2020-02-07 (version 2)

TATA41, Envariabelanalys 1, 6 hp
Kursprogram för D/IT/U/KB/TB, vt 2020

Hans Lundmark, MAI, LiU

Allmänt

Denna kurs är obligatorisk på alla LiU:s civilingenjörsprogram i Linköping. Programmen D, IT, U, KB, TB läser den i period vt1, liksom M, DPU, EMM. Övriga läser den i period ht2. Lisam används för närvarande inte i denna kurs, utan all information finns på kurshemsidan: **courses.mai.liu.se/GU/TATA41/**.

Litteratur

- Kursbok: **Matematisk analys: En variabel** av **Göran Forsling** och **Mats Neymark**, andra upplagan, Liber, 2011, ISBN 978-91-47-10023-1. Kursen omfattar kapitel 3–6.
- Problemsamling: **Problem för envar**, MAI, 2018 eller 2019, Tryckakademin (Kårallen, LiU).

Undervisning och hemarbete

[Länk till Föreläsnings- och lektionsplan på s. 3.]

Undervisningen består av 15 föreläsningar (30 timmar) och 16 lektioner (32 timmar). Kursens storlek är 6 hp, vilket motsvarar c:a 160 timmars arbete, och eftersom den schemalagda tiden är 62 timmar så blir den rekommenderade självstudietiden 98 timmar.

Föreläsningarna behandlar valda delar av teorin, samt tar upp en del exempel. **Föreläsningarna har inte som ambition att täcka allt som ingår i kursen; det finns det inte tid till.** Resten får man läsa sig till i kurslitteraturen.

Lektionerna ägnas åt eget arbete med övningsuppgifterna i problemsamlingen, och gärna även gemensamma diskussioner i klassen. **De flesta lektionerna innehåller många fler uppgifter än vad en typisk student hinner räkna på två timmar, så du bör räkna med att göra de flesta uppgifterna utanför schemalagd tid.** Om du går till en lektion oförberedd kommer du troligen bara att hinna med de första uppgifterna på programmet, vilket oftast är enklare uppvärmningsuppgifter. Detta är inte effektiv användning av lektionstiden. **Gör så många uppgifter du kan i förväg**, så att du har chansen att hinna få hjälp med de svårare uppgifterna på lektionen.

Examination

Kursen examineras genom en skriftlig tentamen (TEN1) med 5 timmars skrivtid och 7 uppgifter. Varje uppgift bedöms som godkänd (vilket ger 2 eller 3 poäng) eller underkänd (0 eller 1 poäng). För att få betyg 3/4/5 behövs 3/4/5 stycken godkända uppgifter och dessutom sammanlagt 8/12/16 poäng. Gamla tentor finns på kurshemsidan.

Förkunskaper

Grundkurs i matematik (TATA79, TATA68, TATM79 eller motsvarande).

Det är helt nödvändigt att behärska grundläggande algebra (bråkräkning, potenslagar, polynomdivision, binomialsatsen, m.m.) samt räkneregler och egenskaper för de funktioner som behandlas i grundkursen: exponentialfunktionen, logaritmer, trigonometriska funktioner, arcusfunktioner. Du behöver också kunna hantera olikheter och absolutbelopp, samt förstå vad orden **definition**, **sats** och **bevis** betyder. Komplexa tal förekommer inte så mycket i den här kursen, förutom att Eulers formler

$$e^{ix} = \cos x + i \sin x, \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}, \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}$$

är användbara vid uträkning av vissa trigonometriska integraler.

Vad handlar kursen om?

Matematisk analys är den del av matematiken som handlar om begreppen **gränsvärde**, **derivata** och **integral**, samt sådant som som bygger på detta, exempelvis **differentialekvationer** som tas upp i TATA42 (Envariabelanalys 2) och i sin tur används i otaliga tillämpningar inom naturvetenskap och andra ämnen. Dessa begrepp bör vara bekanta från gymnasiematematiken, men i denna kurs kommer vi att gå betydligt djupare, med större krav både på räknefärdighet och på teoretisk förståelse. Kursen börjar från grunden, så i princip behöver man inte kunna något om ämnet innan. Det underlättar dock mycket om man från gymnasiet har fått med sig lite känsla för idéerna. En bra överblick (som även innehåller vissa saker som kommer i senare kurser) ges i videoserien *The Essence of Calculus* på YouTube-kanalen *3Blue1Brown*: **3b1b.co/calculus**.

Analysen har en lång och invecklad historia, som t.ex. innefattar area- och volymsberäkningar i antikens Grekland, men man brukar säga att den föddes i slutet av 1600-talet när Isaac Newton och Gottfried Wilhelm Leibniz oberoende av varandra utvecklade **infinitesimalkalkylen**, med vars hjälp vissa typer av resonemang som tidigare varit svåra och komplicerade kunde ersättas med enkla uträkningar, t.ex. när det gällde att bestämma arean under en kurva. Denna kalkyl använde sig av **infinitesimaler**, tal som man föreställde sig som ”oändligt små” men ändå inte noll: en positiv infinitesimal ξ uppfyller $0 < \xi < 1/n$ för alla positiva heltal n. Derivatan definierades som förhållandet mellan två infinitesimaler, och integralen som en summa av oändligt många infinitesimaler. Dåtidens matematiker hade vissa problem att förklara exakt vad för slags tal infinitesimalerna var egentligen, men det hindrade inte att de nya metoderna framgångsrikt användes för att lösa många problem, t.ex. av Leonhard Euler som var den mest framstående matematikern under 1700-talet. Euler publicerade 1748 en berömd lärobok i analys, *Introductio in analysin infinitorum* (Introduktion till analysen av det oändliga).

Den version av analysen som ni kommer att få lära er är annorlunda formulerad, och är resultatet av många matematikers arbete med matematikens logiska grundvalar under 1800-talet (Augustin-Louis Cauchy, Bernhard Riemann, Richard Dedekind, Karl Weierstrass, Georg Cantor, m.fl.). Här används bara det **reella talsystemet**, som **inte** innehåller några nollskilda infinitesimaler: det enda tal $x \in \mathbf{R}$ som uppfyller $|x| < 1/n$ för alla positiva heltal n är $x = 0$. Definitionerna av derivata och integral baseras istället på begreppet **gränsvärde**, som vi tar upp i början av kursen, och som också hänger nära samman med ett annat viktigt begrepp, **kontinuitet**.

Fortfarande används dock vissa skrivsätt med ursprung i infinitesimalkalkylen: derivata skrivs ibland

$$\frac{dy}{dx}$$

där dy och dx från början symboliserade ”infinitesimala ändringar av variablerna y resp. x”, och integraler skrivs

$$\int f(x)\,dx$$

där integraltecknet $\int$ är ett stiliserat S som i *summa*, vilket indikerade att man skulle ”summera” de oändligt många infinitesimalerna $f(x)\,dx$ för alla x i något intervall.

Det kan nämnas att det numera faktiskt också finns sätt att rigoröst formulera analysen med hjälp av infinitesimaler; t.ex. utvecklade Abraham Robinson på 1960-talet *ickestandard-analysen*, som bygger på det *hyperreella talsystemet*, ett talsystem som utvidgar det reella talsystemet till att innehålla oändligt små och oändligt stora tal. Där definieras t.ex. derivatan genom att man dividerar två infinitesimaler och sedan ”avrundar” det hyperreella resultatet till närmaste reella tal. Men detta är ett ämne för en annan kurs!

Föreläsnings- och lektionsplan

- I ditt schema kan det eventuellt vara så att någon lektion kommer senare i förhållande till föreläsningarna jämfört med ordningen i programmet.
- Problem märkta med ”B” finns i boken, problem med ”P” finns i problemsamlingen.
- De stjärnmarkerade uppgifterna är ofta lite mer utmanande; de kan kräva bättre räknefärdighet, eller en djupare teoretisk förståelse, eller att man måste få en ljus idé, osv.
- Till varje föreläsning finns läsanvisningar till kurslitteraturen, en översikt över vad som kommer att tas upp, samt diverse kommentarer. Dessa kommentarer är inte föreläsningsanteckningar, utan det rör sig ofta om saker som **inte** kommer att hinnas tas upp men som kan vara bra att ha på papper ändå: repetition, extra exempel, alternativa förklaringar, rättelser till boken, varningar för vanliga fel, anmärkningar om olika skrivsätt och konventioner, historiska detaljer, med mera.

[Tillbaka till föreläsnings- och lektionsplan.]

Föreläsning 1. Gränsvärde & kontinuitet

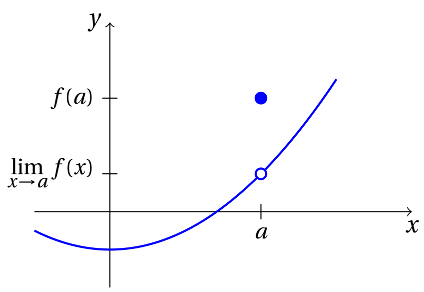
Innehåll

- Avsnitt 3.1–3.2 i kursboken, samt lite från början av avsnitt 3.3 (före Sats 3.5).
- Vi börjar med att säga *ungefär* vad gränsvärde och kontinuitet är, för att snabbt kunna komma igång och räkna på Lektion 1. De *exakta* definitionerna tas upp på nästa föreläsning.
- **Gränsvärdet**

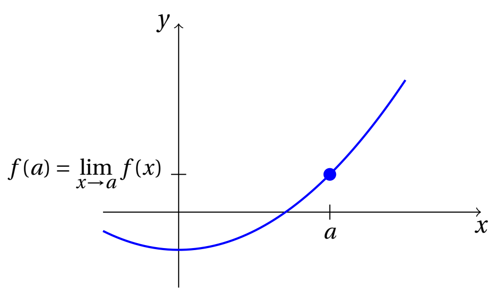
  $$\boxed{\lim_{x\to a} f(x)}$$

  (latin **limes**, engelska **limit**) är det värde som funktionen f "**borde ha**" i punkten a för att det ska stämma överens med funktionens värden i de omkringliggande punkterna. (Det kan hända att det inte finns något sådant värde, dvs. att gränsvärde saknas.)

  Ofta när man undersöker gränsvärdet när $x \to a$ så är det för att f är odefinierad i punkten a. Men det kan även hända att f redan har ett värde i punkten a. I så fall ska man ignorera detta värde $f(a)$ vid gränsvärdesberäkningen, och bara ta hänsyn till värdena $f(x)$ för $x \neq a$.

- Funktionen f sägs vara kontinuerlig i punkten a ifall den har "det rätta" värdet där, alltså om funktionsvärdet $f(a)$ är definierat (dvs. $a \in D_f$) **och** gränsvärdet $\lim_{x\to a} f(x)$ existerar **och** de överensstämmer med varandra:

  Om $a \in D_f$ och f inte är kontinuerlig i a så sägs f vara **diskontinuerlig** i a. (Det kan t.ex. se ut som i den första figuren ovan, där gränsvärdet existerar men $f(a)$ inte överensstämmer med detta värde.)

  Ifall $a \notin D_f$ är det meningslöst att tala om kontinuitet eller diskontinuitet i punkten a. (Vad man däremot *kan* fråga sig i det läget är om det går att *utvidga* funktionen, genom att tilldela den ett värde $f(a)$ i punkten a, så att den *blir* kontinuerlig i a.)

- Två synonyma skrivsätt:

  $$\lim_{x\to a} f(x) = \boxed{\text{gränsvärdet}} \qquad \text{("limes av } f(x) \text{ är lika med …")}$$

  eller

  $$f(x) \to \boxed{\text{gränsvärdet}} \quad \text{då} \quad x \to a \qquad \text{("} f(x) \text{ går mot …")}.$$

[Tillbaka till föreläsnings- och lektionsplan.]

- Andra sorters gränsvärden: **höger- och vänstergränsvärde** ($x \to a^{\pm}$), **gränsvärde i oändligheten** ($x \to \pm\infty$), **oegentligt gränsvärde** ($f(x) \to \pm\infty$).
- Några räknelagar. Några grundläggande enkla gränsvärden ("standardgränsvärden").
- Beräkning av mer komplicerade gränsvärden genom att kombinera standardgränsvärden med hjälp av räknelagar.

  De flesta gränsvärdesuträkningarna i denna kurs görs på detta sätt, och de bör redovisas så att det tydligt framgår vilka standardgränsvärden och räknelagar man har använt sig av.

  Om man vid en gränsvärdesundersökning har ett uttryck där ingen räkneregel är tillämpbar, t.ex. **en kvot där både täljaren och nämnaren går mot noll**, eller **en differens där båda termerna går mot oändligheten**, så får man försöka göra algebraiska omskrivningar för att komma runt det problemet.

Kommentarer

- Gränsvärdesuppgifter formuleras ofta "undersök $\lim_{x\to a} f(x)$". Detta betyder "räkna ut gränsvärdet om det existerar, eller visa att det inte existerar".
- Skrivsättet "$\lim_{x\to a} f(x) = A$" kan jämföras med frasen "mitt resmål är Spanien", medan skrivsättet "$f(x) \to A$ då $x \to a$" är som att säga "jag reser till Spanien". Många studenter mixar tyvärr dessa två skrivsätt på mer eller mindre hårresande vis, vilket leder till påståenden motsvarande "mitt resmål reser till Spanien" eller "jag är Spanien". Försök att undvika detta. Grundregeln är enkel: skriv det du menar, och inte något annat!
- Symbolerna $+\infty$ och ∞ är synonyma (i reell analys åtminstone). För det mesta skriver vi bara $x \to \infty$, men ibland kan man skriva $x \to +\infty$ om man vill betona eller kontrastera lite extra.
- Observera var plus- eller minustecknen skrivs när det gäller höger-/vänstergränsvärde respektive gränsvärden i oändligheten:

  $$x \to a^{\pm} \quad \longleftarrow \text{plus/minus står uppe till höger (betyder: } x \text{ går mot } a \in \mathbf{R} \text{ från höger/vänster på tallinjen)}$$

  $$x \to \underset{\substack{\uparrow \\ \text{plus/minus står framför (betyder: } x \text{ går obegränsat långt åt höger/vänster på tallinjen)}}}{\pm\infty}$$

  Aldrig såhär:

  $$x \to \infty^{\pm} \quad \longleftarrow \text{fel!}$$

  För vad skulle det betyda egentligen? Om $x \to \infty^{+}$, ska man då närma sig ∞ från *höger*, eller vad? Hur skulle det gå till?!?
- Det rekommenderade sättet att redovisa gränsvärdesuträkningar är såhär:

  $$\lim_{x\to a} f(x) = \boxed{\textit{Skriv resultatet här}}, \quad \text{eftersom}$$

  $$\begin{aligned} f(x) &= (\cdots) \\ &= (\cdots) && \textit{Gör lämpliga algebraiska omskrivningar} \\ &= (\cdots) && \textit{tills du kommer i ett läge där du kan tillämpa} \\ &= (\cdots) && \textit{standardgränsvärden och räknelagar.} \\ &\to \boxed{}, \quad \text{då } x \to a. && \textit{Sedan kan själva "gränsövergången" göras.} \end{aligned}$$

  Observera att alla gränsvärden beräknas i ett enda steg, dvs. det får bara finnas **en** gränsvärdespil "$\to$" i uträkningen. Och *efter* pilen, när man har låtit x gå mot a, får det inte finnas några x kvar. Alla "x" i uttrycket representerar ju samma tal, så om ett x går mot a måste alla andra x också göra det samtidigt.

