mathbb{R}$ , steglängd  $k$ , slutpunkt  $t_n$

**Utdata:** Approximativ lösning  $\hat{y} \in \mathbb{R}$  sådan att  $|y(t_n) - \hat{y}| \leq Ck^2$

**Villkor:**  $\max_{t \in [t_0, t_n]} |y'''(t)| \leq M$

```

1:  $i \leftarrow 1$ 
2: while  $t_{i-1} < t_n$  do
3:    $t_i \leftarrow t_{i-1} + k$ 
4:    $y_i \leftarrow y_{i-1} + kf\left(\frac{t_{i-1}+t_i}{2}, \frac{y_{i-1}+y_i}{2}\right) := g(y_i)$    Löses med fixpunktsiteration!
5:    $i \leftarrow i + 1$ 
6: end while
7:  $\hat{y} \leftarrow y_{i-1}$ 

```

beräkna en approximation för  $x < 1$  med bakåt Euler. Vi varje tidssteg behöver vi lösa följande algebraiska ekvation,

$$y_i = y_{i-1} + ky_i^2,$$

där  $k$  anger det konstanta tidssteget. Detta är en andragradsekvation i variabeln  $y_i$  och lösningen ges av

$$y_i = \frac{1}{2k} \pm \sqrt{\frac{1}{4k^2} - \frac{y_{i-1}}{k}}.$$

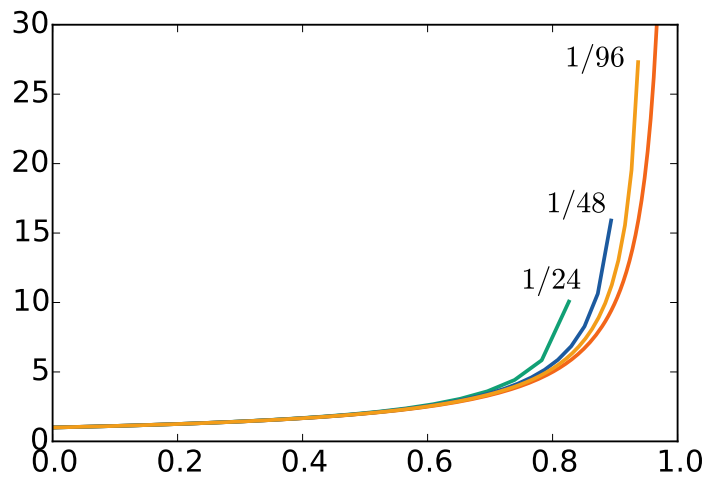
Eftersom  $y_i$  ska närma sig  $y_{i-1}$  då  $k$  går mot noll representerar endast de negativa tecknet framför rotuttrycket en lösning. Efter omskrivning får vi

$$y_i = \frac{1}{2k} \left(1 - \sqrt{1 - 4ky_{i-1}}\right).$$

Vi ser direkt att för  $i = 0$  krävs att  $k < \frac{1}{4}$  för att ekvationen ska ha en reell lösning. När  $y_{i-1}$  ökar krävs kortare och kortare tidssteg för att fixpunktsiterationen ska ge en lösning. I Figur 2.6 studerar vi hur lång tid vi kan genomföra beräkningen med olika konstanta tidssteg innan den algebraiska ekvation inte längre har någon (reell) lösning. I Figur 2.6 plottar vi bakåt Eulerlösningen, med  $k$  lika med  $1/24$ ,  $1/48$  och  $1/96$ , tillsammans med den exakta lösningen  $\frac{1}{1-x}$  i orange. När linjerna avbryts finns inte längre någon reell rot  $y_i$  till den algebraiska ekvationen. Man kan även tänka sig ett varierande tidssteg där  $k_i$  minskar när  $y_i$  ökar. En sådan metod som är anpassad till lösningen kallas en adaptiv metod.

Förutom problem med lösning av de algebraiska ekvationer som uppkommer och numerisk stabilitet som vi behandlade tidigare, finns flera andra utmaningar vid numerisk lösning av ODE. En sådan, som blir viktig om tidssteget är väldigt litet, är avrundningsfel. Vi lämnar nu algoritmerna och går vidare och generaliserar våra resultat till system av ODE.





**Figur 2.6:** Bakåt Eulerlösningen för  $k = 1/24$ ,  $k = 1/48$  och  $k = 1/96$ . Linjerna avbryts när ekvationerna saknar reella lösningar. Den exakta lösningen  $y(x) = \frac{1}{1-x}$  representeras av den röda linjen.

## 2.5 Generalisering till system av ODE

De numeriska metoder vi presenterar kan direkt generaliseras till system av första ordningens differentialekvationer.

Låt  $\vec{y}(t) \in \mathbb{R}^d$  vara lösning till den ordinära differentialekvationen på ett intervall  $I = [t_0, T]$ ,

$$\begin{cases} \dot{\vec{y}}(t) = \vec{f}(t, \vec{y}(t)) \\ \vec{y}(t_0) = \vec{y}_0, \end{cases} \quad (2.16)$$

där  $\vec{y}_0 \in \mathbb{R}^d$  och  $\vec{f}(t, \vec{y}(t)) \in \mathbb{R}^d$ . Vi låter den numeriska approximationen till  $\vec{y}(t_i)$  i tidpunkten  $t_i$  betecknas  $\vec{y}_i$ . Vi presenterar de tre metoderna igen, nu för system av ODE.

**Definition 2.6 (Eulers metod)** Låt  $t_0 < t_1 < \dots < t_n = T$  vara en partition av intervallet  $[t_0, T]$  i  $n$  delintervall av längd  $k_i = t_i - t_{i-1}$ ,  $i = 1, \dots, n$ . Eulers metod för att beräkna en approximation  $\vec{y}_n$  av lösningen  $\vec{y}(t_n)$  till ekvation (2.16) ges av

$$\frac{\vec{y}_i - \vec{y}_{i-1}}{k_i} = \vec{f}(t_{i-1}, \vec{y}_{i-1}), \quad i = 1, \dots, n. \quad (2.17)$$

**Definition 2.7 (Bakåt Euler)** Låt  $t_0 < t_1 < \dots < t_n = T$  vara en partition av intervallet  $[t_0, T]$  i  $n$  delintervall av längd  $k_i = t_i - t_{i-1}$ ,  $i = 1, \dots, n$ . Bakåt Eulermetoden för att beräkna en approximation  $\vec{y}_n$  av lösningen  $\vec{y}(t_n)$  till ekvation (2.16) ges av

$$\frac{\vec{y}_i - \vec{y}_{i-1}}{k_i} = \vec{f}(t_i, \vec{y}_i), \quad i = 1, \dots, n. \quad (2.18)$$

**Definition 2.8 (Mittpunktsmetoden)** Låt  $t_0 < t_1 < \dots < t_n = T$  vara en partition av intervallet  $[t_0, T]$  i  $n$  delintervall av längd  $k_i = t_i - t_{i-1}$ ,  $i = 1, \dots, n$ . Mittpunktsmetoden för att beräkna en approximation  $\vec{y}_n$  av lösningen  $\vec{y}(t_n)$  till ekvation (2.16) ges av

$$\frac{\vec{y}_i - \vec{y}_{i-1}}{k_i} = \vec{f}\left(\frac{t_i + t_{i-1}}{2}, \frac{\vec{y}_i + \vec{y}_{i-1}}{2}\right), \quad i = 1, \dots, n. \quad (2.19)$$

Notera att de implicita metoderna (Bakåt Euler och mittpunktsmetoden) nu kräver lösning av en algebraisk ekvation som i detta fallet är vektorvärd. Dessa ekvationer löses också med fixpunktsiteration. Genom att minska tidssteget  $k$  kan vi garantera att iterationen konvergerar med Banachs fixpunktssats. Konvergensthastigheten påverkas av Lipschitzkonstanten  $L$  till funktionen  $\vec{f}$ ,

$$|\vec{f}(t, \vec{y}) - \vec{f}(t, \vec{z})| \leq L|\vec{y} - \vec{z}|, \quad t \in I, \quad \vec{y}, \vec{z} \in \mathbb{R}^n,$$

som kan bli mer komplicerad att räkna ut i det vektorvärda fallet i jämförelse med det skalärvärda.

**Exempel 2.6 (Lipschitzkonstant)** Vi låter  $n = 2$  och studerar begynnelsevärdesproblemet

$$\dot{\vec{y}}(t) = \vec{f}(t, \vec{y}(t)), \quad \vec{y}(0) = \vec{y}_0,$$

där

$$\vec{f}(t, \vec{y}) = \begin{bmatrix} f_1(y_1, y_2) \\ f_2(y_1, y_2) \end{bmatrix} = \begin{bmatrix} 1 - 3y_2 \\ y_1 \end{bmatrix}.$$

Vi har vi att,

$$\begin{aligned} |\vec{f}(t, \vec{y}) - \vec{f}(t, \vec{z})| &\leq \sqrt{(3z_2 - 3y_2)^2 + (y_1 - z_1)^2} \\ &\leq 3\sqrt{(y_1 - z_1)^2 + (y_2 - z_2)^2} \\ &= 3|\vec{y} - \vec{z}|. \end{aligned}$$

En Lipschitzkonstant ges alltså av  $L = 3$ .

Uppskattningen av felet följer samma mönster som i det skalära fallet. Vi nöjer oss med att formulera resultatet för Eulers metod. Beviset bygger helt på beviset av Sats 2.1 och utelämnas.

**Sats 2.3 (Konvergen analys för Eulers metod för system)** Låt  $\vec{y}(t)$  lösa begynnelsevärdesproblemet

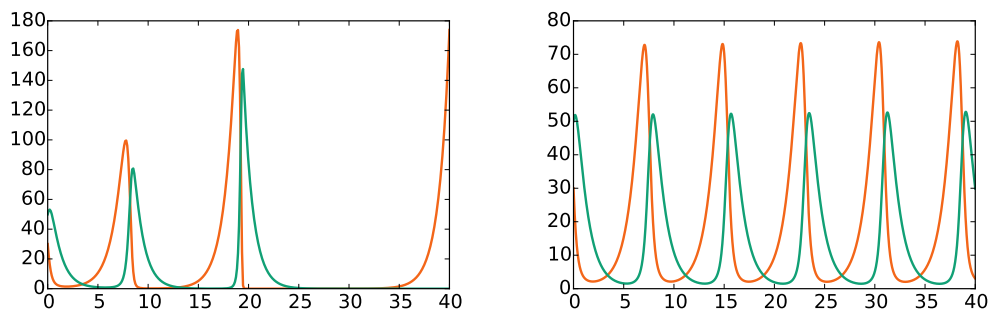
$$\frac{d\vec{y}}{dt} = \vec{f}(t, \vec{y}), \quad \vec{y}(t_0) = \vec{y}_0, \quad (2.20)$$

där den kontinuerliga funktionen  $\vec{f}(t, \vec{y})$  är Lipschitzkontinuerlig i  $\vec{y}$  med Lipschitzkonstant  $L$  och  $|\frac{d^2}{dt^2} \vec{y}| \leq M$ . Låt  $\vec{y}_n$  vara Euler approximationen av  $\vec{y}(t_n)$  med konstant steglängd  $k$ , där  $t_n = nk$ ,  $n = 0, 1, 2, \dots$ . Då gäller

$$|\vec{y}(t_n) - \vec{y}_n| \leq \frac{Mk}{2L} (e^{L(t_n - t_0)} - 1). \quad (2.21)$$

*Bevis.* Beviset följer av beviset av Sats 2.1. Vi behöver Taylorutveckla i alla  $n$  komponenter och tolka absolutbeloppet som längden av vektorerna.  $\square$

Konvergensordningen för Eulers metod är densamma som i det skalära fallet. Detsamma gäller bakåt Euler och mittpunktsmetoden som är av första respektive andra ordningen. Stabilitetsanalysen är mer komplicerad för system av ODE och kräver kunskap om egenvärden från linjär algebra. Vi nöjer oss med att illustrera stabilitetsproblemet genom att använda Eulers metod för att beräkna lösningen till modellen för rävar och kaniner, som vi presenterade redan i Figur 1.4. Vi börjar med att lösa problemet med 400 delintervall, vilket ger en steglängd på  $k = 0.1$ , och sedan med 40000 delintervall vilket motsvarar  $k = 0.001$ . Sluttiden för beräkningen är satt till  $T = 40$ .



**Figur 2.7:** Lösning för systemet med rävar och kaniner med Eulers metod och steglängd  $k = 0.1$  (vänster) respektive  $k = 0.001$  (höger). Vi ser att lösningen för  $k = 0.1$  är instabil medan den för  $k = 0.001$  är stabil på grund av tillräckligt litet tidssteg.

Vi ser att vi har tydliga problem med numerisk instabilitet i det första exemplet. Lösningen ökar i  $t$  medan den exakta lösningen behåller samma magnitud över tid. I det andra exemplet har vi tillräckligt liten steglängd för att Eulers metod ska vara stabil.

Det finns fler egenskaper hos en numerisk lösning som är eftersträvsvärd förutom stabilitet och konvergens. En sådan är bevarande av egenskaper hos lösningen. Vi illustrerar detta med ett exempel.

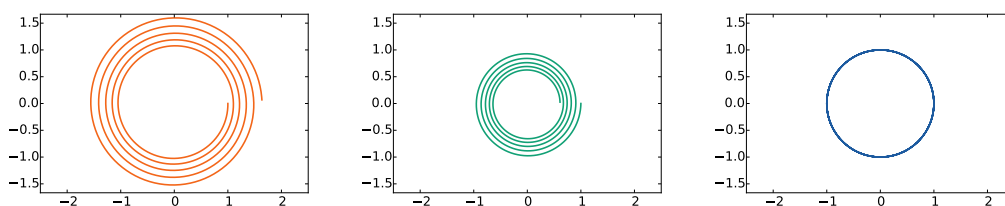
**Exempel 2.7 (Energibevarande)** Vi går tillbaka till kapitel 1 och de alternativa definitionerna av  $\cos(t)$  och  $\sin(t)$  som system av ODE. Vi har följande linjära system av ODE

$$\begin{bmatrix} y'(x) \\ z'(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} y(x) \\ z(x) \end{bmatrix}, \quad \begin{bmatrix} y(0) \\ z(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

vars lösning ges av  $y(t) = \cos(t)$  och  $z(t) = \sin(t)$ . Kvantiteten  $E(t) = \frac{1}{2}(y(t)^2 + z(t)^2)$  kan tolkas som systemets energi vid tiden  $t$ . Trigonometriska ettan ger att  $E(t) = \frac{1}{2}$  för alla  $t$ , det vill säga energin är bevarad.

Vi löser nu systemet på intervallet  $[0, 10\pi]$  med 1000 steg,  $k = 10\pi/1000$ , med Eulers metod, bakåt Euler och mittpunktsmetoden. I Figur 2.8 plottar vi approximationerna till  $(y(t), z(t))$  i planet. Notera att den exakta lösningen ligger på enhetscirkeln. Eulers metod ger en lösning som tillför energi och cirkelarna växer och växer medan bakåt Euler tar bort energi och cirkelarna minskar. Mittpunktsmetoden däremot bevarar energin över tid. Detta är ännu en aspekt som kan vara viktig för en numerisk

approximation och som är en styrka för mittpunktsmetoden.



**Figur 2.8:** Vi ser att Eulers metod leder till att cirkelarna vidgas (vänster) och bakåt Euler till att de krymper (mitten) medan mittpunktsmetoden bevarar energin.

## 2.6 Randvärdesproblem

Så här långt har vi helt och hållet fokuserat på begynnelsevärdesproblem. Det finns dock en annan klass av differentialekvationer som är minst lika viktig i tillämpningar, nämligen randvärdesproblem. Då begynnelsevärdesproblemen endast har (begynnelse)villkor vid en startpunkt har randvärdesproblemen (rand)villkor både i start- och slutpunkten. Vi ger ett exempel från värmeledning i en endimensionell stav.

**Exempel 2.8 (Värmeledningsekvationen)** Vi studerar stationär värmeledning i en endimensionell stav av längd  $L$ . Temperaturen  $u(x)$  löser följande randvärdesproblem som också är en ODE,

$$\begin{cases} -\frac{d}{dx}(\kappa \frac{du}{dx}) = q(x), & 0 < x < L, \\ u(0) = u(L) = 0, \end{cases} \quad (2.22)$$

där  $x$  är rumskoordinaten,  $\kappa$  är värmeledningskoefficienten och  $q(x)$  är värmeförseln som beror av positionen i rummet. Genom att införa hjälpvariabelerna  $u_1 = u$  och  $u_2 = \frac{du_1}{dx}$  kan vi skriva om ekvationen som ett system av första ordningens ODE,

$$\frac{du_1}{dx} = u_2, \quad (2.23)$$

$$\frac{du_2}{dx} = -\frac{q(x)}{\kappa}, \quad (2.24)$$

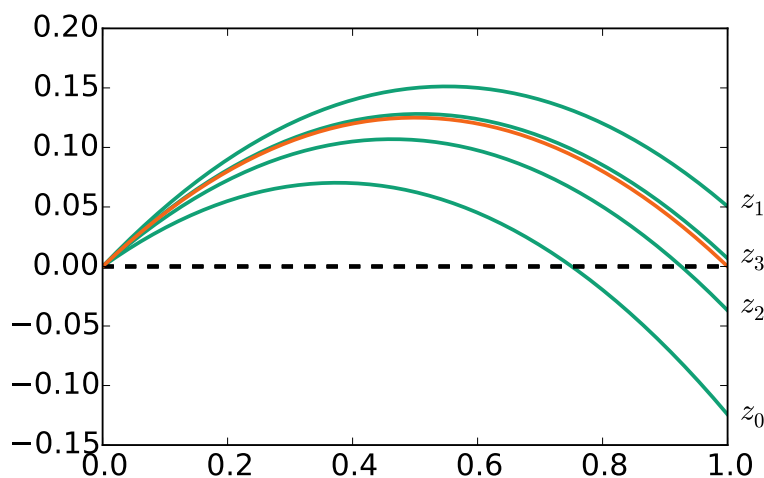
med begynnelsevillkor bara i den första variabeln  $u_1(0) = 0$  fast med det extra randvillkoret  $u_1(L) = 0$ .

Anledningen till att vi så här långt fokuserat på begynnelsevärdesproblem är att randvärdesproblem ofta leder till partiella differentialekvationer eftersom den oberoende variabeln vanligen är rumskoordinater i  $\mathbb{R}^3$ . I fallet med staven gör vi en endimensionell approximation av ett tredimensionellt objekt. Begynnelsevärdesproblemen har typiskt tiden som oberoende variabel vilken naturligt är endimensionell. En annan anledning är

att de mer kraftfulla numeriska metoderna för att lösa randvärdesproblem kräver djupare kunskaper i linjär algebra.

Det finns dock en teknik som gör att våra ODE-lösare även kan användas för att lösa randvärdesproblem. Metoden kallas inskjutning och idén är enkel. Vi gissar helt enkelt ett värde på  $z$ ,  $u_2(0) = z$  i exemplet. Därmed har vi begynnelsevillkor och kan genom att använda en ODE-lösare (tex Eulers metod) beräkna  $u_1(L)$  som kommer bli en funktion av  $z$ . Låt oss kalla den  $f(z) = u_1(L)$ . För varje  $z$  kan alltså  $f(z)$  beräknas genom att lösa ett begynnelsevärdesproblem. Detta innebär att funktionen  $f(z)$  är definierad implicit. Vi vill hitta ett  $z$  sådant att  $f(z) = 0$  eftersom  $u_1$  och  $u_2$  med begynnelsevillkoren  $u_1(0) = 0$  och  $u_2(0) = z$  då kommer lösa randvärdesproblemet. Att funktionen  $f(z)$  är definierad implicit hindrar oss inte från att använda till exempel bisektion eller Newtons metod för att lösa ekvationen  $f(z) = 0$ . Vi återgår till exemplet med endimensionell värmeledning och visar hur en approximativ lösning kan beräknas med inskjutning.

**Exempel 2.9 (Inskjutning)** Vi låter  $\kappa = q = 1$  i exemplet ovan och använder mittpunktsmetoden med konstant steglängd  $k = 0.01$  och startar med gissningarna  $u_2(0) = z$  med  $z_0 = 0.2$  för att beräkna en första approximation och  $z_1 = 0.9$  för att beräkna en andra. Vi ser resultatet i Figur 2.9. Eftersom  $z_0$  och  $z_1$  är valda så att  $g(z_0) < 0$  och  $g(z_1) > 0$  vet vi att det finns en lösning för något  $z \in [z_0, z_1]$  och att vi kan hitta lösningen med hjälp av bisektion. Vi låter  $z_2 = \frac{z_0+z_1}{2}$  vara mittpunkten och får då  $g(z_2) < 0$ . Det innebär att  $z_2$  blir ny undre gräns och vi beräknar  $z_3 = \frac{z_1+z_2}{2}$ . För varje iteration närmar vi oss målet. Namnet inskjutning kommer just ifrån att vi försöker hitta den bana som leder till målet  $u(L) = 0$ .



**Figur 2.9:** Inskjutningsmetoden med start värden  $z_0 = 0.2$  och  $z_1 = 0.9$ . Mittpunktsmetoden tillsammans med bisektion används för att beräkna de approximativa lösningarna.

## Övningar

### 2.1 Numeriska metoder

**Övning 2.1** Studera ekvationen  $y'(x) = 2y(x)$ , med begynnelsevillkor  $y(0) = 1$ . Vad blir Bakåt Euler approximationen till  $y(1)$ , givet att intervallet  $[0, 1]$  delas in i  $n$  lika stora delintervall?

**Övning 2.2** Studera ekvationen  $y'(x) = y(x)$ , med begynnelsevillkor  $y(0) = 1$ . Vad blir approximation av  $y(1)$  beräknad med Eulers metod, givet att intervallet  $[0, 1]$  delas in i  $n$  lika stora delintervall?

**Övning 2.3** Studera ekvationen  $y'(x) = y(x)^2$ , med begynnelsevillkor  $y(0) = 1$ . Vad blir approximation av  $y(0.2)$  beräknad med Eulers metod, givet att intervallet  $[0, 0.2]$  delas in i 2 lika stora delintervall?

**Övning 2.4** Använd mittpunktsmetoden med konstant steglängd för att beräkna en approximation till  $e = y(1)$  där  $y(x)$  löser ekvationen  $y'(x) = y(x)$ ,  $y(0) = 1$ .

**Övning 2.5** Använd Eulers metod med konstant steglängd  $k = 0.2$  för att beräkna en approximation till  $y'(x) = 1 - y(x)^2$ ,  $y(0) = 0$ , i punkten  $y(0.4)$ .

**Övning 2.6** Använd Eulers metod med konstant steglängd  $k = 0.2$  för att beräkna en approximation till  $y(0.4)$  där  $y(x)$  löser ekvationen  $y'(x) = xy(x)$ ,  $y(0) = 1$ .

**Övning 2.7** Heuns metod för att lösa  $y'(t) = f(t, y(t))$  där  $y(0) = y_0$  är en andra ordningen metod som definieras i två steg,

$$\begin{aligned}\tilde{y}_i &= y_{i-1} + kf(t_{i-1}, y_{i-1}) \\ y_i &= y_{i-1} + \frac{k}{2}(f(t_{i-1}, y_{i-1}) + f(t_i, \tilde{y}_i)),\end{aligned}$$

där  $k$  är steglängden. Är Heuns metod explicit eller implicit?

### 2.2 Konvergens

**Övning 2.8** Studera begynnelsevärdesproblemet  $y'''(t) + y'(t) = \sin(t)$ ,  $y(0) = 1$ ,  $y'(0) = 0$ ,  $y''(0) = 1$ . Skriv som ett system av första ordningen ODE. Ange en numerisk metod som kan användas för att lösa ekvationen samt metodens konvergensordning.

**Övning 2.9** Studera ekvationen  $y'(x) + y(x) = x$  med begynnelsevillkor  $y(0) = 0$ . Lös först ekvationen analytiskt. Dela in  $[0, 1]$  i 100 lika långa delintervall. Ange en uppskattning av felet mellan Euler metod och det exakta lösningen  $y(1)$ .

**Övning 2.10** Studera ekvationen  $y'(x) = \cos(y(x))$  med begynnelsevillkor  $y(0) = 0$ . Beräkna en uppskattning av  $|y''(x)|$ . Dela in  $[0, 2]$  i 100 lika långa delintervall. Ange en uppskattning av felet mellan Eulers metod och det exakta värdet  $y(1)$ .

**Övning 2.11** Studera  $y'(x) = -y(x)$ ,  $y(0) = 1$ . Uppskatta felet i Euler approximationen med steglängd  $k = 0.1$  på intervallet  $t \in [0, T]$  med hjälp av Sats 2.1. Efter hur lång tid  $T$  är feluppskattningen större än 1?

**Övning 2.12** I Exempel 2.3 ger en steglängd på  $\pi/10$  ett fel på 0.00064 med mittpunktsmetoden och 0.012 med Eulers metod. Hur stor steglängd behöver vi använda med Eulers metod för att få samma fel som mittpunktsmetoden?

### 2.3 Stabilitet

**Övning 2.13** Studera ekvationen  $y'(x) = -10y(x)$ ,  $y(0) = 1$ . Vilket steglängd krävs för att Eulers metod respektive bakåt Euler ska vara stabila?

**Övning 2.14** Visa att mittpunktsmetodens stabilitetsområde innehåller alla punkter på den negativ reella axeln, alltså alla  $\lambda k \in \mathbb{R}$  sådana att  $\lambda k < 0$ .

**Övning 2.15** Visa att bakåt Eulers stabilitetsområde innehåller alla punkter på den negativ reella axeln, alltså alla  $\lambda k \in \mathbb{R}$  sådana att  $\lambda k < 0$ .

### 2.4 Generalisering till system

**Övning 2.16** Formulera Eulers metod för systemet

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ x \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Låt steglängden  $k = 0.1$  och beräkna lösningen efter ett steg.

**Övning 2.17** Formulera Eulers metod för systemet

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Låt steglängden  $k = 0.1$  och beräkna lösningen efter ett steg.

**Övning 2.18** Formulera bakåt Euler för systemet

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ x \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Låt steglängden vara  $k$ .

**Övning 2.19** Formulera bakåt Euler för systemet

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

med steglängd  $k$ .

**Övning 2.20** Studera begynnelsevärdesproblemet  $y''(t) + y(t)^2 = \sin(t)$ ,  $y(0) = 1$  och  $y'(0) = 0$ . Skriv som ett system av första ordningen och utför ett steg med Eulers metod, med steglängd  $k = 0.1$ .

**Övning 2.21** Studera begynnelsevärdesproblemet  $y''(t) - 4y(t) = 0$ ,  $y(0) = 1$ ,  $y'(0) = 0$ . Skriv som ett system av första ordningen ODE på matrisform. Utför ett steg av Bakåt Euler metoden med steglängd  $k = 0.1$ .

## 2.5 Randvärdesproblem

**Övning 2.22** Studera Exempel 2.8. Låt  $\kappa = L = 1$  och välj  $q(x)$  sådant att lösningen blir  $u(x) = x^2(1 - x)$ .

**Övning 2.23** Studera Exempel 2.8. Låt  $\kappa = L = 1$  och välj  $q(x)$  sådant att lösningen blir  $u(x) = \sin(\pi x)$ .

**Övning 2.24** Låt  $q(x) = x^2$  och  $L = \kappa = 1$  i Exempel 2.8. Räkna ut  $u(x)$  genom att integrera ekvationen två gånger och använda randvillkoren.

**Övning 2.25** Låt  $q(x) = \sin(\pi x)$  och  $L = \kappa = 1$  i Exempel 2.8. Räkna ut  $u(x)$ .

## Problem

### 2.1 Numeriska metoder

**Problem 2.1** Studera  $y'(t) = y^2$  där  $y(0) = 1$ . Låt  $k = 0.1$  och beräkna  $y_2$  med Heuns metod från Övning 2.7.

### 2.2 Konvergens

**Problem 2.2** Bevisa att Bakåt Euler är en första ordningens metod.

### 2.3 Stabilitet

**Problem 2.3** Ge ett uttryck för stabilitetsområdet för Heuns metod presenterad i Ö2.7.



## 2.4 Generalisering till system

**Problem 2.4** Skriv som system av första ordningen och genomför en iteration med bakåt Euler metoden för differentialekvationen  $y''(x) + y'(x) = 1 + x$ ,  $y(0) = 1$ ,  $y'(0) = 0$ , med steglängd  $k = 0.1$ .

## 2.5 Randvärdesproblem

**Problem 2.5** Visa att  $\int_0^L u(x) dx \geq 0$  om  $q \geq 0$  är en konstant och  $\kappa = 1$  i Exempel 2.8.  
*Hint:* Multiplicera ekvationen med  $u(x)$  och integrera från 0 till  $L$ .