[Tillbaka till föreläsnings- och lektionsplan.]

- Man kan också skriva såhär, men det är mindre lämpligt:

$\lim_{x\to a} f(x) = \lim_{x\to a}(\cdots)$	
$= \lim_{x\to a}(\cdots)$	*Obs. att det måste stå "$\lim_{x\to a}$" i varje steg fram till gränsövergången.*
$= \lim_{x\to a}(\cdots)$	
$= \lim_{x\to a}(\cdots)$	
$= \square$	*Vid gränsövergången ska det stå "=", inte "→". (**Gränsvärdet självt**, alltså $\lim_{x\to a}(\cdots)$, går ju ingenstans, det är **funktionen** som går mot gränsvärdet.) Skriv inte heller "då $x \to a$" efteråt, det är redundant.*

Det finns flera anledningar till att detta är sämre:

- Man måste skriva "$\lim_{x\to a}$" en massa gånger.
- Alla likheter "hänger i luften" tills man på slutet verkligen visar att gränsvärdet existerar. Räknelagarna är ju bara giltiga under vissa förutsättningar, t.ex. "**om** $f(x) \to A \in \mathbf{R}$ och $g(x) \to B \in \mathbf{R}$ så gäller $f(x)+g(x) \to A+B$", så om man ska vara riktigt petig kan man egentligen bara skriva
$$\cdots = \lim_{x\to a}\big(f(x)+g(x)\big) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = \cdots$$
eller
$$\cdots = \lim_{x\to a}\big(f(x)\,g(x)\big) = \Big(\lim_{x\to a} f(x)\Big)\Big(\lim_{x\to a} g(x)\Big) = \cdots$$
ifall man **redan har visat** att $\lim_{x\to a} f(x)$ och $\lim_{x\to a} g(x)$ existerar (som vanliga gränsvärden, inte oegentliga).
- Man kan få problem med detta skrivsätt om det dyker upp oegentliga gränsvärden (se kommentar nedan om redovisning av oegentliga gränsvärden).
- Resonemanget kan bli svårare att följa, eftersom det är frestande att börja blanda algebraiska omskrivningar och gränsvärdesuträkningar huller om buller. T.ex. ser man ibland följande typ av uträkning i samband med de mer avancerade standardgränsvärden som kommer lite senare i kursen:
$$\lim_{x\to 0}\frac{\sin 2x}{\ln(1+3x)} = \lim_{x\to 0}\underbrace{\frac{\sin 2x}{2x}}_{\substack{\to 1 \\ \text{(std.gr.v.)}}}\frac{2x}{\ln(1+3x)} \overset{?}{=} \lim_{x\to 0}\frac{2x}{\ln(1+3x)} = \cdots$$
Den rödmarkerade likheten är visserligen sann (vilket kan visas med produktlagen för gränsvärden), men så som det är redovisat ovan tycker jag **inte** att **motiveringen** är tillräckligt bra för att t.ex. godkännas på en tenta. För om du skriver så, hur vet läraren som ska bedöma tentan att du har tänkt rätt, och inte bara har gjort dig skyldig till den klassiska tankevurpan **"stegvis gränsövergång"**, där man låter $x \to a$ för vissa x men inte för alla? Som i det här fallet, där den rödmarkerade likheten **inte** är sann:
$$\text{(Fel!)}\qquad \lim_{x\to 0}(\underbrace{1+x}_{\to 1})^{1/x} \overset{?}{=} \lim_{x\to 0} 1^{1/x} = \lim_{x\to 0} 1 = 1.$$
(Det korrekta svaret är e enligt Sats 3.11e i boken.)

- Och såhär ska man definitivt **inte** skriva:

$\lim_{x\to a} f(x) = f(x)$	(Fel!)	*Utan "lim" i högerledet är likheten inte sann!*
$= (\cdots)$		*(Det står ju en konstant i VL och en funktion av x i HL.)*
$\to \square$		

[Tillbaka till föreläsnings- och lektionsplan.]

Detta är ett mycket vanligt fel, men skriv inte så är du snäll, det ser jättefult ut! Använd bara likhetstecken för att markera att två matematiska objekt (tal, funktioner, vektorer, etc.) verkligen är **lika**, inte som något slags allmänt skiljetecken som kan slängas in var som helst med betydelsen ”nu sätter jag igång” eller ”nu händer det något” eller dylikt. (Likaså med implikationspilen ”$\Longrightarrow$”, den ska bara användas för att markera att ett logiskt påstående **medför** ett annat.)

- En tumregel vid gränsvärdesuträkning är att ”bryta ut dominerande bidrag”. T.ex. såhär:

$$\lim_{x\to\infty}\frac{3x^3+4x^2}{5x^3+7x}=\frac{3}{5},\quad \text{eftersom}\quad \frac{3x^3+4x^2}{5x^3+7x}=\frac{x^3(3+\frac{4}{x})}{x^3(5+\frac{7}{x^2})}=\frac{3+\frac{4}{x}}{5+\frac{7}{x^2}}\to\frac{3+0}{5+0}=\frac{3}{5}\quad \text{då } x\to\infty,$$

eller

$$\lim_{x\to 0}\frac{3x^3+4x^2}{5x^3+7x}=0,\quad \text{eftersom}\quad \frac{3x^3+4x^2}{5x^3+7x}=\frac{x^2(3x+4)}{x(5x^2+7)}=x\cdot\frac{3x+4}{5x^2+7}\to 0\cdot\frac{0+4}{0+7}=0\cdot\tfrac{4}{7}=0\quad \text{då } x\to 0.$$

Ur täljare respektive nämnare bryter vi alltså ut en lämplig faktor (blått), med syftet att vi i varje parentes efter utbrytningen ska få kvar en nollskild konstant plus ”skräptermer” (rött) som **går mot noll** på slutet och inte påverkar gränsvärdet.

- Ibland behöver man bryta ut det dominerande bidraget ur ett *deluttryck* för att komma vidare. T.ex. finns det ingen räknelag för logaritmen av en summa, $\ln(A+B)$, men om man bryter ut en faktor ur summan $A+B$ så får man logaritmen av en produkt, och *det* finns det ju en räknelag för, $\ln(ab)=\ln a+\ln b$.

Exempel.

$$\lim_{x\to\infty}\frac{\ln(x^2+1)}{\ln x}=2,$$

eftersom

$$\frac{\ln(x^2+1)}{\ln x}=\frac{\ln\big(x^2(1+\frac{1}{x^2})\big)}{\ln x}=\frac{\ln(x^2)+\ln(1+\frac{1}{x^2})}{\ln x}=\frac{2\ln x+\ln(1+\frac{1}{x^2})}{\ln x}$$
$$=2+\underbrace{\tfrac{1}{\ln x}}_{\to 0}\cdot\underbrace{\ln(1+\tfrac{1}{x^2})}_{\to\ln(1+0)}\to 2+0\cdot 0=2\quad \text{då } x\to\infty.$$

- Angående **oegentligt gränsvärde**, kom ihåg att ∞ och $-\infty$ bara är symboler, inte reella tal.

Vi kan tillåta oss att skriva ”$\lim_{x\to a}f(x)=\infty$”, men enbart som synonym till ”$f(x)\to\infty$ då $x\to a$”; vi kan inte börja räkna med symbolen ∞ som om det vore ett vanligt tal. Det finns visserligen något som kallas för det ”utvidgade reella talsystemet”, som består av de reella talen tillsammans med symbolerna $\pm\infty$, samt *vissa* räkneoperationer definierade på denna mängd. Av pedagogiska skäl rekommenderas dock att du i denna kurs **inte** använder detta utvidgade talsystem för att skriva saker som ”$\infty+5=\infty$”, för då är det lätt hänt att man frestas att även börja skriva andra saker som **absolut inte** är acceptabla, t.ex. ”$\infty-\infty=0$”. De fyra räknesätten och andra räkneoperationer använder vi alltså bara på vanliga reella (eller komplexa) tal i denna kurs.

T.ex. i räknelagen

$$\lim_{x\to a}\big(f(x)+g(x)\big)=\lim_{x\to a}f(x)+\lim_{x\to a}g(x)$$

förutsätts det att det är vanliga gränsvärden man talar om, inte oegentliga gränsvärden.

Tyvärr är beteckningen ”lim” lite tvetydig. Om det inte framgår ur sammanhanget huruvida man tillåter oegentligt gränsvärde eller inte, så kan man behöva göra en kommentar om detta vid sidan av, t.ex. via uttryck som ”om gränsvärdet *si-och-så* existerar (ändligt), så gäller *det-och-det*”, där den instuckna parentesen med ordet ”ändligt” markerar att det måste vara ett vanligt gränsvärde, inte ∞ eller $-\infty$.

[Tillbaka till föreläsnings- och lektionsplan.]

- Oegentliga gränsvärden uppfyller egna räknelagar som formuleras separat (även om kursboken tyvärr inte är särskilt tydlig på den punkten, utan räknar med att läsaren själv ska kunna inse vilka lagar som gäller).

  Här är några stycken räknelagar för summor:

  - Om $f(x) \to \infty$ och $g(x) \to \infty$ så gäller $f(x) + g(x) \to \infty$.
  - Om $f(x) \to \infty$ och $g(x) \to A \in \mathbf{R}$ så gäller $f(x) + g(x) \to \infty$.
  - Om $f(x) \to \infty$ och $g(x)$ är **nedåt begränsad**[1] så gäller $f(x) + g(x) \to \infty$.
  - (Obs. att det inte finns någon räknelag för $f(x) + g(x)$ om $f(x) \to \infty$ och $g(x) \to -\infty$.)

  Och för produkter:

  - Om $f(x) \to \infty$ och $g(x) \to \infty$ så gäller $f(x)\, g(x) \to \infty$.
  - Om $f(x) \to \infty$ och $g(x) \to A \in \mathbf{R}$, där A är **positivt**, så gäller $f(x)\, g(x) \to \infty$.
  - Om $f(x) \to \infty$ och $g(x)$ är **nedåt begränsad bort från noll**[2] så gäller $f(x)\, g(x) \to \infty$.
  - Motsvarande regler med ombytta tecken på $f(x)$ och/eller $g(x)$ gäller såklart också, t.ex. om $f(x) \to \infty$ och $g(x) \to A \in \mathbf{R}$, där A är **negativt**, så gäller $f(x)\, g(x) \to -\infty$.
  - (Obs. att det inte finns någon räknelag för $f(x)\, g(x)$ om $f(x) \to \infty$ och $g(x) \to 0$.)

  Och för reciproka värden:

  - Om $f(x) \to \infty$ så gäller $1/f(x) \to 0$.
  - Om $f(x) \to 0$ och $f(x) > 0$ (**viktig förutsättning!**) så gäller $1/f(x) \to \infty$.

  Samt ett slags motsvarighet till instängningsregeln (”utknuffningsregeln”, kanske?!):

  - Om $f(x) \to \infty$ och $f(x) \le g(x)$ så gäller $g(x) \to \infty$.

  Man kan alltså t.ex. dra en sådan här slutsats:

$$\underbrace{\underbrace{f(x)}_{\to\infty} + \underbrace{g(x)}_{\to 43}}_{\to\infty} \to \infty, \qquad \text{då } x \to a.$$

  Men redovisa alltså **inte** detta på följande vis:

$$\lim_{x\to a}\Big(f(x) + g(x)\Big) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = \infty + 43 = \infty.$$

- Boken (s. 133) listar följande ”farliga typer av gränsvärden” där räknelag saknas:

$$\text{typ } \frac{0}{0}, \qquad \text{typ } \frac{\infty}{\infty}, \qquad \text{typ } 0\cdot\infty, \qquad \text{typ } \infty - \infty.$$

  Uttrycket ”gränsvärde av typ $0/0$” betyder att man söker gränsvärdet av ett bråk där man vet att täljaren och nämnaren går mot noll. Enbart utifrån denna kunskap kan man alltså **inte** dra någon slutsats, och samma för de andra typerna.

  Det finns dock ytterligare några farliga gränsvärdestyper som också bör nämnas:

$$\text{typ } 1^\infty, \qquad \text{typ } \infty^0, \qquad \text{typ } 0^0.$$

[1] ”Nedåt begränsad” betyder att det finns en konstant $C \in \mathbf{R}$ sådan att $g(x) \ge C$. Det räcker om detta gäller i närheten av det som x går mot, dvs i något punkterat intervall $0 < |x - a| < \delta$ ifall $x \to a \in \mathbf{R}$, eller i något intervall $x > \omega$ ifall $x \to \infty$. De föregående två räknelagarna är specialfall av denna.

[2] Dvs. det finns en konstant C sådan att $g(x) \ge C > 0$. Även här räcker det att detta gäller i närheten av det som x går mot.

[Tillbaka till föreläsnings- och lektionsplan.]

Angående det sistnämnda fallet finns det mycket goda skäl att låta $0^0 = 1$ *per definition* (även om en del mindre upplysta källor säger att 0^0 är odefinierat). Bl.a. är det ju så man vill tolka konstanttermen $a_0 x^0$ när man sätter in $x = 0$ i ett polynom $p(x) = \sum_{k=0}^{n} a_k x^k$, och binomialsatsen $(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$ med $x = 0$, $y = 1$ och $n \in \mathbf{N}$ har 1 i vänsterledet och 0^0 i högerledet, och antalet funktioner från en mängd med n element till en mängd med m element är m^n, vilket borde ge 1 för antalet funktioner från tomma mängden till sig själv, och så vidare... Men detta betyder alltså inte att ett *gränsvärde* av typen

$$(\text{någonting som går mot noll})^{(\text{någonting som går mot noll})}$$

måste bli 1. Ett trivialt motexempel är att $0^x \to 0 \neq 1$ då $x \to 0$.

- Det finns något som kallas **l'Hôpitals regel** (eller **l'Hospitals regel** med den ursprungliga franska 1600-talsstavningen av namnet), som kan användas för att beräkna vissa gränsvärden av typ 0/0 eller ∞/∞ med hjälp av derivering. På vissa platser (t.ex. USA) lär man ofta ut denna regel som ett slags universalverktyg för gränsvärden, men det gör vi inte här, utan snarare tvärtom:

  I denna kurs vill vi att ni **inte** använder l'Hôpitals regel!

  Det finns flera pedagogiska anledningar till detta:

  1. Erfarenheten från tusentals rättade tentor visar att det är *ytterst* sällsynt att de som trots denna uppmaning försöker använda l'Hôpitals regel lyckas göra detta korrekt, dvs. att verkligen redovisa en verifikation av att **alla** de nödvändiga **förutsättningarna** är uppfyllda.
  2. Att använda l'Hôpitals regel slentrianmässigt uppmuntrar inte till förståelse.
  3. Ofta när det gäller enklare gränsvärden här i TATA41 är det logiskt cirkelresonemang (eller åtminstone närapå) att använda l'Hôpitals regel, eftersom man vid deriveringen väsentligen använder det gränsvärde som man är ombedd att räkna ut.
  4. I TATA42 (Envariabelanalys 2) kommer ni att lära er att beräkna mer komplicerade gränsvärden med hjälp av s.k. Taylorutveckling, och då skulle ni ändå inte ha någon större nytta av l'Hôpitals regel längre.
     (I de fall där de inblandade funktionerna är såpass snälla att båda metoderna är tillämpbara så är det inte mera jobb att Taylorutveckla så långt som det behövs än att derivera så många gånger det behövs, och l'Hôpitals regel kan i detta fall bevisas på en rad med Taylorutveckling. Regelns verkliga användbarhet ligger snarare i att den kan tillämpas i teoretiska sammanhang under svagare förutsättningar, men det är inget som ni kommer att råka ut för här, och i detta allmänna fall är beviset inte alls så enkelt.)