# 3. Laplacetransform

3.1	Definition av Laplacetransform	47
3.2	Skalning	51
3.3	Laplacetransform av derivator och integraler	52
3.4	Faltning	54
3.5	Impulser och diskontinuerliga funktioner	56
3.6	Lösning av ODE med Laplacetransform	58

*Laplacetransform är en analytisk metod för att lösa ordinära differentialekvationer och integralekvationer. Till skillnad från andra analytiska tekniker vi studerat kan Laplace-transformen hantera diskontinuerliga data och impulser, som är vanligt förekommande i tillämpningar inom elektriska nät och mekaniska system. Laplacetransformen transformerar differentialekvationen till en algebraisk ekvation, med en ny variabel, som är enklare att lösa. Sedan transformeras lösningen tillbaka till den ursprungliga variabeln som är tiden.*

## 3.1 Definition av Laplacetransform

Laplacetransformen är en integraltransform med vilken ett givet problem transformeras från tidsvariabeln  $t$  till en ny variabel  $s$ . Poängen är att integral- och differentialekvationer är enklare att lösa på transformsidan eftersom de transformeras till algebraiska ekvationer. När lösningen är bestämd kan man sedan gå tillbaka till tidsdomänen med hjälp av invers Laplacetransform.

För att ge en motivering till hur och varför Laplacetransform fungerar börjar vi med ett exempel.

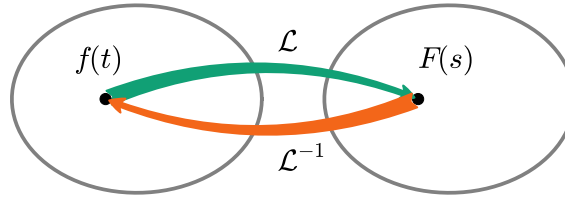
**Exempel 3.1 (Motiverande exempel)** Vi vill lösa begynnelsevärdesproblemet

$$y'(t) + y(t) = e^{-2t}, \quad t > 0,$$

med begynnelsevillkoret  $y(0) = 0$ . Vi har tidigare använt integrerande faktor för att lösa denna ekvation. Nu vill vi istället introducera en ny metod. Vi multiplicerar ekvationen med en funktion  $e^{-st}$  och integrerar från 0 till  $\infty$ . Vi antar tillsvidare att integralerna är väldefinierade, vilket medför att  $\lim_{t \rightarrow \infty} (|y'(t)| + |y(t)|)e^{-st} = 0$ , om vi väljer  $s > 0$ . Vi får

$$\int_0^{\infty} y'(t)e^{-st} dt + \int_0^{\infty} y(t)e^{-st} dt = \int_0^{\infty} e^{-2t} e^{-st} dt.$$

Vi integrerar den första termen partiellt och använder att den övre gränsen blir noll på grund av att  $\lim_{t \rightarrow \infty} |y(t)|e^{-st} = 0$  och den undre gränsen blir noll eftersom  $y(0) =$



**Figur 3.1:** Laplacetransformen  $\mathcal{L}$  transformerar funktioner uttryckta i  $t$  till en ny variabel  $s$ . Den inversa Laplacetransformen  $\mathcal{L}^{-1}$  transformerar tillbaka till  $t$ .

0. Detta ger

$$s \int_0^{\infty} y(t)e^{-st} dt + \int_0^{\infty} y(t)e^{-st} dt = \left[ \frac{-1}{s+2} e^{-st} \right]_0^{\infty} = \frac{1}{s+2}.$$

Vi inför följande beteckning  $Y(s) = \int_0^{\infty} y(t)e^{-st} dt$  och ser vi att vi får en algebraisk ekvation som bestämmer  $Y(s)$ ,

$$(s+1)Y(s) = \frac{1}{s+2},$$

eller med hjälp av partialbråksuppdelning,

$$Y(s) = \frac{1}{s+1} \cdot \frac{1}{s+2} = \frac{A}{s+1} + \frac{B}{s+2} = \frac{1}{s+1} - \frac{1}{s+2},$$

eftersom vi får att  $A+B=0$  och  $2A+B=1$  med lösning  $A=1$  och  $B=-1$  när vi gör liknämngt. Frågan är nu om vi kan hitta en funktion  $y(t)$  sådan att

$$\int_0^{\infty} y(t)e^{-st} dt = \frac{1}{s+1} - \frac{1}{s+2}.$$

Vi vet redan att  $\int_0^{\infty} e^{-2t}e^{-st} dt = \frac{1}{s+2}$ . På samma sätt fås  $\int_0^{\infty} e^{-t}e^{-st} dt = \frac{1}{s+1}$ . Därför har vi också att

$$\int_0^{\infty} (e^{-t} - e^{-2t})e^{-st} dt = \left[ \frac{-1}{s+1} e^{-(s+1)t} - \frac{-1}{s+2} e^{-(s+2)t} \right]_0^{\infty} = \frac{1}{s+1} - \frac{1}{s+2}.$$

Vi noterar här att så länge  $s > 0$  så konvergerar alla integraler i exemplet eftersom  $\lim_{t \rightarrow \infty} e^{-st} = 0$ . Lösningen ges alltså av  $y(t) = e^{-t} - e^{-2t}$ .

I exemplet ser vi att vi transformerar en differentialekvation i variabeln  $t$  till en algebraisk ekvation i variabeln  $s$ . Vi löser sedan den algebraiska ekvationen och transformerar tillbaka till  $t$ .

Efter denna motivering är vi redo att definiera Laplacetransformen. Variabeln  $s$  kommer i definitionen tillåtas vara komplex. Laplacetransformen av en funktion  $f(t)$  är en generaliserad integral med avseende på tiden  $t$  från 0 till  $\infty$  av funktionen viktad med  $e^{-st}$  där  $s \in \mathbb{C}$ .

**Definition 3.1 (Laplacetransform)** Låt  $f(t) : [0, \infty) \rightarrow \mathbb{R}$  vara begränsad av  $|f(t)| \leq Ce^{at}$  för några konstanter  $C$  och  $a$ . Laplacetransformen av  $f(t)$  ges av

$$F(s) = \int_0^{\infty} e^{-st} f(t) dt, \quad s \in \mathbb{C}, \quad \operatorname{Re}(s) > a. \quad (3.1)$$

Vi använder två olika notationer för Laplacetransformen av en funktion  $f(t)$ . Dels  $F(s)$  som ovan och dels  $\mathcal{L}(f(t))(s)$ . Den första är kortare och att föredra om man jobbar med funktioner  $f(t)$  och  $g(t)$ . Den senare är att föredra om man har explicita uttryck  $t^2$ ,  $\sin(t)$  som till exempel i tabeller eller som data i ekvationer.

Laplacetransformer är väldefinierad under förutsättning att integralen är konvergent. I följande sats visar vi varför villkoret  $|f(t)| \leq Ce^{at}$  leder till en väldefinierad Laplacetransform för  $\operatorname{Re}(s) > a$ .

**Sats 3.1 (Väldefinierad Laplacetransform)** Låt  $f(t) : [0, \infty) \rightarrow \mathbb{R}$  vara begränsad av  $|f(t)| \leq Ce^{at}$  för några konstanter  $C, a \geq 0$ . Laplacetransformen av  $f(t)$  är väldefinierad för  $s$  sådana att  $\operatorname{Re}(s) > a$ .

*Bevis.* Låt  $s = x + iy$  där  $x, y$  är reella och  $x > a$ . Vi får

$$|F(s)| = \left| \int_0^{\infty} e^{-st} f(t) dt \right| \quad (3.2)$$

$$\leq \int_0^{\infty} |e^{-st} f(t)| dt \quad (3.3)$$

$$= \int_0^{\infty} |e^{-xt}| |e^{-iyt}| |f(t)| dt \quad (3.4)$$

$$= \int_0^{\infty} |e^{-xt}| |f(t)| dt \quad (3.5)$$

$$\leq C \int_0^{\infty} e^{(a-x)t} dt \quad (3.6)$$

$$= \lim_{R \rightarrow \infty} C \left[ \frac{e^{(a-x)t}}{a-x} \right]_0^R \quad (3.7)$$

$$= \frac{C}{x-a} \quad (3.8)$$

$$< \infty, \quad (3.9)$$

eftersom  $x > a$ . Notera att  $|e^{ix}| = |\cos(x) + i \sin(x)| = \sqrt{\cos^2(x) + \sin^2(x)} = 1$ .  $\square$

I Exempel 3.1 gäller att Laplacetransformerna är väldefinierade eftersom vi kan välja  $C = 1$  och  $a = 0$ . Under förutsättning att  $|f(t)| \leq Ce^{at}$  med  $\operatorname{Re}(s) > a$  kommer bidraget från den övre gränsen i den generaliserade integralen försvinna vid evaluering av primitivfunktionen.

**Exempel 3.2** Låt  $f(t) = e^t$ . Därmed uppfyller  $|f(t)| \leq Ce^{at}$  med  $C = a = 1$ . Vi låter  $\operatorname{Re}(s) > 1$  och får

$$F(s) = \int_0^{\infty} e^{-st} e^t dt = \int_0^{\infty} e^{(1-s)t} dt = \left[ \frac{1}{1-s} e^{(1-s)t} \right]_0^{\infty} = \frac{1}{s-1}, \quad (3.10)$$

eftersom om vi låter  $s = x + iy$  med  $x > 1$  så gäller att

$$\lim_{t \rightarrow \infty} |e^{(1-s)t}| = \lim_{t \rightarrow \infty} |e^{-iyt}| |e^{(1-x)t}| = \lim_{t \rightarrow \infty} e^{(1-x)t} = 0. \quad (3.11)$$

**Exempel 3.3** Låt  $f(t) = 1$ . Därmed uppfyller  $|f(t)| \leq Ce^{at}$  med  $C = 1$  och  $a = 0$ . För  $\operatorname{Re}(s) > 0$  får vi

$$F(s) = \int_0^{\infty} e^{-st} t dt = \left[ \frac{1}{-s} e^{-st} \right]_0^{\infty} = \frac{1}{s}. \quad (3.12)$$

Den övre integrationsgränsen ger inget bidrag eftersom för  $s = x + iy$  med  $x > 0$  gäller att

$$\lim_{t \rightarrow \infty} |e^{-st}| = \lim_{t \rightarrow \infty} |e^{-iyt}| |e^{-xt}| = \lim_{t \rightarrow \infty} e^{-xt} = 0. \quad (3.13)$$

Laplacetransformen är linjär vilket innebär att givet två tal  $\alpha, \beta \in \mathbb{R}$  och två funktioner  $f(t)$  och  $g(t)$ , sådana att Laplacetransformen är väldefinierad, gäller

$$\mathcal{L}(\alpha f(t) + \beta g(t))(s) = \int_0^{\infty} e^{-st} (\alpha f(t) + \beta g(t)) dt \quad (3.14)$$

$$= \alpha \int_0^{\infty} e^{-st} f(t) dt + \beta \int_0^{\infty} e^{-st} g(t) dt \quad (3.15)$$

$$= \alpha \mathcal{L}(f(t))(s) + \beta \mathcal{L}(g(t))(s). \quad (3.16)$$

Det finns även en invers Laplacetransform definierad på följande vis.

**Definition 3.2 (Invers Laplacetransform)** Givet  $F(s)$  definierad för  $\operatorname{Re}(s) > a$  gäller,

$$f(t) = \frac{1}{2\pi i} \lim_{R \rightarrow \infty} \int_{-R}^R F(a + i\omega) e^{(a+i\omega)t} d\omega. \quad (3.17)$$

För att använda den inversa formeln krävs ofta djupare kunskaper om integration i det komplexa talplanet än vad som ingår i denna framställning. Därför nöjer vi oss med att konstatera att en invers finns. Det går även att visa att Laplacetransformen är entydig upp till variationer i enskilda punkter. På grund av entydigheten kan vi använda Laplacetransformen för att beräkna tabeller och sedan hitta den inversa transformen som Laplacetransformen till ett känt uttryck. Detta tillvägagångssätt är analogt med att beräkna primitivfunktioner genom att studera tabeller av derivator.

## 3.2 Skalning

Det finns ett antal tekniker som förenklar beräkningen av Laplacetransformer. Vi introducerar först en teknik som kallas skalning.

**Sats 3.2 (Skalning)** Låt  $f(t)$  vara en given funktion med Laplacetransform  $F(s)$  och  $\alpha > 0$  vara ett givet tal. Vi bildar  $g(t) = f(\alpha t)$ . Då gäller att Laplacetransformen  $G(s)$  av  $g(t)$  ges av  $G(s) = \frac{1}{\alpha} F\left(\frac{s}{\alpha}\right)$ .

*Bevis.* Vi har

$$G(s) = \int_0^{\infty} f(\alpha t) e^{-st} dt = \left\{ \tau = \alpha t, \frac{d\tau}{dt} = \alpha \right\} \quad (3.18)$$

$$= \frac{1}{\alpha} \int_0^{\infty} f(\tau) e^{-\frac{s}{\alpha} \tau} d\tau \quad (3.19)$$

$$= \frac{1}{\alpha} F\left(\frac{s}{\alpha}\right). \quad (3.20)$$

□

**Exempel 3.4 (Skalning)** Låt  $g(t) = e^{3t}$ . Vi vet att Laplacetransformen till  $f(t) = e^t$  ges av  $F(s) = \frac{1}{s-1}$ . Sats 3.2 ger därför att  $G(s) = \frac{1}{3} \frac{1}{\frac{s}{3}-1} = \frac{1}{s-3}$ .

Här skalas alltså tidsvariabeln om mellan  $f(t)$  och  $g(t)$ . Nästa teknik kallas exponentiell skalning.

**Sats 3.3 (Exponentiell skalning)** Låt  $f(t)$  vara en given funktion med Laplacetransform  $F(s)$  och  $\alpha > 0$  vara ett givet tal. Vi bildar  $g(t) = e^{\alpha t} f(t)$ . Då gäller att Laplacetransformen  $G(s)$  av  $g(t)$  ges av  $G(s) = F(s - \alpha)$ .

*Bevis.* Vi har

$$G(s) = \int_0^{\infty} e^{\alpha t} f(t) e^{-st} dt = \int_0^{\infty} f(t) e^{-(s-\alpha)t} dt = F(s - \alpha).$$

□

**Exempel 3.5 (Exponentiell skalning)** Låt  $g(t) = e^{\alpha t}$ . Vi använder nu istället exponentiell skalning med. Vi har  $g(t) = e^{\alpha t} f(t)$ , där  $f(t) = 1$  med Laplacetransform  $F(s) = \frac{1}{s}$ . Vi får genom att använda exponentiell skalning  $G(s) = F(s - \alpha) = \frac{1}{s - \alpha}$ .

Kan vi Laplacetransformera  $f(t)$  går alltså  $e^{\alpha t} f(t)$  enkelt att transformera. Det finns andra tekniker också men vi nöjer oss för tillfället för att gå vidare till Laplacetransform av derivata och integral som är huvudpoängen med Laplacetransform.

### 3.3 Laplacetransform av derivator och integraler

Laplacetransformen används för att lösa differential- och integralekvationer. Anledningen till att Laplacetransformen lämpar sig för dessa problem är att Laplacetransformen av både derivator och integraler är enkel att räkna ut.

**Sats 3.4 (Laplacetransform av derivata)** Låt  $f(t)$  vara en deriverbar funktion sådan att  $|f(t)| \leq Ce^{at}$ . Låt vidare Laplacetransformen av  $f(t)$  vara  $F(s)$ . Då gäller att  $\mathcal{L}(f'(t))(s) = sF(s) - f(0)$ .

*Bevis.* Vi använder partiell integration för att få,

$$\mathcal{L}(f'(t))(s) = \int_0^{\infty} e^{-st} f'(t) dt \quad (3.21)$$

$$= [e^{-st} f(t)]_0^{\infty} + s \int_0^{\infty} e^{-st} f(t) dt \quad (3.22)$$

$$= -f(0) + sF(s), \quad (3.23)$$

om  $Re(s) > a$  eftersom med  $s = x + iy$  gäller att

$$\lim_{t \rightarrow \infty} |e^{-(x+iy)t} f(t)| \leq \lim_{t \rightarrow \infty} |e^{-xt}| |e^{iyt}| |f(t)| \leq \lim_{t \rightarrow \infty} e^{(a-x)t} = 0,$$

då  $x = Re(s) > a$ . □

**Exempel 3.6 (Laplacetransform av högre ordningens derivata)** Genom att upprepa formeln för Laplace transform av derivata kan vi härleda en formel för högre ordningens derivata. Givet att  $f(t)$  är två gånger deriverbar gäller

$$\mathcal{L}(f''(t))(s) = s\mathcal{L}(f'(t))(s) - f'(0) = s^2\mathcal{L}(f(t))(s) - sf(0) - f'(0).$$

för tredjederivatans gäller,

$$\mathcal{L}(f'''(t))(s) = s\mathcal{L}(f''(t))(s) - f''(0) = s^3\mathcal{L}(f(t))(s) - s^2f(0) - sf'(0) - f''(0).$$

Den allmänna formeln ges av följande sats.

**Sats 3.5 (Laplacetransform av högre ordningens derivata)** Låt  $f(t)$  vara  $n$  gånger kontinuerligt deriverbar. Låt vidare Laplacetransformen av  $f(t)$  vara  $F(s)$ . Då gäller att  $\mathcal{L}(f^{(n)}(t))(s) = s^n F(s) - \sum_{i=1}^n s^{n-i} f^{(i-1)}(0)$ .

*Bevis.* Formeln för Laplacetransform av derivata ger att satsen gäller för  $n = 1$ . Vi antar nu att den gäller för  $n - 1$  alltså  $\mathcal{L}(f^{(n-1)}(t))(s) = s^{n-2}F(s) - \sum_{i=1}^{n-1} s^{n-1-i} f^{(i-1)}(0)$ . Vi låter vidare



$g(t) = f^{(n-1)}(t)$ . Vi får  $G(s) = s^{n-2}F(s) - \sum_{i=1}^{n-1} s^{n-1-i} f^{(i-1)}(0)$  och  $g(0) = f^{(n-1)}(0)$  så

$$\mathcal{L}(f^n(t))(s) = \mathcal{L}(g'(t))(s) \quad (3.24)$$

$$= sG(s) - g(0) \quad (3.25)$$

$$= s \cdot s^{n-2}F(s) - s \sum_{i=1}^{n-1} s^{n-1-i} f^{(i-1)}(0) - f^{(n-1)}(0) \quad (3.26)$$

$$= s^n F(s) - \sum_{i=1}^n s^{n-i} f^{(i-1)}(0). \quad (3.27)$$

Satsen följer nu genom induktion. □

Derivata motsvarar alltså multiplikation med  $s$  på transformsidan. För integralen, som är derivatans invers, gäller att Laplacetransformen istället motsvarar division med  $s$ .

**Sats 3.6 (Laplacetransform av integral)** Låt  $f(t)$  vara en integrerbar funktion sådan att  $|f(t)| \leq Ce^{at}$ . Låt vidare Laplacetransformen av  $f(t)$  vara  $F(s)$ . Vi bildar  $g(t) = \int_0^t f(\tau) d\tau$  och antar vidare att  $|g(t)| \leq Ce^{at}$ . Då gäller att  $\mathcal{L}(g(t))(s) = \frac{1}{s}F(s)$  under förutsättning att  $Re(s) > a$ .

*Bevis.* Vi använder partiell integration för att få

$$\mathcal{L}(g(t))(s) = \int_0^\infty e^{-st} \int_0^t f(\tau) d\tau dt \quad (3.28)$$

$$= \left[ \frac{-1}{s} e^{-st} g(t) \right]_0^\infty + \frac{1}{s} \int_0^\infty e^{-st} g'(t) dt \quad (3.29)$$

$$= \frac{1}{s} F(s), \quad (3.30)$$

eftersom  $g(0) = 0$  och gränsvärdet i  $\infty$  också blir 0. Detta ses genom att anta  $Re(s) > a$ , låta  $s = x + iy$  och studera

$$\lim_{t \rightarrow \infty} |e^{-(x+iy)t} g(t)| \leq \lim_{t \rightarrow \infty} |e^{-xt}| |e^{iyt}| |g(t)| \leq \lim_{t \rightarrow \infty} e^{(a-x)t} = 0,$$

då  $x = Re(s) > a$ . □

**Exempel 3.7** Låt  $f(t) = 1$  och  $g(t) = \int_0^t 1 d\tau = t$ . Då ges Laplacetransformen av  $t$  av

$$\mathcal{L}(t)(s) = \mathcal{L}\left(\int_0^t 1 d\tau\right)(s) = \frac{1}{s^2},$$

eftersom  $\mathcal{L}(1) = \frac{1}{s}$ . På samma sätt gäller att,

$$\mathcal{L}(t^2)(s) = \mathcal{L}\left(\int_0^t 2\tau d\tau\right)(s) = \frac{2}{s^3},$$

och genom upprepning,

$$\mathcal{L}(t^n)(s) = \frac{n!}{s^{n+1}}.$$

Derivering på transformsidan motsvaras i sin tur av multiplikation med  $-t$  i tidsdomänen.

**Sats 3.7 (Multiplikation med  $t$ )** Låt  $f(t)$  vara en given funktion med Laplacetransform  $F(s)$ . Då gäller att  $g(t) = tf(t)$  har Laplacetransform  $G(s) = -F'(s)$ .

*Bevis.* Vi har

$$F'(s) = \frac{d}{ds} \int_0^{\infty} e^{-st} f(t) dt = - \int_0^{\infty} t e^{-st} f(t) dt = -G(s).$$

□

**Exempel 3.8** Vi bestämmer Laplacetransformen av  $t^k e^{-t}$ . Genom att använda Sats 3.7 får vi att

$$\mathcal{L}(te^{-t})(s) = -\frac{d}{ds} \frac{1}{s+1} = \frac{1}{(s+1)^2}.$$

På samma sätt gäller

$$\mathcal{L}(t^2 e^{-t})(s) = -\frac{d}{ds} \frac{1}{(s+1)^2} = \frac{2}{(s+1)^3},$$

och i allmänhet

$$\mathcal{L}(t^k e^{-t})(s) = \frac{k!}{(s+1)^{k+1}}.$$

## 3.4 Faltning

Man skulle kunna tro att Laplacetransformen av produkten av två funktioner är lika med produkten av Laplacetransformerna men så är det inte. Istället gäller att Laplacetransformen av en speciell integral, faltningsintegralen, mellan två funktioner är lika med produkten av Laplacetransformerna. När man löser ordinära differentialekvationer med Laplacetransform är det viktigt att kunna transformera tillbaka produkter av kända Laplacetransformer. Därför är faltning ett viktigt begrepp i detta kapitel.

**Definition 3.3 (Faltning)** Faltningen av två funktioner,  $f(t)$  och  $g(t)$ , ges av  $(f * g)(t) := \int_0^t f(t-u)g(u) du$ .

Faltningen mellan två funktioner är kommutativ vilket innebär att,

$$(f * g)(t) = \int_0^t f(u)g(t-u) du = \{\bar{t} = t-u\} = - \int_t^0 f(t-\bar{t})g(\bar{t}) d\bar{t} = (g * f)(t).$$

**Sats 3.8 (Laplacetransform av faltning)** För två kontinuerliga funktioner  $f(t)$  och  $g(t)$  som uppfyller  $|f(t)| + |g(t)| \leq Ce^{at}$  gäller,

$$\mathcal{L}((f * g)(t))(s) = \mathcal{L}(f(t))(s) \cdot \mathcal{L}(g(t))(s).$$

*Bevis.* Vi låter  $h(t) = (f * g)(t) = \int_0^t f(u)g(t-u) du$  och betecknar dess Laplacetransform  $H(s)$ . Vi får då att

$$H(s) = \int_0^\infty e^{-st} h(t) dt \quad (3.31)$$

$$= \int_0^\infty e^{-st} \int_0^t f(u)g(t-u) du dt \quad (3.32)$$

$$= \int_0^\infty \int_0^t e^{-st} f(u)g(t-u) du dt = \{ \text{vi byter integrationsordning} \} \quad (3.33)$$

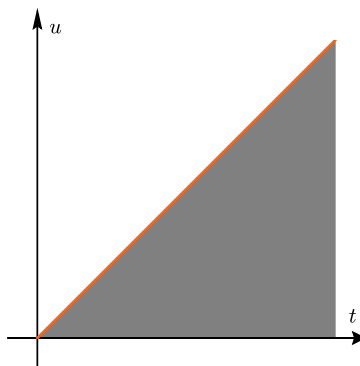
$$= \int_0^\infty \int_u^\infty e^{-st} f(u)g(t-u) dt du = \{ \bar{t} = t - u \} \quad (3.34)$$

$$= \int_0^\infty \int_0^\infty e^{-s(u+\bar{t})} f(u)g(\bar{t}) d\bar{t} du \quad (3.35)$$

$$= \int_0^\infty e^{-su} f(u) du \int_0^\infty e^{-s\bar{t}} g(\bar{t}) d\bar{t} \quad (3.36)$$

$$= \mathcal{L}(f(t))(s) \mathcal{L}(g(t))(s). \quad (3.37)$$

I Figur 3.2 ser vi integrationsområdet i  $tu$ -planet. När vi byter ordning på integralerna kommer gränserna istället gå från 0 till  $\infty$  i  $u$  och från  $u$  till  $\infty$  i  $t$ .  $\square$



**Figur 3.2:** Byte av integrationsordning.

**Exempel 3.9** Bestäm den inversa Laplacetransformen till  $H(s) = \frac{1}{s^2(s+1)}$ . Vi vet sedan tidigare att  $\mathcal{L}(t) = s^{-2}$  och  $\mathcal{L}(e^{-t}) = (s+1)^{-1}$  så  $H(s) = F(s)G(s)$  med  $f(t) = t$  och  $g(t) =$

$e^{-t}$ . Vi använder Sats 3.8 och får

$$h(t) = (f * g)(t) \quad (3.38)$$

$$= \int_0^t (t-u)e^{-u} du \quad (3.39)$$

$$= t \int_0^t e^{-u} du - \int_0^t ue^{-u} du \quad (3.40)$$

$$= t(1 - e^{-t}) + [ue^{-u}]_0^t - \int_0^t e^{-u} du \quad (3.41)$$

$$= t(1 - e^{-t}) + te^{-t} - (1 - e^{-t}) = e^{-t} + t - 1. \quad (3.42)$$

### 3.5 Impulser och diskontinuerliga funktioner

I tillämpningar är det viktigt att kunna hantera diskontinuerliga funktioner och impulser som t.ex. uppkommer när en elektrisk krets slås på. För att kunna uttrycka denna typ av fenomen matematiskt återkommer vi till Heavisides stegfunktion.

**Definition 3.4 (Heavisides stegfunktion)** Heavisides stegfunktion  $\theta(t) = 1$  för  $t > 0$ ,  $\theta(t) = \frac{1}{2}$  för  $t = 0$  och  $\theta(t) = 0$  för  $t < 0$ .

Eftersom  $\theta(t) = 0$  för  $t < 0$  har den Laplacetransform  $\mathcal{L}(\theta(t))(s) = \mathcal{L}(1)(s) = \frac{1}{s}$ . Diskontinuiteten i 0:an påverkar inte integralen eftersom vi har definierat integralen för styckvis kontinuerliga funktioner och integralens värde påverkas inte av funktionens värde i en enskild punkt.

Genom definitionen av Heavisides stegfunktion kan vi även studera fördröjning. I tillämpningar innebär fördröjning t.ex. att en spänning slås på vid en given tidpunkt  $T$  i framtiden.