Lektion 1

Påminnelse: B-problem finns i boken, P-problem i problemsamlingen. Stjärnmarkerade uppgifter är "extraproblem" (oftast lite mer krävande).

- En liten påminnelse om absolutbelopp: P3.2.
- Blandade gränsvärdesproblem: B3.1, B3.2, B3.7, B3.10, B3.8 (instängningsregeln, se Sats 3.3), P3.5a, P3.6, B3.11 (ledtråd till b-uppgiften: förläng **inte** med konjugatet!), B3.12.
- Höger- och vänstergränsvärde: B3.14, B3.15, P3.11.
- Gränsvärde med variabeln x i både basen och exponenten: P3.12.

  Ett allmänt tips i sådana fall är att skriva om uttrycket såhär:

  $$f(x)^{g(x)} = \left(e^{\ln f(x)}\right)^{g(x)} = e^{g(x)\ln f(x)}.$$

  Då får man alla x samlade uppe i exponenten, vilket är enklare att hantera.
- Inledande övningar om kontinuitet: B3.17, B3.18, B3.19, B3.20.
- Limes av limes (spelar ordningen roll?): P3.7*.

[Tillbaka till föreläsnings- och lektionsplan.]

Föreläsning 2. Gränsvärde & kontinuitet, forts.

Innehåll

- Samma avsnitt i boken som till Föreläsning 1.
- Noggrann definition av *exakt* vad som menas med gränsvärde. Detta är kanske det svåraste momentet i hela kursen, så var beredd på att det troligen ligger på en rejält högre svårighetsnivå än all matematik du har sett förut. Försök att tänka igenom Definition 3.2 och 3.3 i förväg, så har du en bättre chans att förstå vad som sägs på föreläsningen.
- Några exempel på hur man bevisar saker utifrån gränsvärdesdefinitionen.

Kommentarer

- **Definition av gränsvärde.** (Def. 3.2a i boken.) Skrivsättet

$$\boxed{f(x) \to A \qquad \text{då } x \to a}$$

(där $a \in \mathbf{R}$ och $A \in \mathbf{R}$) betyder att funktionen f har följande två egenskaper, alltså att **både villkor G1 och villkor G2** gäller:

G1. Definitionsmängden D_f innehåller punkter godtyckligt nära a (men skilda från a), dvs. för varje $\delta > 0$ existerar $x \in D_f$ som uppfyller $0 < |x - a| < \delta$.
(Synonymt uttryckssätt: punkten a är en **hopningspunkt** till mängden D_f.)

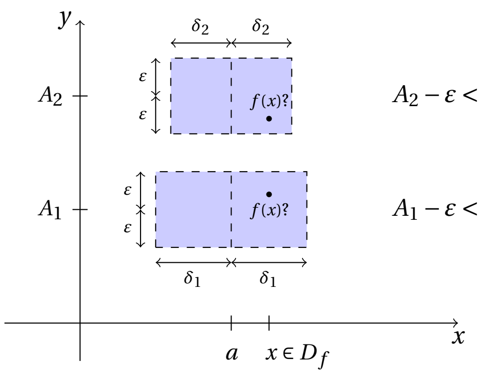
G2. För varje $\varepsilon > 0$ finns $\delta > 0$ sådant att $\left|f(x) - A\right| < \varepsilon$ för alla $x \in D_f$ sådana att $0 < |x - a| < \delta$.

- Observera att denna definition inte säger någonting om hur man ska bära sig åt för att *hitta* ett sådant tal A (om det finns); det är en senare fråga. Den säger bara *vad som måste gälla* för att talet A ska räknas som gränsvärde av $f(x)$ då $x \to a$. För varje givet tal A kan man alltså ställa en ja/nej-fråga till definitionen, och få ett otvetydigt svar: är det sant eller inte att $f(x)$ går mot just detta tal A då $x \to a$?

- Boken saknar följande **entydighetssats** om att en funktion kan ha *högst ett gränsvärde* då $x \to a$:

Sats. Om $f(x) \to A_1$ då $x \to a$ och $f(x) \to A_2$ då $x \to a$, så är $A_1 = A_2$.

Bevis. För att härleda en motsägelse, anta att $A_1 \neq A_2$. Låt ε vara ett positivt tal som är mindre än halva avståndet mellan A_1 och A_2, till exempel $\varepsilon = \frac{1}{3}|A_1 - A_2|$. Det första antagandet "$f(x) \to A_1$ då $x \to a$" medför (enligt definitionen) att olikheten $A_1 - \varepsilon < f(x) < A_1 + \varepsilon$ gäller för alla $x \in D_f$ i något punkterat intervall $0 < |x - a| < \delta_1$. Och det andra antagandet "$f(x) \to A_2$ då $x \to a$" medför på samma sätt att $A_2 - \varepsilon < f(x) < A_2 + \varepsilon$ för alla $x \in D_f$ i något punkterat intervall $0 < |x - a| < \delta_2$. För de $x \in D_f$ som uppfyller $0 < |x - a| < \min(\delta_1, \delta_2)$ skulle alltså $f(x)$ *samtidigt* behöva ligga i intervallet $]A_1 - \varepsilon, A_1 + \varepsilon[$ och i intervallet $]A_2 - \varepsilon, A_2 + \varepsilon[$. Men detta är omöjligt, för dessa intervall överlappar inte varandra (se figuren nedan); vi valde ju ε så litet att avståndet mellan A_1 och A_2 är större än 2ε. □



$$A_2 - \varepsilon < f(x) < A_2 + \varepsilon \quad \text{för} \quad 0 < |x - a| < \delta_2$$

$$A_1 - \varepsilon < f(x) < A_1 + \varepsilon \quad \text{för} \quad 0 < |x - a| < \delta_1$$

[Tillbaka till föreläsnings- och lektionsplan.]

- Motsvarande sats (med liknande bevis) gäller även för gränsvärden då $x \to \infty$ eller $x \to -\infty$, samt för oegentliga gränsvärden (t.ex. kan man inte samtidigt ha både $f(x) \to \infty$ och $f(x) \to A \in \mathbf{R}$ då $x \to a$).

- Det är denna entydighet som gör att det alternativa skrivsättet

$$\boxed{\lim_{x \to a} f(x) = A}$$

  är meningsfullt. För om $f(x)$ skulle kunna ha *flera* gränsvärden då $x \to a$ skulle man inte kunna tala om **gränsvärdet** i bestämd form singularis.

- En annan enkel sak som inte visas i boken är följande extremt grundläggande standardgränsvärde:

$$x \to a \text{ då } x \to a.$$

  Med detta skrivsätt ser det ju helt självklart ut, men man måste faktiskt verifiera att funktionen $f(x) = x$ och talet $A = a$ uppfyller definitionen av vad det betyder att $f(x) \to A$ då $x \to a$. Detta är dock inte svårt: för varje $\varepsilon > 0$ duger $\delta = \varepsilon$. (Eller hur?)

  Ännu enklare är fallet med en konstant funktion $f(x) = c$ (eftersom *varje* $\delta > 0$ duger, oavsett ε):

$$c \to c \text{ då } x \to a.$$

Föreläsning 3. Satser om kontinuerliga funktioner

Innehåll

- Resten av avsnitt 3.3 i kursboken.
- Exakt definition av kontinuitet.
- Kontinuitet hos invers funktion.
- De elementära funktionerna är kontinuerliga.
- Satsen om största och minsta värde.
- Satsen om mellanliggande värde.

Kommentarer

- **Definition av kontinuitet.** (Def. 3.5 i boken.) Funktionen f sägs vara **kontinuerlig** i punkten a ifall den har endera av följande två (ömsesidigt uteslutande) egenskaper, alltså att **antingen villkor K1 eller villkor K2** gäller:

  K1. Funktionsvärdet $f(a)$ är definierat, gränsvärdet $\lim_{x \to a} f(x)$ existerar, och $f(a) = \lim_{x \to a} f(x)$.

  Med andra ord: $\boxed{f(x) \to f(a) \text{ då } x \to a}$.

  K2. Funktionsvärdet $f(a)$ är definierat, men a är en **isolerad punkt** i D_f (dvs. det finns ett tal $\delta > 0$ sådant att det punkterade intervallet $0 < |x - a| < \delta$ inte innehåller några punkter $x \in D_f$).

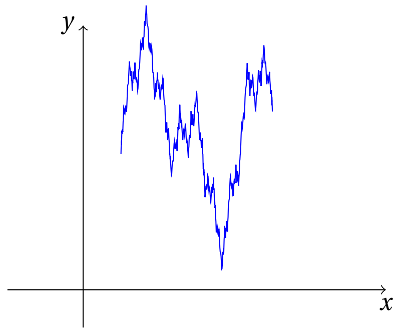
  Och funktionen f sägs vara **kontinuerlig** (rätt och slätt) ifall den är kontinuerlig i *varje* punkt $a \in D_f$.

- Intuition: För kontinuitet i en punkt krävs för det första att ”de omkringliggande punkterna kan enas om vilket värde $f(a)$ borde ha”, dvs. att gränsvärdet existerar, och för det andra krävs att $f(a)$ också ”lyder grupptrycket” och *har* detta värde. (Detta är fall K1, ”huvudfallet”.) *Förutom* om ”ingen av de omkringliggande punkterna *har* någon åsikt”, för då kan $f(a)$ få vara vad det vill, ingen bryr sig! (Detta är fall K2, ”undantagsfallet”.)

[Tillbaka till föreläsnings- och lektionsplan.]

- I boken visas att alla de välbekanta funktionerna från grundkursen är kontinuerliga (polynom, rationella, exp & log, trig & arcus, belopp).

- Om någon ber dig tänka på en typisk kontinuerlig funktion ska du dock inte tänka på ett polynom eller dylikt, utan på någonting ”elakare”, som en kurva över en aktiekurs, ritad av en ”oändligt darrhänt” person, där det förblir taggigt oavsett hur mycket man zoomar in:

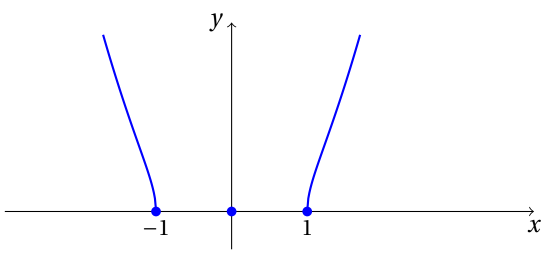


- Ofta sägs att en kontinuerlig funktion är en funktion ”vars graf kan ritas utan att lyfta pennan”, men detta är lite vilseledande! Ifall f:s definitionsmängd inte är ett intervall så *måste* man ju lyfta pennan för att rita grafen, eftersom man inte får rita in något värde i punkter som inte tillhör definitionsmängden.

  Ta t.ex. funktionen $f(x) = 1/x$, definierad för $x \neq 0$. Denna funktion är kontinuerlig, eftersom den är kontinuerlig i varje punkt i sin definitionsmängd. Att den har en singularitet i punkten $x = 0$ är irrelevant. (Man bör f.ö. inte säga att f har en ”diskontinuitet” i punkten $x = 0$, även om det ofta slarvas med detta, även bland matematiker. Det vore ju konstigt om en kontinuerlig funktion skulle ha diskontinuiteter, eller hur? Begreppen ”kontinuerlig i a” och ”diskontinuerlig i a” bör endast tillämpas på punkter a som tillhör funktionens definitionsmängd.)

- Det där undantagsfallet K2 med ”isolerad punkt i D_f” kan verka konstigt, och i boken är det förvirrande nog tillagt som en kommentar *efter* Def. 3.5, som om det inte riktigt hörde till definitionen. Men det måste vara med, om följande viktiga sats ska vara sann: **en sammansättning av kontinuerliga funktioner är kontinuerlig**.

  Exempel. Funktionen $f(x) = \sqrt{x^4 - x^2} = \sqrt{x^2(x^2 - 1)}$ är definierad för $x^2(x^2 - 1) \geq 0$, dvs. för $x = 0$ eller $x \geq 1$ eller $x \leq -1$. Vi har alltså $0 \in D_f$, men det är en isolerad punkt i D_f, eftersom punkterna i de omgivande intervallen $]-1, 0[$ och $]0, 1[$ *inte* tillhör D_f. Eftersom funktionen f är en sammansättning av kontinuerliga funktioner (polynom och kvadratrot) vill vi väldigt gärna att den ska räknas som en kontinuerlig funktion, dvs. kontinuerlig i alla punkter i dess definitionsmängd, och i synnerhet kontinuerlig i punkten 0. Men $\lim_{x \to 0} f(x)$ existerar inte, för villkor 1 i gränsvärdesdefinitionen ovan (att D_f ska innehålla punkter godtyckligt nära 0) är inte uppfyllt. Alltså uppfyller inte denna funktion det första kontinuitetsvillkoret $f(x) \to f(0)$ då $x \to 0$, men den räknas ändå som kontinuerlig i 0, eftersom det andra kontinuitetsvillkoret är uppfyllt, dvs. att 0 är en isolerad punkt i D_f.



Exempel. För att ta ett ännu mer extremt (och kanske lite löjligt) exempel så är funktionen $f(x) = \sqrt{-1-x^2}$ inte definierad för något $x \in \mathbf{R}$, dvs. $D_f = \emptyset$ (tomma mängden). Men den är ju bildad genom sammansättning av kontinuerliga funktioner, så borde den inte vara kontinuerlig enligt den där satsen?!? Ja, det är den faktiskt! Definitionen av att vara kontinuerlig är att f ska vara kontinuerlig i varje punkt i D_f, och om det inte *finns* några punkter i mängden D_f så räknas detta villkor enligt logikens lagar som ”tomt sant” (eng. *vacuously true*) – det finns ju inga punkter som skulle kunna *falsifiera* påståendet.

- Ofta definieras kontinuitet på ett (skenbart) annorlunda sätt än i kursboken:

  Alternativ definition av kontinuitet. Funktionen f sägs vara kontinuerlig i punkten a ifall den har följande egenskap:

  > Funktionsvärdet $f(a)$ är definierat, och för varje $\varepsilon > 0$ finns $\delta > 0$ sådant att $\left|f(x) - f(a)\right| < \varepsilon$ för alla $x \in D_f$ sådana att $|x-a| < \delta$.

  Intuitionen här är att man genom att hålla ändringen av indatavärdet (från a till x) tillräckligt liten kan göra motsvarande ändring av utdatavärdet (från $f(a)$ till $f(x)$) så liten som man vill.

  Man kan visa att **denna definition är ekvivalent med bokens definition**. Beviset är ”trivialt” i den meningen att det enbart består i att jämföra med gränsvärdesdefinitionen och tänka efter noga vad som sägs egentligen. Det krävs med andra ord ingen inspirerad snilleblixt för att kunna bevisa påståendet. Men detta är inte samma sak som att det är ”enkelt”, åtminstone inte första gången man ser det, för vem har sagt att det är enkelt att tänka efter noga?

- Är följande påstående sant?

  > Gränsvärdet $\lim_{x\to a} f(x)$ existerar om och endast om högergränsvärdet $\lim_{x\to a^+} f(x)$ och vänstergränsvärdet $\lim_{x\to a^-} f(x)$ båda existerar och är lika.

  Nej, inte riktigt! Det är sant om funktionens definitionsmängd D_f innehåller punkter godtyckligt nära a, **på båda sidorna** om a. Men säg att funktionen är **odefinierad** i något intervall $]a-\delta, a[$ till vänster om a, medan varje intervall av formen $]a, a+\delta[$ till höger om a innehåller punkter ur D_f (där $\delta > 0$ i båda fallen). Då är det meningslöst att ens tala om vänstergränsvärde, och begreppen **gränsvärde** och **högergränsvärde** kommer att sammanfalla! Dvs. för en sådan funktion gäller istället följande påstående:

  > Gränsvärdet $\lim_{x\to a} f(x)$ är exakt samma sak som högergränsvärdet $\lim_{x\to a^+} f(x)$.

  Och på samma sätt sammanfaller gränsvärde och vänstergränsvärde ifall funktionen är odefinierad till höger om a.

  Exempel. Vi brukar ju säga att funktionen $f(x) = \sqrt{x}$, $x \ge 0$, är kontinuerlig. Detta ska gälla i varje punkt i definitionsmängden, så i synnerhet ska det gälla i punkten $x = 0$:

  $$\sqrt{0} = \lim_{x\to 0} \sqrt{x}.$$

  Och detta är sant; gränsvärdet existerar (och är noll), men inte för att höger- och vänstergränsvärdena är lika, utan för att vänstergränsvärdet saknar mening och gränsvärdet därför är lika med högergränsvärdet:

  $$\lim_{x\to 0} \sqrt{x} = \lim_{x\to 0^+} \sqrt{x} = 0.$$

- Sats 3.7 i boken, om gränsvärde hos sammansatt funktion, är ej helt korrekt formulerad. Det kan nämligen finnas problem med definitionsmängden. Ta t.ex. $f(x) = \sqrt{x}$ och $g(x) = -x^2$. Enligt bokens version av satsen skulle då den sammansatta funktionen $f(g(x))$ gå mot noll då $x \to 0$:

  $$f(g(x)) = f(\underbrace{-x^2}_{\to 0}) \to f(0) = \sqrt{0} = 0,$$

eftersom f är kontinuerlig i punkten 0. Problemet är bara att den sammansatta funktionen $f(g(x)) = \sqrt{-x^2}$ är **odefinierad** för alla $x \neq 0$, så den saknar gränsvärde då $x \to 0$!

Istället för att försöka fixa till denna sats kan vi konstatera att vad som **alltid är sant**, och som också är mycket viktigare, är att **sammansättningen av två kontinuerliga funktioner är kontinuerlig**. Boken hänvisar till Sats 3.7 som argument för detta, men det kan istället visas ganska enkelt direkt ur den alternativa definitionen av kontinuitet som gavs i kommentarerna till Föreläsning 2.

På så vis undgår man problemet med att försöka formulera en korrekt sammansättningsregel för gränsvärden, vilket alltså är lite smålurigt.

- Frasen ”anta att f är kontinuerlig på $[a, b]$” förekommer i flera av satserna i detta avsnitt. Men vad exakt *menar* man med denna fras? Detta sägs aldrig i kursboken (och detsamma gäller många andra källor).

  Den första tolkningen som man skulle gissa på är väl att det betyder att f är kontinuerlig i varje punkt $x \in [a, b]$.

  Men ta t.ex. funktionen

  $$f(x) = \begin{cases} x^2, & \text{om } 3 \le x \le 5, \\ 0, & \text{för övriga } x \in \mathbf{R}. \end{cases}$$

  Är denna funktion ”kontinuerlig på $[3, 5]$”? Nej, inte enligt ovanstående tolkning, för f är ju diskontinerlig i punkterna $x = 3$ och $x = 5$, som ju båda tillhör intervallet $[3, 5]$.

  Å andra sidan är det uppenbart att f:s värden i punkter utanför intervallet $[3, 5]$ är helt irrelevanta ifall man t.ex. skulle vilja undersöka om f har ett största eller minsta värde på $[3, 5]$. Så man kanske vill att den här funktionen faktiskt *ska* räknas som ”kontinuerlig på $[3, 5]$” så att man kan tillämpa bokens formulering av satsen om största och minsta värde på den?

  De böcker[3] som faktiskt är noggranna nog att definiera vad de menar med ”kontinuerlig på $[a, b]$” brukar därför säga att det betyder att f:s **restriktion** till $[a, b]$ är en kontinuerlig funktion.

  I vårt exempel syftar detta alltså på funktionen

  $$g(x) = x^2, \qquad x \in [3, 5],$$

  vars definitionsmängd är *exakt* intervallet $[3, 5]$, och som överensstämmer med $f(x)$ i detta intervall. Och denna funktion **är** kontinuerlig, ty den är kontinuerlig i varje punkt i sin definitionsmängd.[4] Så *med den tolkningen* är det *sant* att f är ”kontinuerlig på $[3, 5]$”.

  (Jag kan dock inte lova att du skulle få detta svar av varje matematiker som du frågar! Somliga skulle nog ändå förespråka den första tolkningen ovan. Det finns alltså en viss tvetydighet i denna fras. För *öppna* intervall $]a, b[$ uppstår inte detta tvetydighetsproblem, utan då sammanfaller de två tolkningarna. Men t.ex. i satsen om största och minsta värde är det ju viktigt att det är ett *slutet* och begränsat intervall.)

  Andra böcker undviker problemet genom att formulera satsen om största och minsta värde såhär:

  Sats. Om $f\colon [a, b] \to \mathbf{R}$ är kontinuerlig så har f ett största och ett minsta värde på $[a, b]$.

  Här anger skrivsättet ”$f\colon [a, b] \to \mathbf{R}$” att f:s definitionmängd ska vara *exakt* $[a, b]$. Då är det upp till användaren av satsen att själv ta restriktionen till $[a, b]$ ifall hen vill tillämpa denna sats på en funktion där $[a, b]$ är en strikt delmängd av D_f.

[3] T.ex. *Calculus* av M. Spivak, eller *An Introduction to Classical Real Analysis* av K. Stromberg.

[4] Det är uppenbart att g är kontinuerlig i det inre av intervallet, alltså för $3 < x < 5$. I den vänstra ändpunkten $x = 3$ gäller $g(3) = \lim_{x \to 3} g(x)$ eftersom g är odefinierad för $x < 3$ så att gränsvärdet sammanfaller med högergränsvärdet, $g(3) = \lim_{x \to 3^+} g(x)$. Analogt i den högra ändpunkten $x = 5$.

[Tillbaka till föreläsnings- och lektionsplan.]

Lektion 2

- Övningar på gränsvärdesdefinitionen: P3.8ab, P3.8c*, B3.6, B3.5*.
- Mera kontinuitet: P3.21, B3.51, B3.27*, P3.13*.
- Extrauppgift utan nummer: Är följande påstående **sanna** eller **falska**?

  (Facit finns längs bak, men gör helst alla deluppgifterna innan du tittar där.)

  (a) Gränsvärdet $\lim_{x\to a} f(x)$ existerar om och endast om höger- och vänstergränsvärdena $\lim_{x\to a^+} f(x)$ och $\lim_{x\to a^-} f(x)$ existerar och är lika.

  (b) Högergränsvärdet $\lim_{x\to 0^+} \sqrt{x}$ existerar.

  (c) Vänstergränsvärdet $\lim_{x\to 0^-} \sqrt{x}$ existerar.

  (d) Gränsvärdet $\lim_{x\to 0} \sqrt{x}$ existerar.

  (e) Funktionen f är kontinuerlig i punkten a om och endast om
  $$\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x).$$

  (f) Funktionen f är kontinuerlig i punkten a om och endast om
  $$\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = f(a).$$

  (g) Funktionen $f(x) = \sqrt{x}$, $x \geq 0$, är kontinuerlig i punkten $x = 0$.

  (h) Funktionen $f(x) = \sqrt{x}$, $x \geq 0$, är kontinuerlig.

- Satsen om största och minsta värde: B3.21.
- Satsen om mellanliggande värde: P3.14a, P3.14b*, B3.22, B3.23, B3.24, B3.25, B3.26.

Föreläsning 4. Standardgränsvärden och talföljder

Innehåll

- Avsnitt 3.4 och 3.5 i kursboken.

  Underavsnittet ”Rekursiva talföljder” är överkurs. Det kan dock vara bra att känna till Sats 3.16 som står där: varje växande och uppåt begränsad talföljd $(a_n)_{n=0}^{\infty}$ har ett gränsvärde $\lim_{n\to\infty} a_n \in \mathbf{R}$.

- Standardgränsvärden då $x \to 0$, t.ex. $\lim_{x\to 0} \dfrac{\sin x}{x} = 1$ och $\lim_{x\to 0} \dfrac{e^x - 1}{x} = 1$.
- Standardgränsvärden då $x \to \infty$, t.ex. $\lim_{x\to\infty} \dfrac{e^x}{x} = \infty$ och $\lim_{x\to\infty} \dfrac{x}{\ln x} = \infty$.
- Ej på föreläsningen, men i boken: Standardgränsvärden (för talföljder) där $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ ingår.

Kommentarer

- Förväxla inte standardgränsvärdet $\lim_{x\to 0} \frac{\sin x}{x} = 1$ med gränsvärdet $\lim_{x\to\infty} \frac{\sin x}{x} = 0$, som ju beräknas med Sats 3.1 (nollgående gånger begränsad):
  $$\frac{\sin x}{x} = \underbrace{\frac{1}{x}}_{\to 0} \cdot \underbrace{\sin x}_{\substack{\text{begränsad} \\ \text{(mellan } -1 \text{ och } 1)}} \to 0, \quad \text{då } x \to \infty.$$

[Tillbaka till föreläsnings- och lektionsplan.]

- Titta noga i Sats 3.13b:
$$\frac{x^\alpha}{a^x} \to 0 \quad \text{då } x \to \infty, \text{ om konstanten } a > 1.$$
Det står inte samma konstant i täljaren och nämnaren! I täljaren står det x upphöjt till den **grekiska bokstaven alfa**, men i nämnaren står det **vanligt latinskt a** (i kursiv stil, som kutymen är med variabler) upphöjt till x. Och det är a:et som ska vara större än 1.

  Det hade antagligen varit tydligare att välja någon annan bokstav än alfa, t.ex. c:
$$\frac{x^c}{a^x} \to 0 \quad \text{då } x \to \infty, \text{ om konstanten } a > 1.$$
Och ännu bättre vore det kanske att formulera detta standardgränsvärde i den reciproka formen
$$\frac{a^x}{x^c} \to \infty \quad \text{då } x \to \infty, \text{ om konstanten } a > 1,$$
ty denna version medför automatiskt den som står i boken, *men inte tvärtom*. Reciproka värdet av någonting som går mot ∞ måste ju automatiskt gå mot noll. Men reciproka värdet av någonting som går mot noll måste *inte* automatiskt gå mot oändligheten, denna slutsats kan man bara dra ifall uttrycket som går mot noll är *positivt*. Och så är det ju i det här fallet, men man vill inte behöva *påpeka* detta varenda gång man vill använda standardgränsvärdet ”uppochnedvänt”. (Samma sak för Sats 3.13a.)

- Ett intressant gränsvärde som saknas i Sats 3.14 är
$$\sqrt[n]{n!} \to \infty \quad \text{då } n \to \infty.$$
Bevis. För jämna $n = 2k$ är
$$n! = \underbrace{1 \cdot 2 \cdot \ldots \cdot k}_{\geq 1} \cdot \underbrace{(k+1)\cdot(k+2)\cdot\ldots\cdot(2k)}_{> k \cdot k \cdot \ldots \cdot k = k^k} > k^k$$
så
$$\sqrt[n]{n!} > (k^k)^{1/(2k)} = k^{1/2} \to \infty \quad \text{då } k \to \infty.$$
Och för udda $n = 2k+1$ är
$$n! = \underbrace{1 \cdot 2 \cdot \ldots \cdot k}_{\geq 1} \cdot \underbrace{(k+1)\cdot(k+2)\cdot\ldots\cdot(2k)\cdot(2k+1)}_{> k \cdot k \cdot \ldots \cdot k \cdot k = k^{k+1} \geq k^{k+\frac{1}{2}}} > k^{k+\frac{1}{2}}$$
så
$$\sqrt[n]{n!} > (k^{k+\frac{1}{2}})^{1/(2k+1)} = k^{1/2} \to \infty \quad \text{då } k \to \infty. \qquad \square$$
Alternativt bevis. Låt ω vara ett godtyckligt tal. Eftersom $\omega^n/n! \to 0$ då $n \to \infty$ (enligt Sats 3.14d) så finns det ett tal N sådant att $\omega^n/n! < 1$ för $n \geq N$. Med andra ord är $n! > \omega^n$ för $n \geq N$, dvs. $\sqrt[n]{n!} > \omega$ för $n \geq N$. Detta är (per definition) vad det betyder att $\sqrt[n]{n!} \to \infty$ då $n \to \infty$. $\square$

Lektion 3

- Standardgränsvärden i origo: P3.18, P3.19 (sätt $x = 1 + t$ i b-uppgiften, för att få ett gränsvärde där $t \to 0$).
- Gör variabelbyten för att få standardgränsvärden: P3.23.
- Fler standardgränsvärden i origo: B3.29, B3.30.
- Lurigt gränsvärde innefattande invers funktion: P3.15*. (Tips: variabelbyte.)

[Tillbaka till föreläsnings- och lektionsplan.]

- Ytterligare några gränsvärden med kvadratrötter: P3.24, B3.48*.
- Sned asymptot: P3.22*.

  Vad detta problem frågar efter (utan att säga det uttryckligen) är en asymptot till kurvan $y = \sqrt{x^2 + x}$, dvs. en linje $y = Ax + B$ som kurvan närmar sig då $x \to \infty$. Jfr. Exempel 3.35 och kommentarerna därefter.

  Sneda asymptoter är lite överkurs här i TATA41, i den meningen att vi på tentauppgifter om kurvritning inte begär att ni ska undersöka om det finns sneda asymptoter $y = Ax + B$ med $A \neq 0$; det räcker att hitta vågräta asymptoter $y = B$ och lodräta asymptoter $x = C$. (Mer om detta senare i kursen.)

Lektion 4

- Standardgränsvärden i origo (ev. efter variabelbyte): B3.31ace.
- Dominerande term då $x \to 0$ resp. $x \to \pm\infty$: B3.32.

  Idén är samma i samtliga fall: bryt ut den dominerande termen ur täljare och nämnare. Men vilken term som dominerar beror på vad x går mot!
- Standardgränsvärden i oändligheten (m.m.): B3.34.
- Gränsvärde av talföljd: P3.10, B3.38, P3.16*.
- Klassiskt gränsvärde (som ibland tas som definition av talet e): P3.20.
- Gränsvärde av geometrisk summa när antalet termer går mot oändligheten: B3.41*.