**Sats 3.9 (Fördröjning)** Låt  $f(t)$  vara en given funktion med Laplacetransform  $F(s)$ ,  $T > 0$  vara ett tal och  $\theta(t)$  vara Heavisides stegfunktion. Då gäller

$$\mathcal{L}(f(t-T)\theta(t-T))(s) = e^{-sT}F(s).$$

*Bevis.* Vi har

$$\mathcal{L}(f(t-T)\theta(t-T))(s) = \int_0^\infty e^{-st} f(t-T)\theta(t-T) dt \quad (3.43)$$

$$= \int_T^\infty e^{-st} f(t-T) dt = \int_T^\infty e^{-s\tau} f(\tau) d\tau \quad (3.44)$$

$$= \int_0^\infty e^{-s(\tau+T)} f(\tau) d\tau = e^{-sT}F(s). \quad (3.45)$$

□

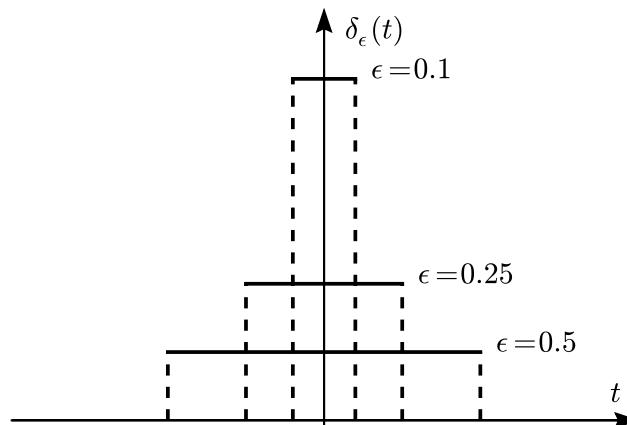
**Exempel 3.10** För att bestämma Laplacetransformen av den funktion som beskriver att en spänning av storlek  $E$  slås på vid tiden  $t = 1$ , kan vi använda Heavisides stegfunktion och Sats 3.9. Vi har  $f(t) = E\theta(t - 1)$ . Vi får

$$\mathcal{L}(E\theta(t - 1))(s) = E \frac{e^{-s}}{s}.$$

Ett annat fenomen som är vanligt i tillämpningar är krafter som verkar med stor styrka under kort tid, så kallade impulser. Om vi låter  $\epsilon > 0$  vara ett litet tal så definierar vi den diskontinuerliga funktionen  $\delta_\epsilon(t)$  på följande vis,

$$\delta_\epsilon(t) = \begin{cases} 0, & t \leq -\epsilon \\ \frac{1}{2\epsilon}, & -\epsilon < t < \epsilon, \\ 0, & t \geq \epsilon. \end{cases}$$

I Figur 3.3 ser vi hur funktionen  $\delta_\epsilon(t)$  för mindre och mindre värden på  $\epsilon$  växer mot en impuls i origo men att integralen förblir konstant  $\int_{-\infty}^{\infty} \delta_\epsilon(t) dt = 1$  för alla  $\epsilon$ . Vi har även att  $\lim_{\epsilon \rightarrow 0} \delta_\epsilon(t) = 0$  för alla  $t \neq 0$ . Vi vill studera den generaliserade funktion  $\delta$  som uppkommer då vi låter  $\epsilon \rightarrow 0$ .



**Figur 3.3:** Funktionen  $\delta_\epsilon(t)$  där  $\epsilon = 0.5, 0.25, 0.1$ .

**Definition 3.5 (Diracs delta)** Diracs delta är en idealiserad enhetspuls i  $t = 0$  sådan att  $\delta(t) = 0$  för alla  $t \neq 0$  och  $\int_{-\infty}^{\infty} \delta(t) dt = 1$ .

En impuls i tidpunkt  $t_0 > 0$  uttrycks som  $\delta(t - t_0)$ . En viktig egenskap för Diracs delta

är att givet en kontinuerlig funktion  $f(t)$  gäller att,

$$\begin{aligned} \int_0^{\infty} \delta(t-t_0)f(t) dt &= \lim_{\epsilon \rightarrow 0} \int_0^{\infty} \delta(t-t_0)f(t) dt \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_{t_0-\epsilon}^{t_0+\epsilon} f(t) dt \\ &= \lim_{\epsilon \rightarrow 0} f(t_0 + c_\epsilon) \\ &= f(t_0). \end{aligned}$$

Vi använder medelvärdessatsen för integraler som säger att  $\int_{t_0-\epsilon}^{t_0+\epsilon} f(t) dt = 2\epsilon f(t_0 + c_\epsilon)$  för något  $c_\epsilon \in [-\epsilon, \epsilon]$ . Integralen av produkten mellan  $\delta(t-t_0)$  och en kontinuerlig funktion  $f(t)$  blir alltså funktionens värde i  $t_0$ . Denna egenskap brukar användas för att definiera Diracs delta. Diracs delta uppfyller inte de krav vi har satt upp på funktioner för att de ska kunna ha en Laplacetransform, nämligen  $|f(t)| \leq Ce^{at}$ . Man kan ändå definiera Laplacetransformen av  $\delta$  genom att använda  $\delta_\epsilon$  och gå i gräns.

**Sats 3.10 (Laplacetransform av Dirac delta)** Vi har att

$$\mathcal{L}(\delta(t-t_0))(s) = e^{-st_0},$$

för alla  $t_0 \geq 0$ . Speciellt gäller  $\mathcal{L}(\delta(t))(s) = 1$ .

*Bevis.* Vi antar först att  $t_0 > 0$  och  $\epsilon < t_0$  och studerar Laplacetransformen av  $\delta_\epsilon(t-t_0)$ . Vi har

$$\mathcal{L}(\delta_\epsilon(t-t_0))(s) = \int_0^{\infty} e^{-st} \delta_\epsilon(t-t_0) dt \quad (3.46)$$

$$= \int_{t_0-\epsilon}^{t_0+\epsilon} \frac{e^{-st}}{2\epsilon} dt \quad (3.47)$$

$$= \frac{1}{2\epsilon s} e^{-st_0} (e^{s\epsilon} - e^{-s\epsilon}). \quad (3.48)$$

Vi kan nu låta  $\epsilon \rightarrow 0^+$ . L'Hopitals regel ger att  $\lim_{\epsilon \rightarrow 0^+} \frac{e^{s\epsilon} - e^{-s\epsilon}}{2s\epsilon} = 1$ . Därför gäller

$$\mathcal{L}(\delta_\epsilon(t-t_0))(s) = e^{-st_0}.$$

Vi utvidgar nu resultatet till  $t_0 = 0$  genom att gå i gräns,

$$\mathcal{L}(\delta(t))(s) = \lim_{t_0 \rightarrow 0^+} e^{-st_0} = 1.$$

□

## 3.6 Lösning av ODE med Laplacetransform

Vi går nu tillbaka till det ursprungliga problemet att lösa ODE som motiverade införandet av Laplacetransformen. Metoden fungerar om koefficienterna till ekvationen är konstan-

ta. Fördelen med Laplacetransform är att den ger ett systematiskt sätt att beräkna partikulärlösningar utan att man behöver ansätta en lösning. Dessutom kan vi hantera diskontinuerliga högerled och impulser. Vid lösning med Laplacetransform följer vi fyra steg som vi redan såg i det första motiverande exemplet. Principen är följande:

1. Laplacetransformera differentialekvationen.
2. Sätt in begynnelsevillkoren.
3. Lös den algebraiska ekvation som uppkommer. Här behövs ofta partialbråksuppdelning användas.
4. Använd tabell för att beräkna den inversa Laplacetransformen som ger lösningen till differentialekvationen.

I nedanstående sats sammanfattar vi alla Laplacetransformer från exemplen i boken.

**Sats 3.11 (Tabell av Laplacetransformer)** Låt  $F(s)$  och  $G(s)$  vara Laplacetransformer till  $f(t)$  respektive  $g(t)$ . Då gäller,

$$\mathcal{L}(\alpha f(t) + \beta g(t))(s) = \alpha F(s) + \beta G(s) \quad (3.49)$$

$$\mathcal{L}(f(at))(s) = \frac{1}{a} F\left(\frac{s}{a}\right) \quad (3.50)$$

$$\mathcal{L}(f(t-T)\theta(t-T))(s) = e^{-sT} F(s) \quad (3.51)$$

$$\mathcal{L}(\delta(t))(s) = 1 \quad (3.52)$$

$$\mathcal{L}(\theta(t))(s) = \frac{1}{s} \quad (3.53)$$

$$\mathcal{L}(t^n)(s) = \frac{n!}{s^{n+1}} \quad (3.54)$$

$$\mathcal{L}(t^n e^{-t})(s) = \frac{n!}{(s+1)^{n+1}} \quad (3.55)$$

$$\mathcal{L}(e^{-at})(s) = \frac{1}{s+a} \quad (3.56)$$

$$\mathcal{L}(\sin(at))(s) = \frac{a}{s^2 + a^2} \quad (3.57)$$

$$\mathcal{L}(\cos(at))(s) = \frac{s}{s^2 + a^2} \quad (3.58)$$

$$(3.59)$$

*Bevis.* Alla utom de sista två ges som exempel i boken. Vi låter  $f(t) = \sin(at)$ . Definitionen av den komplexa exponentialfunktionen  $e^{iat} = \cos(at) + i \sin(at)$  kan användas för att

förenkla uträkningen. Vi har  $\sin(at) = \frac{e^{iat} - e^{-iat}}{2i}$  vilket ger

$$F(s) = \int_0^{\infty} e^{-st} \frac{e^{iat} - e^{-iat}}{2i} dt \quad (3.60)$$

$$= \frac{1}{2i} \int_0^{\infty} e^{(ia-s)t} - e^{(-ia-s)t} dt \quad (3.61)$$

$$= \frac{1}{2i} \left[ \frac{1}{ia-s} e^{(ia-s)t} + \frac{1}{s+ia} e^{-(s+ia)t} \right]_0^{\infty} \quad (3.62)$$

$$= \frac{1}{2i} \left( \frac{1}{s-ia} - \frac{1}{s+ia} \right) \quad (3.63)$$

$$= \frac{1}{2i} \frac{2ia}{s^2 + a^2} \quad (3.64)$$

$$= \frac{a}{s^2 + a^2}. \quad (3.65)$$

Vi kan använda Laplacetransformen av en derivata för att transformera  $\cos(t)$ . Vi har

$$\mathcal{L}(\cos(at))(s) = \mathcal{L}\left(\frac{1}{a} \frac{d}{dt} \sin(at)\right) = \frac{s}{a} F(s) - \sin(0) = \frac{s}{s^2 + a^2}.$$

□

Vi börjar med två exempel på ordinära differentialekvationer som kan lösas med Laplacetransform.

**Exempel 3.11** Vi studerar följande ordinära differentialekvation. Låt  $y(t)$  vara lösning till

$$y''(t) + 2y'(t) + y(t) = e^{-t}, \quad y(0) = 0, \quad y'(0) = 1.$$

Vi Laplacetransformerar ekvationen och låter  $Y(s)$  beteckna Laplacetransformen av  $y(t)$ ,

$$s^2 Y(s) - sy(0) - y'(0) - 2sY(s) + Y(s) = \mathcal{L}(e^{-t})(s) = \frac{1}{s+1}.$$

Här använder vi Sats 3.4 om Laplacetransformen av derivator. Genom att sätta in begynnelsevärdena och flytta om termer får vi

$$(s^2 + 2s + 1)Y(s) = \frac{1}{s+1} + 1,$$

vilket ger,

$$Y(s) = \frac{1}{(s+1)^3} + \frac{1}{(s+1)^2}.$$

Vi använder nu Sats 3.11,  $\mathcal{L}(t^n e^{-t})(s) = \frac{n!}{(s+1)^{n+1}}$ , för att dra slutsatsen att  $y(t) = \frac{1}{2}t^2 e^{-t} + t e^{-t}$ .

I exemplet ser vi att den karakteristiska ekvationen  $r^2 + 2r + 1 = 0$  har en dubbelrot  $r = -1$ . Laplacetransformen hanterat detta, utan speciella ansatser, med hjälp av partialbråksuppdelning. Vi ser även hur tabellen av Laplacetransformer används för att transformera tillbaka till tidsdomänen.



Ett stort tillämpningsområde för Laplacetransform är elektriska nät. Relationen mellan ström, spänning, resistans och induktans kan beskrivas med hjälp av en ODE.

**Exempel 3.12 (Elektrisk krets)** En spole med resistans  $R$  och induktans  $L$  är seriekopplad till en generator med spänning  $u(t) = E$ . Strömmen  $i(t)$  kan beräknas som lösningen till ekvationen

$$Ri(t) + L \frac{di(t)}{dt} = u(t) = E, \quad i(0) = 0.$$

Vi låter Laplacetransformen till  $i(t)$  vara  $I(s)$  och får

$$RI(s) + LsI(s) = \frac{E}{s}.$$

Omskrivning ger

$$I(s) = \frac{E}{s(R + Ls)} = \frac{\frac{E}{L}}{s(s + \frac{R}{L})}.$$

Faltning ger att,

$$i(t) = \frac{E}{L} \int_0^t 1 \cdot e^{-\frac{R}{L}\tau} d\tau = \frac{E}{L} \left[ -\frac{L}{R} e^{-\frac{R}{L}\tau} \right]_0^t = \frac{E}{R} (1 - e^{-\frac{R}{L}t}).$$

Vi ger nu ännu ett exempel på en elektrisk krets som leder till något som kallas en integro-differentialekvation. Det är en differentialekvation som även innehåller (minst) en integral.

**Exempel 3.13 (Elektrisk krets)** En spole med induktans  $L$ , en resistor med resistans  $R$  och en kondensator med kapacitans  $C$  är seriekopplade med en spänningskälla med spänning  $E(t) = e^{-2t}$ . Strömmen  $i(t)$  ges då som lösningen till ekvationen

$$Li'(t) + Ri(t) + \frac{1}{C} \int_0^t i(u) du = E(t).$$

Låt  $L = 1$ ,  $R = 4$ ,  $C = \frac{1}{3}$  och beräkna strömmen  $i(t)$  om  $i(0) = 0$ . Vi Laplacetransformerar och får,

$$sI(s) + 4I(s) + \frac{3}{s}I(s) = \frac{1}{s+2}.$$

Vi multiplicerar med  $s$  och löser ut  $I(s)$ ,

$$I(s) = \frac{s}{s+2} \cdot \frac{1}{s^2 + 4s + 3} = \frac{s}{(s+2)(s+1)(s+3)},$$

eftersom  $s^2 + 4s + 3 = 0$  har lösningar  $s_1 = -1$  och  $s_2 = -3$ . Vi genomför nu partialbråksuppdelning,

$$\frac{s}{(s+1)(s+2)(s+3)} = \frac{A}{s+1} + \frac{B}{s+2} + \frac{C}{s+3}.$$

Genom att göra liknämning får vi ekvationen

$$s = A(s+2)(s+3) + B(s+1)(s+3) + C(s+1)(s+2) \quad (3.66)$$

$$= As^2 + 5As + 6A + Bs^2 + 4Bs + 3B + Cs^2 + 3Cs + 2C \quad (3.67)$$

som ska gälla för alla  $s$ . Detta leder till tre ekvationer med tre obekanta, nämligen ekvationerna för  $s^2$ ,  $s$  och 1 termerna,

$$0 = A + B + C$$

$$1 = 5A + 4B + 3C$$

$$0 = 6A + 3B + 2C$$

Eftersom  $A+B+C = 0$  ger den andra ekvationen att  $1 = 2A+B$  och den tredje  $0 = 4A+B$  eller  $B = -4A$ . Sätter vi in det får vi  $1 = 2A - 4A$  vilket innebär att  $A = -\frac{1}{2}$ ,  $B = 2$ , och därmed  $C = -A - B = -\frac{3}{2}$ . Vi får alltså att

$$I(s) = -\frac{1}{2} \frac{1}{s+1} + \frac{2}{s+2} - \frac{3}{2} \frac{1}{s+3}.$$

Genom att använda Sats 3.11 kan vi Laplacetransformera tillbaka och slutligen få

$$i(t) = -\frac{1}{2}e^{-t} + 2e^{-2t} - \frac{3}{2}e^{-3t}.$$

## Övningar

### 3.1 Definition av Laplacetransform

Övning 3.1 För vilka  $s \in \mathbb{C}$  är Laplacetransformen av  $e^t \cos(t)$  väldefinierad?