  En sådan här ”summa med oändligt många termer”,

  $$\sum_{k=0}^{\infty} x^k = \lim_{n\to\infty} \sum_{k=0}^{n} x^k,$$

  kallas för en **geometrisk serie** (med kvot x) och kommer att spela en fundamental roll i TATA42 (Envariabelanalys 2).
- Överskattning av cirkelns area med omskrivna polygoner: B3.54*.

Föreläsning 5. Derivata

Föreläsning 6. Derivata, forts.

Innehåll

- Avsnitt 4.1–4.3, samt delar av avsnitt 4.4.
- Definition av derivata och deriverbarhet.
- Deriverbarhet medför kontinuitet.
- Räknelagar (produktregeln, kedjeregeln, etc.).
- Derivator av de elementära funktionerna.
- Derivata av invers funktion.
- Lokalt extremvärde (lokalt maximum/minimum). Derivatan i en lokal extrempunkt måste (om den existerar) ha värdet noll.
- Rolles sats. Medelvärdessatsen för derivator.

[Tillbaka till föreläsnings- och lektionsplan.]

Kommentarer

- **Definition av deriverbarhet.** (Def. 4.1 i boken.) Funktionen f sägs vara **deriverbar** i punkten a ifall den har följande två egenskaper, alltså att **både villkor D1 och villkor D2** gäller:

  D1. Funktionen f är definierad i en **omgivning** av a.

  Vad detta betyder är att det finns något $\delta > 0$ sådant att **varje** punkt i intervallet $a-\delta < x < a+\delta$ tillhör definitionsmängden D_f. Med andra ord ska detta intervall vara en delmängd av D_f, dvs. villkoret kan uttryckas såhär:

  $$\text{Det existerar } \delta > 0 \text{ sådant att }]a-\delta, a+\delta[\subseteq D_f.$$

  D2. Gränsvärdet

  $$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \quad \left(\text{eller } \lim_{x\to a} \frac{f(x)-f(a)}{x-a}, \text{ vilket är samma sak}\right)$$

  existerar ändligt (dvs. oegentligt gränsvärde ∞ eller $-\infty$ räknas inte som tillåtet här).

  Om dessa två villkor är uppfyllda så kallas gränsvärdet i villkor D2 för f:s **derivata** i a, och betecknas $f'(a)$:

  $$\boxed{f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = \lim_{x\to a} \frac{f(x)-f(a)}{x-a}}.$$

  (Flera andra skrivsätt förekommer också, t.ex $\frac{df}{dx}(a)$ eller $Df(a)$. Observera skillnaden i skrivsätt mellan definitionsmängden D_f och derivatan Df.)

- För att det ska vara meningsfullt att tala om gränsvärdet i villkor D2 måste villkor G1 i gränsvärdesdefinitionen vara uppfyllt, dvs. det måste finnas punkter x godtyckligt nära a där uttrycket $\frac{f(x)-f(a)}{x-a}$ är definierat. Och dessutom måste såklart värdet $f(a)$ (som ingår i uttrycket ifråga) vara definierat.

  Villkor G1 kräver dock inte att uttrycket är definierat för **alla** x tillräckligt nära a, vilket är vad villkor D1 kräver. Anledningen till att man ställer detta starkare krav D1 i derivatans definition är bl.a. att förutsättningarna för **kedjeregeln** då blir enklare; se nedan.

- Benämningen ”derivata” kommer från att derivatan f' är ”den härledda funktionen”, dvs. den funktion som härleds från f. Ordets rötter går tillbaka till att ”avleda”, i betydelsen leda bort vatten.[5]

  Såhär heter det på några andra språk:

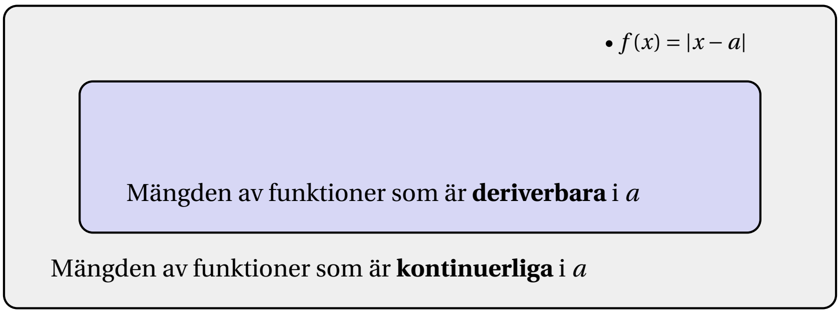
svenska	engelska	tyska	franska
att derivera	**to differentiate** (obs!)	ableiten	dériver
att härleda	to derive	ableiten/herleiten	dériver
derivatan	the derivative	die Ableitung	la dérivée
härledningen	the derivation	die Ableitung/Herleitung	la dérivation

  Engelskans *differentiate* syftar på att beräkna ”differentialen” $df = f'(x)\,dx$. (Jfr. ordet ”differentialkalkyl”.) En äldre term för derivatan $f'(x)$ är ”differentialkoefficienten”, alltså den koefficient som står framför dx i differentialen df. Man kan också säga **to take the derivative of [something]**.

- Sats 4.1, ”om f är deriverbar i punkten a så är f också kontinuerlig i a”, kan illustreras med ett Venndiagram, där man också kan lägga in exemplet $f(x) = |x-a|$ som visar att den omvända implikationen inte gäller:

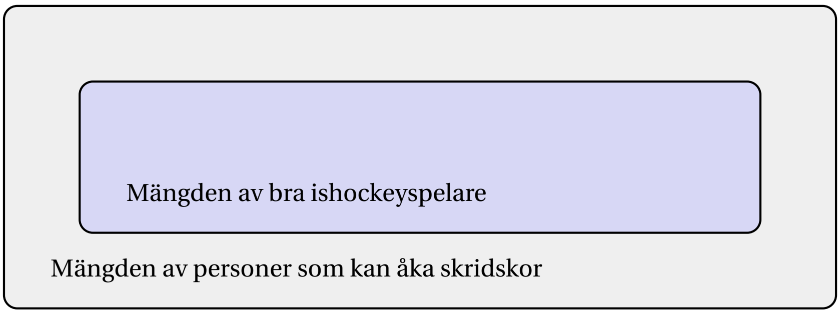
[5] Oxford English Dictionary anger följande ursprung:
derivative Late Middle English (in the adjective sense ‘having the power to draw off’, and in the noun sense ‘a word derived from another’): from French *dérivatif, -ive*, from Latin *derivativus*, from *derivare* (see derive).
derive Late Middle English (in the sense ‘draw a fluid through or into a channel’): from Old French *deriver* or Latin *derivare*, from *de-* ‘down, away’ + *rivus* ‘brook, stream’.

[Tillbaka till föreläsnings- och lektionsplan.]



Kontinuitet är alltså ett **nödvändigt** (men **inte tillräckligt**) villkor för deriverbarhet. Deriverbarhet medför kontinuitet, men inte tvärtom. Deriverbarhet är ett starkare krav att ställa på en funktion än kontinuitet.

Som en en liknelse kan man säga att för att bli en bra ishockeyspelare är det *nödvändigt* att kunna åka skridskor, men det är *inte tillräckligt* (man måste ju även kunna hantera klubba och puck, osv.):



- Derivatan av en sammansatt funktion ges av **kedjeregeln**,

$$\frac{d}{dx} f(g(x)) = \underbrace{f'(g(x))}_{\substack{\text{yttre}\\ \text{derivatan}}} \cdot \underbrace{g'(x)}_{\substack{\text{inre}\\ \text{derivatan}}}.$$

Alla brukar snabbt lära sig att hantera kedjeregeln i enkla fall som när man ska derivera ”sinus av något uttryck”, t.ex. sinus av $5x$ eller sinus av x^3:

$$\tfrac{d}{dx} \sin(5x) = \cos(5x) \cdot 5, \qquad \tfrac{d}{dx} \sin(x^3) = \cos(x^3) \cdot 3x^2.$$

Yttre derivatan fås ju av att ”derivatan av sinus är cosinus” ($\tfrac{d}{dx} \sin x = \cos x$, eller kortare uttryckt $\sin' = \cos$, dvs. om f är **sinus-funktionen** så är f' **cosinus-funktionen**), och inre derivatan blir vad den blir.

Men av någon anledning har många otroligt svårt att göra rätt i motsvarande situation när f är **arctan-** eller **arcsin-funktionen**. Förklara det, den som kan! En teori är att det möjligen beror på att derivatorna av dessa funktioner inte har några egna *namn* (som ”cosinus”) utan bara ges av formler:

$$\tfrac{d}{dx} \arctan x = \frac{1}{1+x^2}, \qquad \tfrac{d}{dx} \arcsin x = \frac{1}{\sqrt{1-x^2}}.$$

Men låt oss då *ge* ett namn åt arctan-funktionens derivata, varför inte (tillfälligt!) kalla den för ”smurf-funktionen”? Såhär:

$$\operatorname{smurf} x = \frac{1}{1+x^2}.$$

Då blir det precis som för sinus-funktionen ovan; yttre derivatan fås av att ”derivatan av arctan är smurf”, och inre derivatan blir vad den blir:

$$\tfrac{d}{dx} \arctan(5x) = \operatorname{smurf}(5x) \cdot 5, \qquad \tfrac{d}{dx} \arctan(x^3) = \operatorname{smurf}(x^3) \cdot 3x^2.$$

[Tillbaka till föreläsnings- och lektionsplan.]

Och om vi sedan skriver ut vad smurf-funktionen var för något, så framgår det förhoppningsvis tydligt vad som är det korrekta sättet att derivera ”arctan av något uttryck” med kedjeregeln:

$$\tfrac{d}{dx}\arctan(5x) = \frac{1}{1+(5x)^2}\cdot 5 = \frac{5}{1+25x^2}, \qquad \tfrac{d}{dx}\arctan(x^3) = \frac{1}{1+(x^3)^2}\cdot 3x^2 = \frac{3x^2}{1+x^6}.$$

Och på liknande sätt inser man att det blir såhär när man deriverar ”arcsin av något uttryck”:

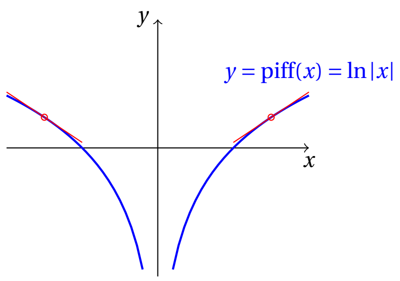
$$\tfrac{d}{dx}\arcsin(5x) = \frac{1}{\sqrt{1-(5x)^2}}\cdot 5 = \frac{5}{\sqrt{1-25x^2}}, \qquad \tfrac{d}{dx}\arcsin(x^3) = \frac{1}{\sqrt{1-(x^3)^2}}\cdot 3x^2 = \frac{3x^2}{\sqrt{1-x^6}}.$$

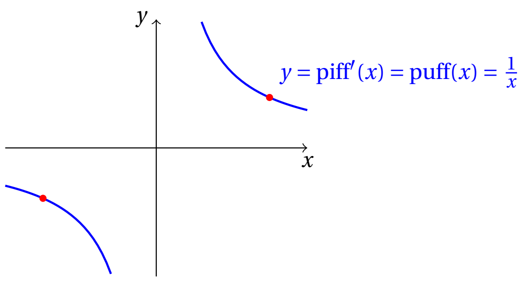
- En liknande sak som ibland också ställer till problem är derivering av ”logaritmen av absolutbeloppet av någon uttryck” med kedjeregeln. Även här kan det kanske hjälpa att införa tillfälliga namn för att förstå hur man ska hantera yttre derivatan. Ge funktionen ”logaritm sammansatt med absolutbelopp” beteckningen

$$\operatorname{piff}(x) = \ln|x|, \qquad x \neq 0,$$

och kalla dess derivata för

$$\operatorname{puff}(x) = \frac{1}{x}, \qquad x \neq 0.$$





När man deriverar med kedjeregeln får man alltså t.ex.

$$\tfrac{d}{dx}\ln\left|x^2+5\right| = \tfrac{d}{dx}\operatorname{piff}(x^2+5) = \operatorname{puff}(x^2+5)\cdot 2x = \frac{1}{x^2+5}\cdot 2x = \frac{2x}{x^2+5},$$

där yttre derivatan fås av att ”derivatan av piff är puff”, och inre derivatan blir vad den blir. Det allmänna fallet blir

$$\tfrac{d}{dx}\ln\left|f(x)\right| = \tfrac{d}{dx}\operatorname{piff}\big(f(x)\big) = \operatorname{puff}\big(f(x)\big)\cdot f'(x) = \frac{1}{f(x)}\cdot f'(x) = \frac{f'(x)}{f(x)}.$$

Observera här att det *inte* ska vara några absolutbeloppstecken i högerledet. Absolutbeloppet ingår som en del i ”piff-funktionen” som deriveras, inte i ”puff-funktionen” som fås som resultat av deriveringen.

- Bokens **bevis av kedjeregeln** (på s. 196) innehåller mycket formler och nästan inga förklaringar, så det är kanske inte helt lättbegripligt. Här kommer en utförligare version, för den som är intresserad av att verkligen förstå detaljerna.

  En första sak som nog borde förklaras är vad som är fel med följande resonemang:

  > ”Anta att $z = f(y)$ och $y = g(x)$, så att den sammansatta funktionen är $z = f(g(x))$. Om man gör en ändring $\Delta x \neq 0$ i x:s värde, så ger detta upphov till en ändring Δy i y:s värde, som i sin tur ger upphov till en ändring Δz i z:s värde. Om vi nu skriver
  >
  > $$\frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y}\cdot\frac{\Delta y}{\Delta x}$$

och låter Δx gå mot noll, så kommer även Δy att gå mot noll, och därmed får vi

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx},$$

vilket skulle bevisas.”

Problemet här är att steget där man förlänger med Δy bara är giltigt ifall $\Delta y \neq 0$. Om funktionen $y = g(x)$ har egenskapen att $\Delta y \neq 0$ för alla tillräckligt små $\Delta x \neq 0$ så fungerar alltså resonemanget, men det är inte säkert att det är så bara för att g är deriverbar i den punkt som det handlar om, för ifall derivatan $g' = dy/dx$ är noll i den punkten kan mycket väl Δy vara noll för Δx godtyckligt nära noll.

Nästa sak att förklara är idén bakom hur man lagar detta hål i resonemanget. Ifall $\Delta y = 0$ så är uppenbart även $\Delta z = 0$, så följande är sant för alla $\Delta x \neq 0$:

$$\frac{\Delta z}{\Delta x} = \begin{cases} \dfrac{\Delta z}{\Delta y} \cdot \dfrac{\Delta y}{\Delta x}, & \Delta y \neq 0, \\ 0, & \Delta y = 0. \end{cases}$$

(Eller, om man ska vara noga: det är åtminstone sant för alla *tillräckligt små* $\Delta x \neq 0$, så att man håller sig i någon omgivning av utgångspunkten x där funktionerna $y = g(x)$ och $z = f(g(x))$ är definierade.) Så vi behöver visa att *detta* uttryck har gränsvärdet $\frac{dz}{dy} \cdot \frac{dy}{dx}$ när $\Delta x \to 0$. För detta syfte, skriv om uttrycket såhär:

$$\begin{aligned} \frac{\Delta z}{\Delta x} &= \begin{cases} \dfrac{\Delta z}{\Delta y} \cdot \dfrac{\Delta y}{\Delta x}, & \Delta y \neq 0, \\ \dfrac{dz}{dy} \cdot 0, & \Delta y = 0 \end{cases} \\ &= \begin{cases} \dfrac{\Delta z}{\Delta y} \cdot \dfrac{\Delta y}{\Delta x}, & \Delta y \neq 0, \\ \dfrac{dz}{dy} \cdot \dfrac{\Delta y}{\Delta x}, & \Delta y = 0 \end{cases} \\ &= \frac{\Delta y}{\Delta x} \cdot \begin{cases} \dfrac{\Delta z}{\Delta y}, & \Delta y \neq 0, \\ \dfrac{dz}{dy}, & \Delta y = 0. \end{cases} \end{aligned}$$

För att se vad som händer här när $\Delta x \to 0$, och verkligen utföra beviset i detalj, låt oss uttrycka detta i termer av funktionerna f och g istället.

Sats (Kedjeregeln)**.** Om funktionen g är deriverbar i punkten a och funktionen f är deriverbar i punkten $b = g(a)$ så är den sammansatta funktionen $f \circ g$ deriverbar i punkten a, med derivatan

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a) \quad \big(= f'(b) \cdot g'(a)\big).$$

Bevis. Notera först att den sammansatta funktionen $f \circ g$ kommer att uppfylla krav D1 i derivatans definition, dvs. vara definierad i en omgivning av punkten a, på grund av att f och g enligt deriverbarhetsförutsättningen också uppfyller krav D1. Såhär:

> g är definierad i en omgivning A till punkten a, och f är definierad i en omgivning B till punkten $b = g(a)$, och eftersom g är kontinuerlig i a (enligt satsen att deriverbarhet medför kontinuitet) så har punkten a en omgivning $A' \subseteq A$ som avbildas av g in i B, och då är ju $f \circ g$ definierad i A'.

[Tillbaka till föreläsnings- och lektionsplan.]

(Detta är huvudanledningen till att krav D1 är med i derivatans definition. Om inte det vore med skulle det kunna hända att a inte är en hopningspunkt till definitionsmängden för $f \circ g$, och då vore det meningslöst att ens börja tala om gränsvärdet i definitionen av $(f \circ g)'(a)$.)

Åter nu till uträkningen ovan. Låt oss skriva h istället för Δx. När x-värdet ändras från $x = a$ till $x = a + h$ kommer y-värdet att ändras från $y = g(a)$ till $y = g(a+h)$, så ändringen Δy, som vi här kommer att beteckna med k istället, är $k = \Delta y = g(a+h) - g(a)$. Och när y ändras från $y = g(a) = b$ till $y = g(a+h) = b + k$ kommer z-värdet att ändras från $z = f(b)$ till $z = f(b+k)$, så att $\Delta z = f(b+k) - f(b) = f(g(a+h)) - f(g(a))$.

Det sista uttrycket för $\Delta z/\Delta x$ ovan (just före satsen) kan alltså skrivas såhär:

$$\frac{f(g(a+h)) - f(g(a))}{h} = \frac{g(a+h) - g(a)}{h} \cdot \begin{cases} \dfrac{f(b+k) - f(b)}{k}, & k \neq 0, \\ f'(b), & k = 0, \end{cases}$$

där k underförstått beror på h enligt formeln $k = g(a+h) - g(a)$.

Låt oss ge ”klammerfunktionen” i högerledet beteckningen $\Psi(k)$, alltså

$$\Psi(k) = \begin{cases} \dfrac{f(b+k) - f(b)}{k}, & k \neq 0, \\ f'(b), & k = 0. \end{cases}$$

Notera att $\frac{f(b+k)-f(b)}{k} \to f'(b)$ då $k \to 0$, eftersom f enligt förutsättning är deriverbar i b. Med andra ord: $\Psi(k) \to \Psi(0)$ då $k \to 0$, dvs. funktionen $\Psi(k)$ är kontinuerlig i punkten $k = 0$.

Med denna notation kommer den andra faktorn i uttrycket ovan att vara $\Psi(k) = \Psi\big(g(a+h) - g(a)\big)$, och hela formeln blir

$$\frac{(f \circ g)(a+h) - (f \circ g)(a)}{h} = \frac{g(a+h) - g(a)}{h} \cdot \Psi\big(g(a+h) - g(a)\big).$$

I vänsterledet står här det uttryck vars gränsvärde då $h \to 0$ (om det existerar ändligt) ger derivatan $(f \circ g)'(a)$. I högerledet står en produkt av två faktorer. Den första faktorn går mot $g'(a)$ då $h \to 0$, eftersom g enligt förutsättning är deriverbar i a. Vad går den andra faktorn mot? Deriverbarhetsförutsättningen medför att g är kontinuerlig i a, så $g(a+h) \to g(a)$ då $h \to 0$. Detta, tillsammans med att Ψ är kontinuerlig i 0, gör att

$$\lim_{h \to 0} \Psi\big(\underbrace{g(a+h) - g(a)}_{\to 0}\big) = \Psi(0) = f'(b).$$

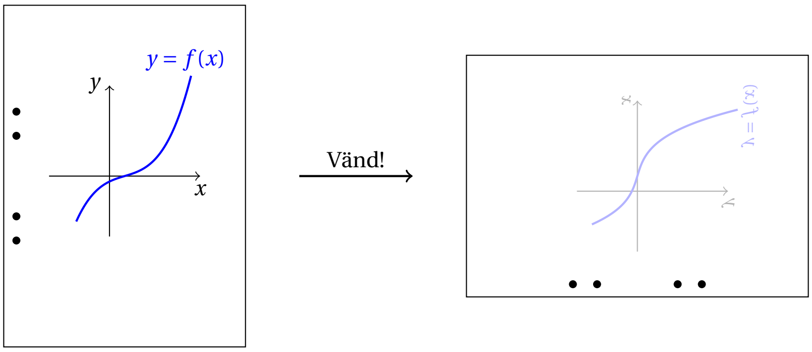
Alltså $(f \circ g)'(a) = g'(a) \cdot f'(b)$, vilket skulle visas. $\square$

Man skulle kunna sammanfatta hela idén med detta bevis i en enda mening: ”tolka den problematiska faktorn $\Delta z/\Delta y$ som dz/dy när $\Delta y = 0$”.

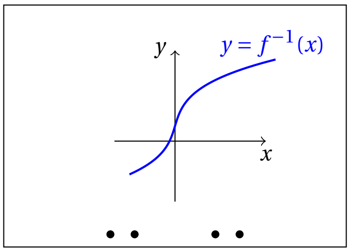
- En påminnelse från grundkursen: grafen för f^{-1} fås genom att reflektera grafen för f i linjen $y = x$ (förutsatt att f är inverterbar till att börja med, annars blir ju den reflekterade kurvan ingen graf).

  Ett mera handfast sätt att beskriva detta är att man utgår från grafen $y = f(x)$ och sedan helt enkelt vänder på papperet och tittar bakifrån, genom papperet, på ett sådant sätt att axlarna ser ut att byta roller (dvs. x-axeln pekar uppåt och y-axeln åt höger):

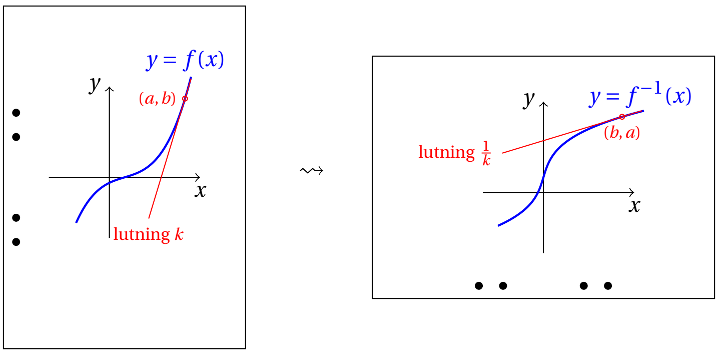
[Tillbaka till föreläsnings- och lektionsplan.]



I denna omvända figur ser man nu grafen $x = f^{-1}(y)$. Ifall man vill att det ska vara mer som vanligt, med x-axeln pekande åt höger och y-axeln uppåt, så kan man ju fylla i figuren ordentligt på papperets baksida (eller rita av den någon annanstans), men skriva dit de vanliga namnen på axlarna istället. Den graf man då ser är $y = f^{-1}(x)$, eftersom man har låtit x och y byta roller:



Och om den ursprungliga kurvan har lutningen k i en viss punkt (a, b) så kommer den omvända kurvan att ha lutningen $1/k$ i den motsvarande punkten (b, a), eller hur?



T.ex. en ursprunglig lutning $k = \frac{5}{2}$ motsvarar ju att man går 2 steg höger och 5 steg uppåt, men när man vänder på papperet så kastar man om ”höger” och ”uppåt”, så att det blir 2 steg uppåt och 5 steg höger istället, alltså lutning $\frac{2}{5} = 1/\frac{5}{2} = 1/k$.

- Ovanstående figur illustrerar vad Sats 4.6 handlar om (derivata av invers funktion, ”$\frac{dx}{dy} = 1/\frac{dy}{dx}$”).

  Om man ska vara petig är dock denna sats inte helt korrekt formulerad i boken. I definitionen av vad som menas med att f^{-1} är deriverbar i b ingår att f^{-1} ska vara definierad i en omgivning av b, och detta följer faktiskt inte med automatik från de förutsättningar som anges i boken, utan måste läggas till som en separat förutsättning:

  Sats 4.6 (korrigerad version)**.** Ifall f är en inverterbar funktion, som är deriverbar i en punkt a och har derivatan $f'(a) \neq 0$, så är den inversa funktionen f^{-1} deriverbar i motsvarande punkt $b = f(a)$ och har där derivatan

  $$(f^{-1})'(b) = \frac{1}{f'(a)}$$

  förutsatt att f^{-1} är kontinuerlig i b och definierad i en omgivning av b.

[Tillbaka till föreläsnings- och lektionsplan.]

Följande följdsats till Sats 4.6 är enklare att tillämpa, och fullt tillräckligt för denna kurs:

Sats. Anta att f är definierad på ett öppet intervall I, och deriverbar på I med **positiv derivata i varje punkt** $a \in I$. Då är f inverterbar, med f^{-1} definierad på ett öppet intervall J, och **inversen är deriverbar i varje punkt** $b \in J$, med

$$\left(f^{-1}\right)'(b) = \frac{1}{f'(a)}, \qquad \text{där } b = f(a), \text{ dvs. } a = f^{-1}(b).$$

(Samma sak gäller om f har **negativ derivata i varje punkt** $a \in I$.)

Bevis. Deriverbarhet medför kontinuitet, så f är kontinuerlig på I, och eftersom $f' > 0$ eller $f' < 0$ på I så är f antingen strängt växande eller strängt avtagande på I (enligt sats som kommer på Föreläsning 7). Enligt Sats 3.5 är f därmed inverterbar, med en kontinuerlig invers f^{-1} som är definierad på ett motsvarande öppet intervall J. I detta läge är förutsättningarna för Sats 4.6 (inklusive den extra förutsättningen ovan) uppfyllda för varje $a \in I$, och slutsatsen följer. □

Slutsatsen kan även formuleras såhär:

$$\left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))}, \qquad x \in J.$$

- **Exempel.** Funktionen

$$f(x) = \tan x, \qquad -\frac{\pi}{2} < x < \frac{\pi}{2},$$

dvs. restriktionen av tangens-funktionen till intervallet $I = \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, uppfyller ovanstående förutsättningar, eftersom $f'(x) = 1 + \tan^2 x \geq 1 > 0$ för alla $x \in I$. Den inversa funktionen är $f^{-1}(x) = \arctan x$, definierad på intervallet $J = \mathbf{R}$. Alltså är

$$\tfrac{d}{dx} \arctan x = \left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))} = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}, \qquad x \in \mathbf{R}.$$

- **Exempel.** Funktionen

$$f(x) = \sin x, \qquad -\frac{\pi}{2} < x < \frac{\pi}{2},$$

dvs. restriktionen av sinus-funktionen till intervallet $I = \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, uppfyller ovanstående förutsättningar, eftersom $f'(x) = \cos x > 0$ för alla $x \in I$. Den inversa funktionen är

$$f^{-1}(x) = \arcsin x, \qquad -1 < x < 1,$$

dvs. restriktionen av arcsin-funktionen till intervallet $J =]-1, 1[$. Alltså är

$$\tfrac{d}{dx} \arcsin x = \left(f^{-1}\right)'(x) = \frac{1}{f'(f^{-1}(x))} = \frac{1}{\cos(\arcsin x)} \overset{*}{=} \frac{1}{\sqrt{1 - \sin^2(\arcsin x)}} = \frac{1}{\sqrt{1 - x^2}}, \qquad -1 < x < 1.$$

Här har vi i det stjärnmarkerade steget använt trigonometriska ettan $\cos^2 t + \sin^2 t = 1$ för att lösa ut $\cos t = \pm\sqrt{1 - \sin^2 t}$, och eftersom vi i detta fall har $t = \arcsin x \in I$ så är $\cos t > 0$, vilket visar att det är plustecknet som är det rätta.

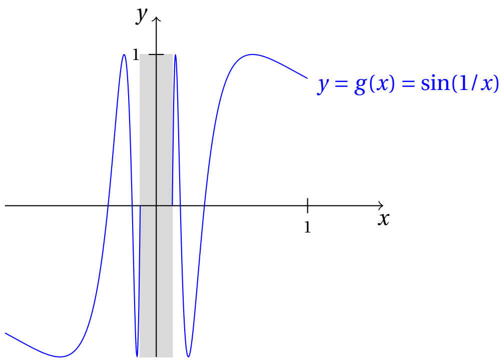
(Funktionen $\arcsin x$ är som bekant definierad och kontinuerlig på det slutna intervallet $-1 \leq x \leq 1$, men den är bara *deriverbar* på det öppna intervallet $-1 < x < 1$.)

- En subtil detalj är att även om f är deriverbar överallt så behöver inte derivatan f' vara en kontinuerlig funktion. Standardmotexemplet är

$$f(x) = \begin{cases} x^2 \sin(1/x), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

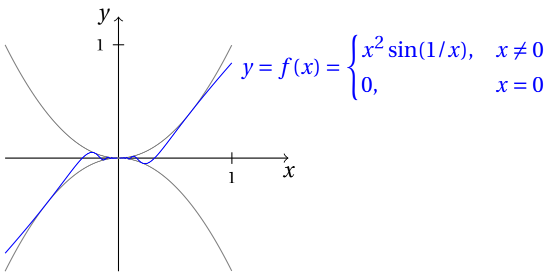
För att förstå hur denna funktion ser ut, betrakta först funktionen $g(x) = \sin(1/x)$, $x \neq 0$:

[Tillbaka till föreläsnings- och lektionsplan.]