Övning 3.2 Finn  $a$  och  $C$  så att  $t^2 \leq Ce^{at}$ .

Övning 3.3 Givet att  $\mathcal{L}(t)(s) = \frac{1}{s^2}$  och  $\mathcal{L}(t^2)(s) = \frac{2}{s^3}$  vad blir  $\mathcal{L}(3t + t^2)(s)$ ?

Övning 3.4 Givet  $F(s) = \frac{7}{s^2}$  bestäm  $f(t)$ .

Övning 3.5 Ange en funktion som inte uppfyller  $|f(t)| \leq Ce^{at}$  för alla  $t \geq 0$ .

### 3.2 Tekniker för att beräkna Laplacetransform

Övning 3.6 Använd skalning för att bestämma Laplacetransformen till  $f(t) = e^{5t} + (3t)^2$ .

Övning 3.7 Givet att  $\mathcal{L}(\cos(t))(s) = \frac{s}{s^2+1}$  bestäm Laplacetransformen till  $f(t) = e^{4t} \cos(t)$ . För vilka  $s \in \mathbb{C}$  är Laplacetransformen väldefinierad?

Övning 3.8 Bestäm den funktion vars Laplacetransform är  $\frac{5}{s^2+3s-4}$ .

Övning 3.9 Bestäm den funktion vars Laplacetransform är  $\frac{s}{s^2-6s+5}$ .

Övning 3.10 Bestäm den funktion vars Laplacetransform är  $\frac{1}{s(s+1)}$ .

### 3.3 Laplacetransform av derivator och integraler

Övning 3.11 Låt  $f(t)$  vara en given funktion sådan att  $f(0) = 3$  och  $f'(0) = 7$ . Låt Laplacetransformen av  $f(t)$  betecknas med  $F(s)$ . Uttryck Laplacetransformen av  $f''(t)$  i termer av  $F(s)$ .

Övning 3.12 Laplacetransformera följande ODE,

$$f''(t) + 3f'(t) = e^{-t}, \quad f(0) = 1, \quad f'(0) = 0.$$

Övning 3.13 Laplacetransformera  $t \cos(t)$ .

Övning 3.14 Bestäm Laplacetransformen till  $\int_0^t \cos(\tau) d\tau$ .

Övning 3.15 Laplacetransformera följande ODE,

$$f'''(t) - f'(t) + f(t) = t, \quad f(0) = 1, \quad f'(0) = 1, \quad f''(0) = 0.$$

### 3.4 Faltning, impuls och diskontinuerliga funktioner

Övning 3.16 Beräkna den funktion vars Laplacetransform ges av  $H(s) = \frac{1}{s(s+1)}$  genom att använda faltning.

Övning 3.17 Bestäm Laplacetransformen till  $3\delta(t-2) + \delta(t-1)$ .

Övning 3.18 Bestäm gränsvärdet  $\lim_{\epsilon \rightarrow 0} \int_{-1}^1 \delta_\epsilon \cos(x) dx$ .

Övning 3.19 Bestäm Laplacetransformen till  $\theta(t-T)e^{-t}$ , där  $\theta$  är Heavisides stegfunktion.

Övning 3.20 Bestäm Laplacetransformen av  $f(t) = \theta(t) - \theta(t-2)$ .

### 3.5 Lösning av ODE med Laplacetransform

Övning 3.21 Använd Laplacetransform för att lösa ekvationen,  $y''(t) = \delta(t)$ ,  $y(0) = 0$  och  $y'(0) = 0$ .

Övning 3.22 Använd Laplacetransform för att lösa ekvationen,  $y' + ky = t$ ,  $y(0) = 0$ .

Övning 3.23 Använd Laplacetransform för att lösa ekvationen  $2y'(t) + 4y(t) = 1$  med begynnelsevillkor  $y(0) = 3$ .

Övning 3.24 Lös integralekvationen  $\int_0^t y(\tau) d\tau + 3y(t) = 2$  med Laplacetransform.

Övning 3.25 Använd Laplacetransform för att lösa ekvationen  $y''(t) + y(t) = 0$  med begynnelsevillkor  $y(0) = 1$  och  $y'(0) = -1$ .

## Problem

---

### 3.1 Definition av Laplacetransform

**Problem 3.1** Antag att  $|f(t)| + |f'(t)| \leq Ce^{at}$  för några  $a, C \geq 0$ . Visa att  $\lim_{s \rightarrow \infty} F(s) = 0$  genom att använda partiell integration.

### 3.2 Tekniker för att beräkna Laplacetransform

**Problem 3.2** Antag att  $\lim_{t \rightarrow 0} \frac{f(t)}{t}$  existerar. Visa att  $\mathcal{L}\left(\frac{f(t)}{t}\right)(s) = \int_s^\infty F(z) dz$ .

### 3.3 Laplacetransform av derivator och integraler

**Problem 3.3** Beräkna Laplacetransformen av  $f(t) = t^n$ . Gäller  $|f(t)| \leq Ce^{at}$  för några  $a, C \geq 0$ ?

### 3.4 Faltning, impuls och diskontinuerliga funktioner

**Problem 3.4** Bestäm Laplacetransformen av  $\sin(t + 1)$ .

### 3.5 Lösning av ODE med Laplacetransform

**Problem 3.5** Använd Laplacetransform för att lösa ekvationen,  $y'' + 2y' + y = f(t)$ ,  $y(0) = 0$ ,  $y'(0) = 1$ , där  $f(t)$  uppfyller  $|f(t)| \leq Ce^{at}$ .

## Kapitel 1

## Övningar

Ö1.1 (a)  $[2 \ 4 \ -3]$  (b)  $[6 \ 0 \ -5]$  (c)  $[0 \ 9 \ -3]$  (d)  $[2 \ -5 \ 0]$

Ö1.2 (a)  $\sqrt{29}/2$  (b)  $\sqrt{5}$  (c) 4 (d)  $\sqrt{n}$

Ö1.3 (a) -5 (b) 2 (c) -1 (d) 30

Ö1.4 (a)  $\theta = \frac{\pi}{6}$  (b)  $\theta = \arccos(\frac{-\sqrt{2}}{3})$  (c)  $\theta = \frac{\pi}{3}$  (d)  $\theta = \arccos(\frac{1}{\sqrt{n}})$

Ö1.5 (a)  $x = 1$  och  $y = 2$  (b)  $x = 1$  och  $y = 1$  (c)  $x = 1, y = 2$  och  $z = 3$  (d)  $x = 7, y = 1$  och  $z = 1$

Ö1.6 (a)

$$A = \begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 8 \\ -1 \end{bmatrix}$$

(b)

$$A = \begin{bmatrix} 1 & 1 \\ 2 & -2 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

(c)

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & -1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 6 \\ 7 \\ 1 \end{bmatrix}$$

(d)

$$A = \begin{bmatrix} 1 & -1 & -3 \\ 0 & 1 & 1 \\ 0 & -3 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$$

Ö1.7 (a)  $[7 \ -1 \ 2]^T$  (b)  $[-1 \ 7 \ 6]^T$  (c)  $[6.5 \ -0.5 \ 2]^T$  (d)  $[4 \ 1 \ 3]^T$

Ö1.8 Första ekvationen ger  $\dot{u} - v = 0$  alltså  $v = \dot{u}$ . Sätt in detta i andra ekvationen,  $0 = \dot{v} + 4u = \ddot{u} + 4u$ . Lösningen till ekvationen  $\ddot{u} + 4u = 0$  ges av  $u(t) = A \sin(2t) + B \cos(2t)$  (den karaktäristiska ekvationen är  $r^2 + 4 = 0$  med lösning  $r_1 = 2i$  och  $r_2 = -2i$  vilket ger lösningar  $u(t) = C_1 e^{2it} + C_2 e^{-2it}$ , realdelen av detta ger lösningen), begynnelsevillkoren ger  $u(t) = \frac{1}{2} \sin(2t)$  och  $v(t) = \cos(t)$ .

Ö1.9 Första ekvationen ger  $\dot{u} + v = 0$ , alltså  $v = -\dot{u}$ . Andra ekvationen ger då med  $v = -\dot{u}$  att  $\dot{v} + v^2 = -\ddot{u} + (\dot{u})^2 = 0$ ,  $u(0) = 0$ , och  $\dot{u}(0) = 1$ . Lösningen till  $\dot{v} + v^2 = 0$  med  $v(0) = 1$  ges av  $v(t) = \frac{1}{1+t}$  (separabel  $v^{-1} = -\int v^{-2} dv = \int dt = t + C$  medför  $v(t) = \frac{1}{t+C}$ ,  $v(0) = 1$  medför  $C = 1$ )  $u(t) = -\int_0^t \frac{1}{1+\tau} d\tau = -\log(1+t) + C$ ,  $u(0) = 0$  ger  $u(t) = -\log(1+t)$  och  $v(t) = \frac{-1}{1+t}$ .

Ö1.10 Vi får  $y''(t) = ty(t)^2$ .

**Ö1.11** Vi deriverar den första  $\ddot{u}_1 + \dot{u}_2 = 2\cos(2t)$  och sätter in  $\dot{u}_2 = u_1 + \cos(2t)$  vilket ger  $\ddot{u}_1 + u_1 = \cos(2t)$ . Vi har  $u_1(0) = 1$  och  $\dot{u}_1(0) = -u_2(0) + \sin(2 \cdot 0) = -1$ . Den karakteristiska ekvationen  $r^2 + 1 = 0$  har rötter  $r_{1/2} = \pm i$  så  $u_h(t) = C_1 \cos(t) + C_2 \sin(t)$ . Vi ansätter  $u_p(t) = A \cos(2t) + B \sin(2t)$  vilket ger  $-3A \cos(2t) - 3B \sin(2t) = \cos(2t)$  som uppfylls då  $A = -1/3$  och  $B = 0$ . Vi har  $u_1(t) = C_1 \cos(t) + C_2 \sin(t) - 1/3 \cos(2t)$ . Begynnelsevillkoren ger  $C_1 = 4/3$  och  $C_2 = -1$ . Vi får då att  $u_2(t) = \sin(2t) - \dot{u}_1(t) = \sin(2t) + \frac{4}{3} \sin(t) + \cos(t) - \frac{2}{3} \sin(2t)$ .

**Ö1.12** Matrisekvationen ger  $u_1' = -u_3$ ,  $u_2' = -u_1 + u_3$  och  $u_3' = 2u_1 + u_2 + 0.5u_3$ . Vi deriverar den första relationen två gånger  $u_1^{(3)} = -u_3''$  och den tredje en gång  $u_3' = 2u_1' + u_2' + 0.5u_3'$  för att få  $u_1^{(3)} = -2u_1' - u_2' - 0.5u_3'$ . Här får vi  $u_2'$  från den andra ekvationen och  $u_3' = -u_1''$  från den första ekvationen. Alltså  $u_1^{(3)} = -2u_1' + u_1 - u_3 + 0.5u_1'' = -2u_1' + u_1 + u_1' + 0.5u_1''$ . Begynnelsevillkoren ger  $u_1(0) = 2$ . Första ekvationen i systemet ger  $u_1'(0) = -u_3(0) = 1$ . Genom att derivera den första ekvationen och använda den tredje får vi  $u_1''(0) = -u_3'(0) = -2u_1(0) - u_2(0) - 0.5u_3(0) = -9/2$ . Genom att använda alla ekvationer får vi slutligen också

$$\begin{aligned} u_1^{(3)}(0) &= -u_3''(0) = -2u_1'(0) - u_2'(0) - 0.5u_3'(0) = \\ &= -2(-u_3(0)) - (-u_1(0) + u_3(0)) - \\ &\quad -0.5(2u_1(0) + u_2(0) + 0.5u_3(0)) = -5/4. \end{aligned}$$

**Ö1.13** Låt  $u = y$  och  $v = \dot{y}$ . Vi får då,

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ x \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Karaktäristiska ekvationen  $r^2 - 2r - 3 = 0$  har lösningar  $r = -1$  och  $r = 3$ . Homogönlösningen ges av  $y_h(x) = C_1 e^{-x} + C_2 e^{3x}$  och partikulärlösningen  $y_p(x) = -\frac{1}{3}x + \frac{2}{9}$ . Begynnelsevillkoren ger  $y(0) = C_1 + C_2 + \frac{2}{9} = 1$  och  $y'(0) = -C_1 + 3C_2 - \frac{1}{3} = 0$  alltså  $y(x) = \frac{1}{2}e^{-x} + \frac{5}{18}e^{3x} - \frac{1}{3}x + \frac{2}{9}$ .

**Ö1.14** Låt  $u = y$  och  $v = y'$ ,

$$\begin{bmatrix} u' \\ v' \end{bmatrix} + \begin{bmatrix} -v \\ \sin(u) \end{bmatrix} = \begin{bmatrix} 0 \\ x \cos(x) \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

**Ö1.15** Låt  $y_1 = y$ ,  $y_2 = y'$  och  $y_3 = y''$ . Vi får,

$$\begin{bmatrix} y_1' \\ y_2' \\ y_3' \end{bmatrix} = \begin{bmatrix} y_2 \\ y_3 \\ x - x \sin(y_2) \end{bmatrix}, \quad \begin{bmatrix} y_1(0) \\ y_2(0) \\ y_3(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

**Ö1.16** Låt  $u_1 = y$  och  $u_2 = y'$ . Vi har  $u_1' = y' = u_2$  och  $u_2' = y'' = -3u_1 + 8u_2 + xe^x$ . Begynnelsevillkoren blir  $u_1(0) = y(0) = 2$  och  $u_2(0) = y'(0) = 3$ . Systemet kan även skrivas på matrisform

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & 8 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 \\ xe^x \end{bmatrix}, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Ö1.17 Sätt  $u_1 := y$ ,  $u_2 := y'$ ,  $u_3 := y''$  och  $u_4 := y^{(3)}$ . Vi uttrycker de nya variablernas förstaderivator med hjälp av  $u_1, u_2, u_3, u_4$ ,

$$\begin{aligned}u_1' &= y' = u_2, \\u_2' &= y'' = u_3, \\u_3' &= y^{(3)} = u_4, \\u_4' &= y^{(4)} = -xu_1 - 3u_2 + 2u_3 + e^{2x} + \cos(x).\end{aligned}$$

Begynnelsevillkoren ger  $u_1(0) = y(0) = 2$ ,  $u_2(0) = y'(0) = 3$ ,  $u_3(0) = y''(0) = 1$  och  $u_4(0) = y^{(3)}(0) = 2$ . Systemet även skrivs på matrisform

$$\begin{bmatrix} u_1' \\ u_2' \\ u_3' \\ u_4' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -x & -3 & 2 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ e^{2x} + \cos(x) \end{bmatrix}, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \\ u_3(0) \\ u_4(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 2 \end{bmatrix}.$$

Ö1.18 Låt  $u_1 = y$  och  $u_2 = y'$ . Vi får

$$\begin{aligned}u_1' &= y' = u_2 \\u_2' &= y'' = -u_1^2 u_2 + e^x.\end{aligned}$$

Begynnelsevillkoren ger  $u_1(0) = y(0) = -1$  och  $u_2(0) = y'(0) = 2$ . Systemet kan skrivas med vektornotation

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} u_2 \\ -u_1^2 u_2 \end{bmatrix} + \begin{bmatrix} 0 \\ e^x \end{bmatrix}, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Ö1.19 Vi har

$$|\vec{f}(t, \vec{y}_1) - \vec{f}(t, \vec{y}_2)| \leq \sqrt{(v_2 - v_1)^2 + (4u_1 - 4u_2)^2} \leq 4\sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2} = 4|\vec{y}_1 - \vec{y}_2|.$$

Sats 1.4 ger entydig lösning för alla  $t$ .

Ö1.20 Vi har

$$|\vec{f}(t, \vec{y}_1) - \vec{f}(t, \vec{y}_2)| = \sqrt{(v_1 - v_2)^2 + (v_1^2 - v_2^2)^2} = |v_1 - v_2| \sqrt{1 + (v_1 + v_2)^2}.$$

Vi drar slutsatsen att det ej finns någon Lipschitzkonstant eftersom för varje  $L > 0$  sådant att

$$|\vec{f}(t, \vec{y}_1) - \vec{f}(t, \vec{y}_2)| \leq L|\vec{y}_1 - \vec{y}_2| = L\sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2},$$

kan vi välja  $u_1 = u_2$  och  $v_1, v_2 \in \mathbb{R}$  sådana att  $\sqrt{1 + (v_1 + v_2)^2} > L$ .

Ö1.21 Lipschitzkontinuitet i  $y = 0$  eftersom derivatan där är oändlig.

Ö1.22 Vi har att  $u' = (y^2)' = 2yy' = 2(y^2) = 2u$  dessutom  $u(0) = (y^2)(0) = y(0)^2 = 1$ .

Ö1.23 Vi har att  $xy = e^{\ln(xy)}$  eftersom  $\ln(x)$  är invers till  $e^x$ . Vi har även  $xy = e^{\ln(x)}e^{\ln(y)} = e^{\ln(x)+\ln(y)}$ . Genom att ta logaritmen på båda sidor får vi  $\ln(xy) = \ln(x) + \ln(y)$ .

Ö1.24 Vi har att  $e^{\ln(x^p)} = x^p = (e^{\ln(x)})^p = e^{p\ln(x)}$ . Genom att ta logaritmen på båda sidor får vi  $\ln(x^p) = p\ln(x)$ .

Ö1.25 Följer direkt genom insättning.



## Problem

**P1.1** Vi har  $|\vec{v} - \vec{w}|^2 = (\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w}) = |\vec{v}|^2 + |\vec{w}|^2 - 2\vec{v} \cdot \vec{w}$ . Cosinussatsen ger då att  $-2\vec{v} \cdot \vec{w} = -2|\vec{v}||\vec{w}|\cos(\theta)$  vilket ger resultatet efter division med  $-2$ .

**P1.2** Ja, högerledet är Lipschitzkontinuerligt eftersom det består av (upp till kvadratis- ka) polynom som har begränsade derivator på området  $0 \leq x, y \leq 10$ . De stationära punkterna ges av  $x = y = 1$  och  $x = y = 0$ .

**P1.3** Vi har att

$$m_1 \ddot{x}_1(t) = -k_1 x_1(t) + k_2(x_2(t) - x_1(t)), \quad (3.68)$$

$$m_2 \ddot{x}_2(t) = -k_2(x_2(t) - x_1(t)) - k_3 x_2(t). \quad (3.69)$$

Vi inför variablerna  $x_3(t) = \dot{x}_1(t)$  och  $x_4(t) = \dot{x}_2(t)$ . Vi får följande system,

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2k_1/m_1 & k_2/m_1 & 0 & 0 \\ k_2/m_2 & -2k_3/m_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}. \quad (3.70)$$

**P1.4** Variabelseparation ger  $\int \frac{dy}{1+y^2} = \int 1 dx = x + C$ . Därför får vi  $\arctan(y(x)) = x + C$ . Begynnelsevillkoret  $y(0) = 0$  ger  $0 = \arctan(0) = 0 + C$ . Alltså får vi  $y(x) = \tan(x)$ . Högerledet är inte Lipschitzkontinuerligt för alla  $y$  eftersom  $|1 - y_1^2 - 1 + y_2^2| = |y_1 + y_2||y_1 - y_2|$  och  $|y_1 + y_2|$  inte är begränsad av något  $L$  för alla  $y_1, y_2 \in \mathbb{R}$ . Vi därför inte global existens men lokal existens i en omgivning till  $0$ ,  $-\delta < x < \delta$  för något  $\delta$ .

**P1.5** Vi har  $\frac{y_n - y_{n-1}}{1/n} = y_n$  eller

$$y_n = \frac{y_{n-1}}{1 - \frac{1}{n}} = \dots = \frac{y_0}{(1 - 1/n)^n} = (1 - 1/n)^{-n},$$

eftersom  $y_0 = 1$ . Därför gäller

$$e = \lim_{n \rightarrow \infty} (1 - 1/n)^{-n}.$$

## Kapitel 2

---

### Övningar

**Ö2.1** Vi får  $y_n = (1 - \frac{2}{n})^{-1} y_{n-1} = \dots = (1 - \frac{2}{n})^{-n}$ .

**Ö2.2**  $y_n = (1 + \frac{1}{n}) y_{n-1} = (1 + \frac{1}{n})^n y_0 = (1 + \frac{1}{n})^n$ .

**Ö2.3** Vi får  $y_1 = y_0 + 0.1 \cdot y_0^2 = 1.1$  och  $y_2 = y_1 + 0.1 \cdot y_1^2 = 1.1 + 0.1 \cdot 1.1^2 = 1.221$ .

**Ö2.4** Mittpunktsmetoden ger  $y_i = y_{i-1} + \frac{1}{n} \frac{y_i + y_{i-1}}{2}$  vilket ger

$$y_n = \frac{1 + \frac{1}{2n}}{1 - \frac{1}{2n}} y_{n-1} = \left( \frac{1 + \frac{1}{2n}}{1 - \frac{1}{2n}} \right)^n y_0 = \left( \frac{1 + \frac{1}{2n}}{1 - \frac{1}{2n}} \right)^n.$$

Ö2.5 Eulers metod ger  $y_1 = y_0 + k(1 - y_0^2) = 0.2$ . Vi får sedan att  $y_2 = y_1 + k(1 - y_1^2) = 0.2 + 0.2(1 - 0.04) = 0.3920$ .

Ö2.6 Vi har  $y_1 = 1 + 0.2 \cdot 0 \cdot 1 = 1$  och  $y_2 = 1 + 0.2 \cdot 0.2 \cdot 1 = 1.04$ .

Ö2.7 Explicit. Ingen algebraisk ekvation behöver lösas.

Ö2.8 Låt  $y_1 = y$ ,  $y_2 = y'$  och  $y_3 = y''$ .

$$\begin{bmatrix} y_1' \\ y_2' \\ y_3' \end{bmatrix} = \begin{bmatrix} y_2 \\ y_3 \\ \sin(t) - y_2 \end{bmatrix}, \quad \begin{bmatrix} y_1(0) \\ y_2(0) \\ y_3(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Systemet kan lösas med text Bakåt Euler som är en första ordningens metod, alltså felet beror linjärt på steglängden.

Ö2.9 Integrerande faktor ger att  $y(x) = 1 - x + e^{-x}$ . Därmed är  $|y''(x)| = |e^{-x}| \leq 1 = M$ . Lipschitzkonstanten med avseende på  $y$  för  $f(y) = x - y$  är  $L = 1$ . Vi får enligt Sats 2.1 att  $|y(1) - y_{100}| \leq \frac{1}{200}(e - 1) = \frac{e-1}{200}$ .

Ö2.10 Vi har  $|y''(x)| = |\sin(y(x))y'(x)| = |\sin(y(x))\cos(y(x))| = 0.5|\sin(2y(x))| \leq 0.5$ . Vidare är  $L = 1$  eftersom derivatan av  $\cos(y)$  är begränsad av 1. Sats 2.1 ger då  $|y(2) - y_{100}| \leq \frac{0.5 \cdot 0.02}{2 \cdot 1}(e^2 - 1) = 0.005(e^2 - 1)$ .

Ö2.11 Vi har  $|y''| = |e^{-t}| \leq 1$  och även  $L = 1$  eftersom derivatan också är begränsad av 1. Sats 2.1 ger  $|y(T) - y_{10T}| \leq 0.05(e^T - 1)$ . För  $T = 3.045$  får vi  $0.05(e^T - 1) > 1$ .

Ö2.12 Felet i Eulers metod är proportionellt mot steglängden  $Ck = C\pi/10 = 0.012$ . Det ger att  $C = 0.12/\pi$ . Ska felet vara 0.00064 krävs att  $Ck = 0.00064$  eller  $k = 0.0168$ .

Ö2.13 Vi har  $\lambda = -10$ . Eulers metod är stabil om  $|1 + k\lambda| < 1$ . Det innebär att  $|1 - 10k| < 1$  eller  $k < 0.1$ . För bakåt Euler gäller  $|1 - k\lambda| > 1$  eller  $|1 + 10k| > 1$  vilket gäller för alla  $k$ .

Ö2.14 Eftersom  $k\lambda < 0$  gäller att  $|1 - \frac{k\lambda}{2}| = 1 + |\frac{k\lambda}{2}| \geq |1 + \frac{k\lambda}{2}|$  vilket innebär stabilitet enligt Sats 2.2.

Ö2.15 Vi har att  $|1 - k\lambda| = 1 + |k\lambda| > 1$  vilket innebär stabilitet enligt Sats 2.2.

Ö2.16 Eulers metod:

$$\begin{bmatrix} u_i - u_{i-1} \\ v_i - v_{i-1} \end{bmatrix} + 0.1 \begin{bmatrix} 0 & -1 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} u_{i-1} \\ v_{i-1} \end{bmatrix} = 0.1 \begin{bmatrix} 0 \\ x_{i-1} \end{bmatrix}, \quad \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

där  $x_i = i \cdot k$ ,  $i = 0, 1, \dots$   $[u_1, v_1] = [1, 0.3]$ .

Ö2.17 Euler metod

$$\begin{bmatrix} u_i - u_{i-1} \\ v_i - v_{i-1} \end{bmatrix} + 0.1 \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} u_{i-1} \\ v_{i-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

där  $x_i = i \cdot h$ ,  $i = 0, 1, \dots$   $[u_1, v_1] = [0.1, 1]$ .

Ö2.18 Bakåt Euler:

$$\begin{bmatrix} u_i - u_{i-1} \\ v_i - v_{i-1} \end{bmatrix} + 0.1 \begin{bmatrix} 0 & -1 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} u_i \\ v_i \end{bmatrix} = k \begin{bmatrix} 0 \\ x_i \end{bmatrix}, \quad \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

där ,  $x_i = i \cdot k, i = 0, 1, \dots$

Ö2.19 Bakåt Euler:

$$\begin{bmatrix} u_i - u_{i-1} \\ v_i - v_{i-1} \end{bmatrix} + k \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

där ,  $x_i = i \cdot k, i = 0, 1, \dots$

Ö2.20 Låt  $y_1 = y, y_2 = y'$ .

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} y_2 \\ \sin(t) - y_1^2 \end{bmatrix}, \quad \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Med  $t_i = ik, i = 0, 1, \dots$ , ges Eulers metods lösning av:

$$\begin{bmatrix} y_1^i - y_1^{i-1} \\ y_2^i - y_2^{i-1} \end{bmatrix} = k \begin{bmatrix} y_2^{i-1} \\ \sin t_{i-1} - (y_1^{i-1})^2 \end{bmatrix}, \quad \begin{bmatrix} (y_1)_0 \\ (y_2)_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Vi får  $y_1^1 = 1$  och  $y_2^1 = -0.1$ .

Ö2.21 Med  $y_1 = y$  och  $y_2 = y'$  får vi,

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Bakåt Euler ger,

$$\begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0.1 \cdot \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} \cdot \begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \end{bmatrix}$$

Vi får att  $y_2^{(1)} = 0.4y_1^{(1)}$ . Insättning i första ekvationen ger  $y_1^{(1)} = 1 + 0.04y_1^{(1)}$  vilket ger  $y_1^{(1)} = \frac{25}{24}$  och  $y_2^{(1)} = \frac{5}{12}$ .

Ö2.22  $q(x) = -\frac{d^2}{dx^2}x^2(1-x) = 6x - 2$ .

Ö2.23  $q(x) = -\frac{d^2}{dx^2}\sin(\pi x) = \pi^2 \sin(\pi x)$ .

Ö2.24 Vi får  $\frac{d}{dx}u(x) = -\frac{x^3}{3} + C$  och  $u(x) = Cx + D - \frac{x^4}{12}$ . Randvillkoren ger  $0 = u(0) = D$  och  $u(1) = C - \frac{1}{12}$  vilket innebär  $C = \frac{1}{12}$ . Vi får  $u(x) = \frac{x}{12}(1-x^3)$ .

Ö2.25 Vi får  $\frac{d}{dx}u(x) = \frac{1}{\pi} \cos(\pi x) + C$  och  $u(x) = Cx + D + \frac{\sin(\pi x)}{\pi^2}$ .  $0 = u(0) = D$  och  $0 = u(1) = C$  ges  $u(x) = \frac{\sin(\pi x)}{\pi^2}$ .

### Problem

**P2.1** Vi får att  $\tilde{y}_1 = 1 + 0.11^2 = 1.1$  och därmed  $y_1 = 1 + 0.05(1 + 1.1) = 1.105$ . Detta ger  $\tilde{y}_2 = y_1 + hy_1^2 = 1.2271$  och slutligen  $y_2 = 1.2216$ .