Vad som händer i det skuggade området är svårt att åskådliggöra grafiskt, så du får använda din fantasi lite! Kurvan $y = g(x)$ kommer att oscillera mellan $y = -1$ och $y = 1$ oändligt många gånger; den svänger fortare och fortare ju närmare $x = 0$ den kommer. Funktionens nollställen ges av $1/x = n\pi$, dvs. $x = 1/(n\pi)$ (för $n \in \mathbf{Z}$, $n \neq 0$). Eftersom det finns punkter godtyckligt nära 0 där $g(x) = 1$, nämligen $x = 1/(\pi/2 + 2\pi n)$ (för $n \in \mathbf{Z}$), och andra punkter godtyckligt nära 0 där $g(x) = -1$, nämligen $x = 1/(-\pi/2 + 2\pi n)$ (för $n \in \mathbf{Z}$), så saknar $g(x)$ gränsvärde då $x \to 0$.

Vår funktion $f(x)$ har bildats genom att multiplicera denna funktion $g(x) = \sin(1/x)$ med en amplitudfaktor x^2, vilket ger en kurva $y = f(x)$ som oscillerar mellan parablerna $y = -x^2$ och $y = x^2$ istället för mellan linjerna $y = -1$ och $y = 1$, och så har man fyllt i värdet $f(0) = 0$ så att f blir kontinuerlig i 0 (instängningsregeln ger ju att $f(x) \to 0$ då $x \to 0$):



Även här är det svårt att se i grafen vad som händer nära $x = 0$, men vi ju kan ju *räkna* ut det. För $x \neq 0$ ges derivatan av

$$f'(x) = \tfrac{d}{dx}\Big(x^2 \sin\big(\tfrac{1}{x}\big)\Big) = 2x \sin\big(\tfrac{1}{x}\big) + x^2 \cos\big(\tfrac{1}{x}\big)\big(\tfrac{-1}{x^2}\big) = \underbrace{\underbrace{2x}_{\to 0}\ \underbrace{\sin\big(\tfrac{1}{x}\big)}_{\text{begr.}}}_{\to 0} - \underbrace{\cos\big(\tfrac{1}{x}\big)}_{\text{osc.}}, \qquad x \neq 0.$$

Den första termen går mot noll då $x \to 0$, medan den andra termen $\cos(1/x)$ saknar gränsvärde; den oscillerar ju mellan -1 och 1 ungefär som $g(x)$ ovan. Geometriskt betyder detta att kurvan $y = f(x)$ växlar frenetiskt mellan att luta ungefär 45 grader uppåt och ungefär 45 grader neråt. Alltså har inte $f'(x)$ något gränsvärde då $x \to 0$.

Å andra sidan så *har* faktiskt derivatan ett värde i punkten 0, nämligen $f'(0) = 0$, vilket geometriskt sett beror på att kurvan $y = f(x)$ är instängd mellan parablerna $y = \pm x^2$ som båda har lutning noll i origo. Låt oss verifiera detta direkt med derivatans definition också:

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = 0, \qquad \text{eftersom} \quad \frac{f(h) - f(0)}{h} = \frac{h^2 \sin(1/h) - 0}{h} = \underbrace{h}_{\to 0}\ \underbrace{\sin\big(\tfrac{1}{h}\big)}_{\text{begr.}} \to 0.$$

Sammanfattningsvis är derivatan

$$f'(x) = \begin{cases} 2x \sin(1/x) - \cos(1/x), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

alltså **definierad** i varje punkt $x \in \mathbf{R}$, men den är **diskontinuerlig** i punkten 0.

- Som en liten varning för ett vanligt tankefel kommer här en variation av föregående exempel, där det har lagts till en extra term x:

$$u(x) = \begin{cases} x^2 \sin(1/x) + x, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

  Vad är derivatan $u'(x)$?

  Liksom ovan kan vi beräkna $u'(x)$ för $x \neq 0$ genom att helt enkelt derivera uttrycket $x^2 \sin(1/x) + x$ med vanliga deriveringsregler. Derivatan i en punkt $x \neq 0$ beror ju bara på u:s värden i de närliggande punkterna, som alla också är nollskilda, så ”specialvärdet” $u(0) = 0$ har ingen inverkan där.

  Det som däremot många gör fel på är vad $u'(0)$ blir. Det felaktiga resonemanget går såhär:

  > ”$u(0) = 0$, och derivatan av konstanten noll är noll, så $u'(0) = 0$.”

  Men detta är helt vansinnigt, om man tänker efter lite! Från funktionsvärdet *i en enda punkt*, dvs. från *en enda prick på kurvan*, kan man ju omöjligt avgöra hur kurvan $y = u(x)$ *lutar*, eller hur? För att avgöra lutningen i punkten $x = 0$ måste man ta hänsyn *både* till själva funktionsvärdet $u(0) = 0$ och värdena $u(x)$ för $x \neq 0$.

  Vad som är sant är att derivatan av en **konstant funktion** är noll, men det är nonsens att säga att derivatan av ett enda ”konstant” funktionsvärde är noll. Alla funktionsvärden är ju konstanter (tal), så om det resonemanget vore giltigt skulle *alla* derivator vara identiskt noll: $\frac{d}{dx}x^2 = 0$, $\frac{d}{dx}\sin x = 0$, $\frac{d}{dx}e^x = 0$, och så vidare…

  Derivatans definition ger däremot det rätta svaret:

$$u'(0) = \lim_{h \to 0} \frac{u(h) - u(0)}{h} = 1,$$

  eftersom

$$\frac{u(h) - u(0)}{h} = \frac{h^2 \sin(1/h) + h - 0}{h} = \underbrace{h}_{\to 0} \underbrace{\sin\left(\tfrac{1}{h}\right)}_{\text{begr.}} + 1 \to 0 + 1 = 1 \quad \text{då } h \to 0.$$

  Svar, alltså:

$$u'(x) = \begin{cases} 2x \sin(1/x) - \cos(1/x) + 1, & x \neq 0, \\ 1, & x = 0. \end{cases}$$

  Och detta stämmer även med räknelagarna för derivata. Jämfört med det föregående exemplet har vi ju

$$u(x) = f(x) + x$$

  för alla $x \in \mathbf{R}$ (även för $x = 0$, eller hur?), vilket medför att

$$u'(x) = f'(x) + 1$$

  för alla x (inklusive $x = 0$). Och vi hade ju tidigare räknat ut att $f'(0) = 0$, så det måste bli $u'(0) = f'(0) + 1 = 0 + 1 = 1$.

- De föregående exemplen illustrerar att man måste vara lite försiktig vid beräkning av derivatan av en styckvis definierad funktion; i ”undantagspunkter” eller ”skarvningspunkter” kan man inte använda vanliga räkneregler utan vidare, utan man behöver man tänka efter vad derivatans *definition* säger egentligen.

  Här är en annan sådan situation, där många får rätt svar, men med felaktigt resonemang:

[Tillbaka till föreläsnings- och lektionsplan.]

Låt u och v vara deriverbara (och därmed även kontinuerliga) funktioner, säg på hela $\mathbf{R}$ för enkelhets skull, och betrakta den styckvis definierade funktionen

$$f(x) = \begin{cases} u(x), & x \le a, \\ v(x), & x > a. \end{cases}$$

Då kan man direkt säga att

$$f'(x) = \begin{cases} u'(x), & x < a, \\ \boxed{\quad ? \quad}, & x = a, \\ v'(x), & x > a, \end{cases}$$

där det enda som behöver undersökas ytterligare är $f'(a)$. Kravet för att $f'(a)$ ska existera är att vänsterderivatan $f'_-(a)$ och högerderivatan $f'_+(a)$ båda existerar, och att de är lika.

Det felaktiga resonemanget som man ofta ser lyder såhär:

> ”Vänster- och högerderivatan ges av $f'_-(a) = \lim_{x \to a^-} u'(x)$ resp. $f'_+(a) = \lim_{x \to a^+} v'(x)$.”

Problemet med detta är att derivatorna u' och v' *i princip* inte behöver vara kontinuerliga funktioner (även om de är det i de flesta enkla fall som förekommer i de övningsuppgifter som ni kommer att stöta på), och dessutom har man inte tagit hänsyn till om f är kontinuerlig i a.

Här är det korrekta resonemanget:

> Direkt ur definitionerna följer att
>
> $$f'_-(a) = \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a} = \Big[f(x) = u(x) \text{ för } x \le a \Big] = \lim_{x \to a^-} \frac{u(x) - u(a)}{x - a} = u'_-(a) = u'(a).$$
>
> **Om** dessutom $u(a) = v(a)$, så att f är **kontinuerlig** i a, så är
>
> $$\begin{aligned} f'_+(a) &= \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} = \Big[f(x) = v(x) \text{ för } x > a, \text{ men } f(a) = u(a) \Big] \\ &= \lim_{x \to a^+} \frac{v(x) - u(a)}{x - a} = \Big[\text{om } u(a) = v(a) \Big] = \lim_{x \to a^+} \frac{v(x) - v(a)}{x - a} = v'_+(a) = v'(a). \end{aligned}$$
>
> (Men observera att ifall $v(a) \ne u(a)$ så existerar **inte** högerderivatan $f'_+(a)$, och i synnerhet är det inte sant att den är lika med $v'(a)$, och inte är den lika med $\lim_{x \to a^+} v'(x)$ heller.)

Man måste alltså först undersöka om f är kontinuerlig i a (annars kan ju omöjligt $f'(a)$ existera!), och i så fall räknar man ut derivatorna $u'(a)$ och $v'(a)$ i punkten a och ser om de är lika.

Om $u(x)$ och $v(x)$ ges av snälla formler så kan derivatorna $u'(x)$ och $v'(x)$ beräknas med vanliga räkneregler, varefter $u'(a)$ och $v'(a)$ fås genom att **sätta in** $x = a$, men det är alltså här principiellt fel att titta på **gränsvärdena** av $u'(x)$ och $v'(x)$ då $x \to a$, även om detta ofta ger samma svar.

(Man kan såklart även alltid gå direkt på definitionen av höger- och vänsterderivata och försöka beräkna de gränsvärden som då uppkommer, men en nackdel med detta är att man då inte har lika direkt tillgång till räkneregler för derivata, som produktregeln och kedjeregeln.)

Lektion 5

- Definitionen av derivata: P4.6.
- Beräkning av derivator med hjälp av deriveringsregler: P4.7, P4.8, P4.9.
- Hur man ser var det ska vara minustecken i derivatorna av sinus och cosinus: P4.1.
- Kontinuitet och deriverbarhet: P4.10.

[Tillbaka till föreläsnings- och lektionsplan.]

- Tangentlinje och normallinje till en kurva: P4.12.

  Kom ihåg att linjerna $y = k_1 x + m_1$ och $y = k_2 x + m_2$ är vinkelräta om och endast $k_1 k_2 = -1$.

  Detta är inte något särskilt djupsinnigt påstående; om du t.ex. på rutat papper ritar en linje med lutning $k_1 = 5$ (en ruta höger ger fem rutor uppåt) och sedan **vrider papperet 90 grader och räknar rutor** så ser du ju själv att du får en linje med lutning $k_2 = -1/5$ (fem rutor höger ger en ruta neråt).

- Kedjeregeln: P4.13.

  Ledning: Arean A beror på radien r, som i sin tur beror på tiden t. Enligt kedjeregeln är $\frac{dA}{dt} = \frac{dA}{dr}\frac{dr}{dt}$.

- Beräkning av derivata direkt med definitionen: B4.2ab, B4.6a.

- Derivatan av x^x: B4.10c.

  Ledning: Vad gjorde vi vid gränsvärdesuträkningar när variabeln x fanns i både basen och exponenten? Gör samma omskrivning här.

- Mera kedjeregeln: P4.19.

  Ledning: Skuggans längd L beror på solens vinkel φ över horisonten, som i sin tur beror på tiden t. Då blir $\frac{dL}{dt}$ lika med vadå, enligt kedjeregeln?

  Och kom ihåg att vinkeln måste mätas i **radianer** när man sysslar med derivator av trigonometriska funktioner, annars blir det fel!

- Höger-/vänsterderivata, styckvis definierade funktioner: P4.14, P4.15. P4.16*.

  Se kommentarerna ovan för tips och varningar!

- Beräkning av gränsvärden genom att känna igen derivatans definition och använda kända derivator: P4.11*.

Föreläsning 7. Vad derivatan säger om funktionen

Innehåll

- Avsnitt 4.4, lite från avsnitt 4.5.

- Med hjälp av medelvärdessatsen för derivator bevisar vi följande grundläggande sats[6]:

$$\text{Om } \left\{ \begin{array}{c} f' > 0 \\ f' \geq 0 \\ f' = 0 \\ f' \leq 0 \\ f' < 0 \end{array} \right\} \text{ på ett öppet intervall så är } f \left\{ \begin{array}{c} \text{strängt växande} \\ \text{växande} \\ \text{konstant} \\ \text{avtagande} \\ \text{strängt avtagande} \end{array} \right\} \text{ på det intervallet.}$$

  Motsvarande gäller även för slutna intervall, ifall f är kontinuerlig på hela intervallet och deriverbar (med $f' > 0$ etc.) i det inre av intervallet.

- Inledande exempel på funktionsundersökning med hjälp av derivata.

  En **checklista för funktionsundersökning** återfinns i menyn på kurshemsidan.

  [Direktlänk till PDF-filen.]

[6] Med skrivsättet "$f' > 0$ på intervallet I" menas att $f'(x) > 0$ för varje $x \in I$.

[Tillbaka till föreläsnings- och lektionsplan.]

Kommentarer

- Kom ihåg *definitionen* av ”strängt växande”:

  Funktionen $f(x)$ sägs vara strängt växande på mängden $M \subseteq \mathbf{R}$ ifall olikheten
  $$f(x_1) < f(x_2)$$
  gäller för alla $x_1 \in M$ och $x_2 \in M$ sådana att $x_1 < x_2$.

  Här sägs inte ett knyst om derivata! Det handlar bara om att *jämföra* funktionsvärden i *olika* punkter i mängden M (som oftast är ett intervall, men det måste inte vara det).

  T.ex. är funktionen $f(x) = x^3$ strängt växande på mängden $\mathbf{R}$ (som är ett intervall). Detta har somliga lite svårt att förlika med faktumet att derivatan $f'(x) = 3x^2$ är noll i punkten $x = 0$. ”Hur kan funktionen vara *strängt* växande i en punkt där derivatan är noll?” Tankefelet här ligger i själva formuleringen ”strängt växande i en punkt” – detta är inte något begrepp som vi ens har definierat, och så länge man inte har angett en tydlig definition av vad man ska *mena* med detta så är det en helt meningslös utsaga! Vi har bara definierat ”strängt växande på en mängd”, vilket som sagt handlar om att jämföra värden i *olika* punkter. Och tar man två olika reella tal $x_1 < x_2$, vilka som helst, så är det otvetydigt sant att $x_1^3 < x_2^3$. Slut på diskussionen.

  Exemplet $f(x) = x^3$ visar också att satsen ovan *inte* är en ekvivalens, dvs. den säger *inte* att f är strängt växande på intervallet *om och endast om* $f' > 0$. Det står bara ”om”, inte ”endast om”.

- Likaså nämner inte definitionen av lokala extrempunkter (Def. 4.3 i boken) någonting om derivata, utan det handlar också enbart om att jämföra funktionsvärden i olika punkter.