**P2.2** Vi studerar först det lokala trunkeringsfelet på intervallet  $[t_{i-1}, t_i]$ . Taylors formel ger att  $y(t_{i-1})$  kan uttryckas med en Taylor serie kring punkten  $y(t_i)$  som,

$$y(t_{i-1}) = y(t_i) - k \cdot y'(t_i) + \frac{k^2}{2} y''(\eta_i),$$

för något  $\eta_i \in [t_{i-1}, t_i]$ . Av ekvationen  $y'(t) = f(t, y(t))$  följer att,

$$y(t_{i-1}) = y(t_i) - kf(t_i, y(t_i)) + \frac{k^2}{2} y''(\eta_i).$$

Bakåt Euler ger att  $y_i = y_{i-1} + k \cdot f(t_i, y_i)$  eller  $y_{i-1} = y_i - k \cdot f(t_i, y_i)$ . Om vi låter  $e_i = y(t_i) - y_i$  har vi,

$$e_{i-1} = e_i - k(f(t_i, y(t_i)) - f(t_i, y_i)) + \frac{k^2}{2} y''(\eta_i).$$

och eftersom  $f$  är Lipschitz kontinuerlig med konstant  $L$  i det andra argumentet och andra derivatan av  $y$  är begränsad av  $M$  får vi,

$$|e_i| \leq |e_{i-1}| + kL|e_i| + \frac{k^2 M}{2}.$$

Med  $kL < 1$  får vi,

$$|e_i| \leq (1 - kL)^{-1} |e_{i-1}| + \frac{k^2 M}{2(1 - kL)}.$$

Om vi upprepar denna räkning från  $i = n$  ner till  $i = 1$  får vi det globala trunkeringsfelet,

$$|e_n| \leq (1 - kL)^{-1} |e_{n-1}| + \frac{k^2 M}{2(1 - kL)} \quad (3.71)$$

$$\leq (1 - kL)^{-n} |e_0| + \frac{k^2 M}{2(1 - kL)} (1 + (1 - kL)^{-1} + (1 - kL)^{-2} + \dots + (1 - kL)^{-(n-1)}) \quad (3.72)$$

$$= \frac{k^2 M}{2(1 - kL)} \sum_{i=0}^{n-1} (1 - kL)^{-i}, \quad (3.73)$$

eftersom  $|e_0| = |y(t_0) - y_0| = 0$ . Summationsformeln för geometrisk summa ger

$$|y(t_n) - y_n| := |e_n| \leq \frac{kM}{2} \frac{(1 - kL)^{-n} - 1}{L}.$$

**P2.3** Vi låter  $f(t, y) = \lambda y$ . Vi få att  $\tilde{y}_i = (1 + k\lambda)y_{i-1}$  och

$$y_i = y_{i-1} + \frac{k}{2} (\lambda y_{i-1} + \lambda \tilde{y}_i) = (1 + \lambda k + \frac{\lambda^2 k^2}{2}) y_{i-1}.$$

Därför ges stabilitetsområdet av  $\{k\lambda \in \mathbb{C} : |1 + \lambda k + \frac{\lambda^2 k^2}{2}| < 1\}$ .

P2.4 Låt  $u = y$  och  $v = y'$ . Vi får,

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} v \\ 1 + x - v \end{bmatrix}, \quad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Med  $x_i = ik$ ,  $i = 0, 1, \dots$ , ges Bakåt Euler lösningen av:

$$\begin{bmatrix} u^i - u^{i-1} \\ v^i - v^{i-1} \end{bmatrix} = h \begin{bmatrix} v^i \\ 1 + x_i - v^i \end{bmatrix}, \quad \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Vi får  $u^1 = 1 + 0.1v^1$  och  $v^1 = 0.1(1.1 - v^1)$  vilket ger  $v^1 = 0.1$  och därmed  $u^1 = 1.01$ .

P2.5 Vi får genom partiell integration att  $0 \leq \int_0^L u'(x)u'(x) dx = \int_0^L -u(x)''u(x) dx + 0 = q \int_0^L u(x) dx$ . Det går även att räkna ut lösningen analytiskt och använda att den är positiv.

## Kapitel 3

### Övningar

Ö3.1 Vi har att  $e^t \cos(t) \leq e^t$  för  $t > 0$ . Alltså gäller att  $|e^t \cos(t)| \leq Ce^{at}$  håller med  $C = a = 1$  och Laplacetransformen är väldefinierad för  $Re(s) > 1$ .

Ö3.2 Vi har att  $1 + t \leq e^t$  för  $t > 0$  eftersom funktionerna är lika för  $t = 0$  och högerledet har större derivata för att  $t > 0$ . Därför gäller att  $t \leq e^t$  och även  $t^2 \leq e^{2t}$ . Begränsningen  $t^2 \leq Ce^{at}$  gäller alltså med  $C = 1$  och  $a = 2$  alltså  $Re(s) > 2$ . Det går att hitta mindre värden på  $a$  med en skarpare kalkyl.

Ö3.3 På grund av linjäritet får vi  $\mathcal{L}(3t + t^2)(s) = 3\mathcal{L}(t)(s) + \mathcal{L}(t^2)(s) = \frac{3}{s^2} + \frac{2}{s^3}$ .

Ö3.4 Vi har att  $\mathcal{L}(t)(s) = \frac{1}{s^2}$ . Linjäritet ger att  $f(t) = 7t$ .

Ö3.5 En funktion på formen  $e^{g(t)}$  där  $g(t)$  växer snabbare än linjärt, till exempel  $e^{t \ln(t)} = te^t$  eller  $e^{t^2}$ . För varje fix  $C$  och  $a$  finns då  $t$  så att till exempel  $e^{t^2} > Ce^{at} = e^{\ln(C)at}$ . I detta fallet räcker  $t > a \ln(C)$ .

Ö3.6 Eftersom  $\mathcal{L}(e^t)(s) = \frac{1}{s-1}$  och  $\mathcal{L}(t^2)(s) = \frac{2}{s^3}$  får vi via skalning att  $\mathcal{L}(e^{5t} + (3t)^2)(s) = \frac{1}{s-5} + \frac{2 \cdot 3^2}{s^3} = \frac{1}{s-5} + \frac{18}{s^3}$ .

Ö3.7 Exponentiell skalning ger  $\mathcal{L}(e^{4t} \cos(t)) = \frac{s-4}{(s-4)^2+1} = \frac{s-4}{s^2-8s+17}$ . Eftersom  $|\cos(t)| \leq 1$  gäller att Laplacetransformen är väldefinierad för  $s > 4$ .

Ö3.8 Vi faktorerar nämnaren och utför partialbråksuppdelning  $\frac{5}{(s-1)(s+4)} = \frac{A}{s-1} + \frac{B}{s+4}$  vilket ger  $A+B = 0$  och  $5 = 4A - B$  med lösning  $A = 1$  och  $B = -1$ . Vi får  $\frac{5}{(s-1)(s+4)} = \frac{1}{s-1} - \frac{1}{s+4}$  med invers Laplacetransform  $f(t) = e^t - e^{-4t}$ .

Ö3.9 Vi faktorerar nämnaren och utför partialbråksuppdelning  $\frac{s}{(s-1)(s-5)} = \frac{A}{s-1} + \frac{B}{s-5}$  vilket ger  $1 = A + B$  och  $0 = -5A - B$  med lösning  $A = -\frac{1}{4}$  och  $B = \frac{5}{4}$ . Den inversa Laplacetransformen ges av  $f(t) = -\frac{1}{4}e^t + \frac{5}{4}e^{5t}$ .

Ö3.10 Partialbråksuppdelning ger  $\frac{1}{s(s+1)} = \frac{1}{s} - \frac{1}{s+1}$ . Funktionen ges då av  $1 - e^{-t}$ .

Ö3.11 Laplacetransformen ges av  $s^2F(s) - 3s - 7$ .

Ö3.12 Vi låter  $f(t)$ 's Laplacetransform vara  $F(s)$ . Vi får genom att sätta in begynnelsevillkoren,

$$s^2F(s) - s + 3sF(s) - 3 = \frac{1}{s+1}.$$

Ö3.13 Låt  $g(t) = tf(t)$  där  $f(t) = \cos(t)$ . Då gäller  $F(s) = \frac{s}{s^2+1}$  enligt Sats 3.7  $G(s) = -F'(s) = \frac{s^2-1}{(s^2+1)^2}$ .

Ö3.14 Sats 3.6 ger  $\frac{1}{s}\mathcal{L}(\cos(t)) = \frac{1}{s^2+1}$ .

Ö3.15 Vi får

$$s^3F(s) - s^2 - s - sF(s) + 1 + F(s) = \frac{1}{s^2}.$$

Ö3.16 Vi vet att funktionen  $f(t) = 1$  har transform  $F(s) = \frac{1}{s}$  och funktionen  $g(t) = e^{-t}$  har transform  $G(s) = \frac{1}{s+1}$ . Eftersom  $H(s) = F(s) \cdot G(s)$  ges  $h(t) = \int_0^t f(t-\tau)g(\tau)d\tau = \int_0^t 1 \cdot e^{-\tau}d\tau = 1 - e^{-t}$ .

Ö3.17 Vi får  $3e^{-2s} + e^{-s}$ .

Ö3.18  $\int_{-1}^1 \delta_\epsilon \cos(x)dx = \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} \cos(x)dx = \frac{1}{2\epsilon} \cos(c_\epsilon)$  för något  $c_\epsilon \in [-\epsilon, \epsilon]$ . När vi låter  $\epsilon \rightarrow 0$  går  $c_\epsilon$  mot 0 och  $\lim_{\epsilon \rightarrow 0} \int_{-1}^1 \delta_\epsilon \cos(x)dx = 1$ .

Ö3.19 Vi har att  $f(t) = \theta(t-T)e^{-t}$ . Vi får  $F(s) = \frac{e^{-Ts}}{s+1}$ .

Ö3.20  $F(s) = \int_0^2 e^{-st} dt = \left[-\frac{e^{-st}}{s}\right]_0^2 = \frac{1-e^{-2s}}{s}$ .

Ö3.21 Vi får  $s^2Y(s) = 1$  vilket ger  $Y(s) = \frac{1}{s^2}$  eller  $y(t) = t, t \geq 0$ .

Ö3.22 Vi får  $Y(s) = \frac{1}{s^2(s+k)} = \frac{-1/k^2}{s} + \frac{1/k}{s^2} + \frac{1/k^2}{s+k}$ . Går vi tillbaka till  $y(t)$  får vi då  $y(t) = -1/k^2 + t/k + e^{-kt}/k^2$ .

Ö3.23 Laplacetransform ger  $2sY(s) - 6 + 4Y(s) = \frac{1}{s}$ . Alltså  $Y(s) = \frac{3}{s+2} + \frac{1/2}{s(s+2)}$ . Vi partialbråksuppdelar den andra termen  $\frac{1}{s(s+2)} = \frac{A}{s} + \frac{B}{s+2}$  där  $A = 1/2$  och  $B = -1/2$ . Alltså  $Y(s) = \frac{11/4}{s+2} + \frac{1/4}{s}$ . Därför blir  $y(t) = \frac{1}{4} + \frac{11}{4}e^{-2t}$ .

Ö3.24 Vi får  $\frac{1}{s}Y(s) + 3Y(s) = \frac{2}{s}$  vilket ger  $Y(s) = \frac{2}{3s+1} = \frac{2/3}{s+\frac{1}{3}} = \frac{2}{3}e^{-t/3}$ .

Ö3.25 Vi får  $s^2Y(s) - s + 1 + Y(s) = 0$  vilket ger  $Y(s) = \frac{s-1}{s^2+1}$  och därmed  $y(t) = \cos(t) - \sin(t)$ .

## Problem

**P3.1** Låt  $s \in \mathbb{R}$  sådan att  $s > a$ . Vi har  $F(s) = \int_0^\infty f(t)e^{-st} dt = [f(t)\frac{e^{-st}}{s}]_0^\infty - \frac{1}{s} \int_0^\infty f'(t)e^{-st} dt$ . Eftersom både  $|f(t)| \leq Ce^{at}$  och  $|f'(t)| \leq Ce^{at}$  gäller att,

$$|F(s)| \leq \lim_{t \rightarrow \infty} \frac{Ce^{(a-s)t}}{s} + \frac{|f(0)|}{s} + \frac{1}{s} \int_0^\infty Ce^{(a-s)t} dt \leq \frac{C'}{s},$$

vilket innebär att  $\lim_{s \rightarrow \infty} F(s) = 0$ .

**P3.2** Vi har

$$\begin{aligned} \int_s^\infty F(z) dz &= \int_z^\infty \int_0^t f(t)e^{-zt} dt dz \\ &= \int_0^\infty f(t) \int_s^\infty e^{-zt} dz dt \\ &= \int_0^\infty f(t) \left[ -\frac{e^{-zt}}{t} \right]_s^\infty dt \\ &= \int_0^\infty \frac{f(t)}{t} e^{-st} dt \\ &= \mathcal{L}\left(\frac{f(t)}{t}\right)(s). \end{aligned}$$

**P3.3** Först noterar vi att  $t^n = e^{n \ln(t)} \leq e^{nt}$  alltså gäller begränsningen med  $C = 1$ ,  $a = n$ . Låt  $I_n = \int_0^\infty t^n e^{-st} dt = [t^n \frac{e^{-st}}{-s}]_0^\infty + \frac{n}{s} \int_0^\infty t^{n-1} e^{-st} dt = \frac{n}{s} I_{n-1}$ . Vi har även att  $I_0 = \frac{1}{s}$ . Rekursion ger  $F(s) = \frac{n!}{s^{n+1}}$ .

**P3.4** Efter substitutionen  $\tau = t+1$  har vi  $\int_0^\infty \sin(t+1)e^{st} dt = e^{-s} \int_1^\infty \sin(\tau)e^{s\tau} d\tau$ . Vi använder den komplexa omskrivningen av sinus för att få,

$$\int_0^\infty \sin(t+1)e^{st} dt = \frac{e^{-s}}{2i} \int_1^\infty e^{(s+i)\tau} - e^{(s-i)\tau} d\tau = \frac{e^{-s}}{2i} \left[ \frac{e^{(s+i)\tau}}{s+i} - \frac{e^{(s-i)\tau}}{s-i} \right]_1^\infty = \frac{s \cdot \sin(1) + \cos(1)}{s^2 + 1}.$$

**P3.5** Vi har  $(s^2 + 2s + 1)Y(s) + 1 = F(s)$  detta ger  $Y(s) = F(s)(1+s)^{-2} - (1+s)^{-2}$ . Transformerar vi tillbaka får vi:  $y(t) = \int_0^t f(t-\tau)\tau e^{-\tau} d\tau - te^{-t}$  eftersom  $te^{-t}$  har Laplacetransform  $(1+s)^{-2}$ .