  Det enkla exemplet $f(x) = |x|$ visar att en funktion mycket väl kan ha en lokal extrempunkt ($x = 0$ i detta fall) där derivatan inte är noll. Men om derivatan *existerar* i en (inre) lokal extrempunkt så måste den vara noll där; detta är Sats 4.9.

- Deriverbarhet i en enda punkt säger inte så mycket. T.ex är följande funktion deriverbar i punkten $x = 0$ (enbart), med $f'(0) = 0$, men den är inte ens kontinuerlig i övriga punkter:
  $$f(x) = \begin{cases} x^2, & x \in \mathbf{Q}, \\ 0, & x \in \mathbf{R} \setminus \mathbf{Q}. \end{cases}$$
  Och även om f' existerar överallt, så säger det inte så mycket att bara veta att $f' > 0$ i en viss punkt. T.ex. (jämför med de liknande exemplen några sidor tillbaka) är funktionen
  $$f(x) = \begin{cases} x^2 \sin(1/x) + x/2, & x \neq 0, \\ 0, & x = 0, \end{cases}$$
  deriverbar överallt, med derivatan
  $$f'(x) = \begin{cases} 2x \sin(1/x) - \cos(1/x) + 1/2, & x \neq 0, \\ 1/2, & x = 0, \end{cases}$$
  så $f'(0) = 1/2 > 0$, dvs. derivatan i origo är positiv, men f är trots detta inte växande på något intervall kring origo, säg $]-\delta, \delta[$ med $\delta > 0$, ty varje sådant intervall innhåller delintervall där $f'(x) < 0$.

  Det är därför man måste anta saker om f' på ett helt *intervall* för att kunna dra några intressanta slutsatser.

- För den teoretiskt lagde som funderar över varför man gör som man gör i beviset av satsen ”derivatan positiv medför funktionen strängt växande” kan jag rekommendera en intressant liten text på nätet av den kände Cambridgematematikern Timothy Gowers, som inte bara är en framstående forskare (Fieldsmedaljör 1998) utan även en utmärkt pedagog: **What is the point of the mean value theorem?**

[Tillbaka till föreläsnings- och lektionsplan.]

Lektion 6

- Funktionsundersökning: P4.21, P4.22.
- Derivata av invers funktion: P4.27, P4.18.
- Några begreppsfrågor: P4.2, P4.3, P4.30, P4.31.
- Medelvärdessatsen: P4.29, B4.27, P4.5.
- Hantering av absolutbelopp: P4.23.
- Ett par exempel med $x^2 \sin(1/x)$: B4.26*.

Föreläsning 8. Användning av derivata

Innehåll

- Avsnitt 4.5 och 4.6, förutom materialet om sneda asymptoter och konvexa/konkava funktioner.
- Derivator av högre ordning.
- Fler exempel på funktionsundersökning.

Kommentarer

- Avsnittet om konvexa/konkava funktioner betraktas som överkurs i denna kurs, men läs det gärna ändå, för det är viktigt i många tillämpningar (t.ex. optimering).

  Några minnesregler: f är kon**vex** om f' är **väx**ande, f är konk**av** om f' är **av**tagande, funktionen e^x är konv**ex**.

- Varför sitter tvåorna som de gör i notationen $\dfrac{d^2 f}{dx^2}$ för andraderivata?

  Förklaringen är att skrivsättet ”$\frac{d}{dx}$” har kommit att betyda ”derivera det som står omedelbart till höger, med avseende på x”, efter att f:et så att säga halkat ner på sidan i uttrycket

  $$\frac{df}{dx} = \frac{d}{dx} f.$$

  Derivatan av derivatan av f skrivs då

  $$\frac{d}{dx}\frac{d}{dx} f.$$

  Om man här använder samma skrivsätt som om det vore en multiplikation,

  $$\left(\frac{d}{dx}\right)^2 f,$$

  och sedan (som om det vore en division) skriver

  $$\frac{d^2}{(dx)^2} f,$$

  samt tar bort parenteserna runt dx (dvs. underförstått betraktar ”dx” som en enda sammansatt symbol) och låter f:et hoppa upp på bråkstrecket igen, så blir det ju

  $$\frac{d^2 f}{dx^2}.$$

[Tillbaka till föreläsnings- och lektionsplan.]

Lektion 7

- Enkel rimlighetskontroll: P4.4.
- Funktionsundersökning: P4.24 (rita grafen $y = f(x)$ först, så kan du bara läsa av alla svaren från den sedan), P4.25, P4.26 (tänk noga efter vad *definitionsmängden* är för den funktion du undersöker!).
- Mera funktionsundersökning: P4.32ac, B4.28a, P4.39, P4.40, P4.42, B4.34, P4.28*, P4.35*, P4.43*.

Lektion 8

- Högre ordningens derivator: B4.41a, B4.43b*, P4.20*, B4.44*
- Ännu mera funktionsundersökning: P4.44abc, P4.46, P4.48, P4.49, P4.34*, B4.40.
- En kär gammal sats igen: P4.47.
- (Konvexitet: B4.47*, B4.48*.)

Föreläsning 9. Primitiv funktion (obestämd integral)

Innehåll

- Avsnitt 5.1 och 5.2.
- F kallas för en primitiv funktion till f om $F' = f$ (dvs. det är motsatsen till derivata).

  Ofta säger man bara ”en primitiv” istället för ”en primitiv funktion”.
- Standardprimitiver.
- Partiell integration (produktregeln baklänges).
- Variabelbyte (kedjeregeln baklänges).

Kommentarer

- När man har deriverat en funktion kan man fråga sig ifall det med kännedom enbart om derivatan går att tänka ut vilken den **ursprungliga** funktionen var (den som man deriverade).

  Det är därför som det heter **primitiv funktion**, för den ursprungliga betydelsen av ordet **primitiv** är ju just **ursprunglig**.
- Det är väl relativt lätt att komma ihåg standardprimitiven

  $$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C,$$

  men den här brukar vara värre:

  $$\boxed{\int \frac{dx}{\sqrt{1+x^2}} = \ln\left(x + \sqrt{1+x^2}\right) + C}.$$

  Var i all världen kommer det konstiga uttrycket i högerledet ifrån, och vad gör man om man sitter på en öde ö (eller på en analystenta) och har glömt bort hur det såg ut?

  Den korta versionen är att högerledet är funktionen **area sinus hyperbolicus**, vilket är **inversen till sinus hyperbolicus**:

  $$\boxed{\int \frac{dx}{\sqrt{1+x^2}} = \operatorname{arsinh} x + C}.$$

Den långa versionen lyder som följer. Börja med att komma ihåg **Eulers formler** för **cosinus** och **sinus**:

$$\cos t = \frac{e^{it}+e^{-it}}{2}, \qquad \sin t = \frac{e^{it}-e^{-it}}{2i}.$$

Dessa funktioner uppfyller ju **trigonometriska ettan** $\cos^2 t + \sin^2 t = 1$, vilket såklart är nära kopplat till enhetscirkelns ekvation $x^2+y^2=1$. Jag påminner om att trig-ettan spelade en viktig roll i härledningen av **derivatan av arcus sinus**, där ju formeln

$$\tfrac{d}{dt}\sin t = \cos t = \pm\sqrt{1-\sin^2 t} = \sqrt{1-\sin^2 t} \quad (\text{för } |t| \le \pi/2)$$

var upphovet till rotuttrycket i $\frac{d}{dx}\arcsin x = 1/\sqrt{1-x^2}$.

Om man tar bort alla ”i” ur Eulers formler får man definitionerna av två andra funktioner som heter **cosinus hyperbolicus** och **sinus hyperbolicus**:

$$\cosh t = \frac{e^{t}+e^{-t}}{2}, \qquad \sinh t = \frac{e^{t}-e^{-t}}{2}.$$

Som du lätt kan verifiera själv uppfyller dessa funktioner den **hyperboliska ettan** $\cosh^2 t - \sin^2 t = 1$, vars namn kommer av att $x^2-y^2=1$ är ekvationen för en hyperbel. Funktionen $\sinh t$ är strängt växande (eftersom dess derivata är $\cosh t > 0$), och dess värdemängd är $\mathbf{R}$ (eftersom $\sinh t \to \pm\infty$ då $t \to \pm\infty$), så den **har en invers** med definitionsmängd $\mathbf{R}$, som går under namnet **area sinus hyperbolicus**[7]:

$$\operatorname{arsinh} = \sinh^{-1}.$$

Ur hyperboliska ettan får man

$$\tfrac{d}{dt}\sinh t = \cosh t = \pm\sqrt{1+\sinh^2 t} = \sqrt{1+\sinh^2 t} \quad (\text{eftersom } \cosh t > 0),$$

vilket innebär att **derivatan av inversen till sinus hyperbolicus** är

$$\tfrac{d}{dx}\operatorname{arsinh} x = \frac{1}{\sqrt{1+x^2}}.$$

Detta visar att $\int \frac{dx}{\sqrt{1+x^2}} = \operatorname{arsinh} x + C$, så nu är bara frågan om $\operatorname{arsinh} x$ går att uttrycka på ett annat sätt. Svaret är ja, och detta är en enkel grunk-övning:

$$\begin{aligned}
\boxed{y = \operatorname{arsinh} x} &\iff x = \sinh y \iff x = \frac{e^y - e^{-y}}{2} \iff 2x = e^y - e^{-y} \\
&\iff (e^y)^2 - 2x\,e^y - 1 = 0 \iff (e^y - x)^2 = 1 + x^2 \iff e^y - x = \pm\sqrt{1+x^2} \\
&\iff e^y = x \pm \sqrt{1+x^2} \overset{*}{\iff} e^y = x + \sqrt{1+x^2} \iff \boxed{y = \ln\left(x + \sqrt{1+x^2}\right)},
\end{aligned}$$

där det stjärnmarkerade steget består i att förkasta det omöjliga fallet $e^y = x - \sqrt{1+x^2}$ eftersom $e^y > 0$ och $x - \sqrt{1+x^2} < 0$.

- För att beräkna primitiv funktion till ett uttryck av formen ”polynom gånger $\cos(kx)$ eller $\sin(kx)$ eller e^{kx}” används **upprepad partiell integration**. I varje steg integrerar man trig/exp-faktorn och deriverar polynom-faktorn, vilket ger en ny liknande integral med ett polynom vars gradtal är ett snäpp lägre, och så fortsätter man tills gradtalet är noll. Och det gäller att se upp en aning, så att man inte klantar sig med parenteser och ”minus-minus”.

[7] Se t.ex. **Wikipedia: Inverse hyperbolic functions** för en förklaring till varför det heter ”area”, till skillnad från de inversa trigonometriska funktionerna som ju heter ”arcus” = båge = båglängd = vinkel mätt i radianer.

[Tillbaka till föreläsnings- och lektionsplan.]

Såhär kan det se ut, exempelvis:

$$
\begin{aligned}
\boxed{\int e^{-5x}(x^3+4x^2+2)\,dx} &= \frac{e^{-5x}}{-5}(x^3+4x^2+2) - \int \frac{e^{-5x}}{-5}(3x^2+8x)\,dx \\
&= \frac{e^{-5x}}{-5}(x^3+4x^2+2) - \left(\frac{e^{-5x}}{25}(3x^2+8x) - \int \frac{e^{-5x}}{25}(6x+8)\,dx\right) \\
&= \frac{e^{-5x}}{-5}(x^3+4x^2+2) - \frac{e^{-5x}}{25}(3x^2+8x) + \int \frac{e^{-5x}}{25}(6x+8)\,dx \\
&= \frac{e^{-5x}}{-5}(x^3+4x^2+2) - \frac{e^{-5x}}{25}(3x^2+8x) + \frac{e^{-5x}}{-125}(6x+8) - \int \frac{e^{-5x}}{-125}\cdot 6\,dx \\
&= \boxed{\frac{e^{-5x}}{-5}(x^3+4x^2+2) - \frac{e^{-5x}}{25}(3x^2+8x) + \frac{e^{-5x}}{-125}(6x+8) - \frac{e^{-5x}}{625}\cdot 6 + C} \\
&= -\frac{e^{-5x}}{625}\Big(125(x^3+4x^2+2)+25(3x^2+8x)+5(6x+8)-6\Big)+C \\
&= -\frac{e^{-5x}}{625}\Big(125x^3+575x^2+230x+296\Big)+C.
\end{aligned}
$$

Om man tänker igenom vad som egentligen händer i denna uträkning så bör man inse att det går att **göra alla dessa partiella integrationer i ett enda svep**, så att man kommer direkt till det inramade uttrycket ovan, genom att upprepat **integrera exp-faktorn** och **derivera polynomet**, och sätta **varannan gång plus och minus mellan termerna**, och fortsätta tills man är nere på ett polynom som bara är en konstant (då är det klart, eftersom nästa derivata skulle ge noll-polynomet):

$$
\int e^{-5x}(x^3+4x^2+2)\,dx = \left(\frac{e^{-5x}}{-5}\right)(x^3+4x^2+2) - \left(\frac{e^{-5x}}{25}\right)(3x^2+8x) + \left(\frac{e^{-5x}}{-125}\right)(6x+8) - \left(\frac{e^{-5x}}{625}\right)\cdot 6 + C.
$$

(Parenteserna runt de blå faktorerna är lite överflödiga, men jag satte dit dem för att det skulle vara enklare att se skillnad på de teckenväxlingar som kommer från inre derivatan -5 vid integrationerna och de svarta minustecken som kommer från regeln ”varannan plus och minus”.) Därefter är det bara att snygga till, som i de två sista stegen ovan. Om man utför uträkningen på detta vis blir det mycket mindre att skriva, och risken för teckenfel minskar också.

- **Variabelbyte** består i grund och botten bara av att känna igen **kedjeregeln baklänges**. Derivering med kedjeregeln innebär ju att

$$\tfrac{d}{dx}F(g(x)) = f(g(x))\,g'(x)$$

(där f är derivatan av F), så när man deriverar baklänges, dvs. räknar ut primitiv funktion, blir det såhär:

$$\int f(g(x))\,g'(x)\,dx = F(g(x)) + C$$

Mellanstegen som man skriver vid variabelbyte är på sätt och vis bara bokföring där man avlastar hjärnan genom att låta pennan göra en del av jobbet:

$$
\boxed{\int \underbrace{f(g(x))}_{f(t)}\,\underbrace{g'(x)\,dx}_{dt} = \begin{bmatrix} t = g(x) \\ \frac{dt}{dx} = g'(x) \\ dt = g'(x)\,dx \end{bmatrix} = \int f(t)\,dt = F(t) + C = F(g(x)) + C.}
$$

Ifall f har en så enkel primitiv att man omedelbart kan utföra steget $\int f(t)\,dt = F(t)+C$ i huvudet finns det inte så stor anledning att skriva ut de där mellanstegen, utan då kan man lika gärna integrera direkt. T.ex. har $f(t) = \cos t$ den enkla primitiva funktionen $F(t) = \sin t$, så man kan med ett ögonkast se att

$$\int \cos(x^3)\cdot 3x^2\,dx = \sin(x^3) + C,$$

[Tillbaka till föreläsnings- och lektionsplan.]

bara man lägger märke till att $3x^2$ är derivatan av uttrycket x^3 som står inne i cosinusfunktionen.

Om däremot steget $\int f(t)\,dt = F(t) + C$ är mera komplicerat, t.ex. att det kräver partiell integration eller annat som man kanske inte klarar direkt med huvudräkning, så är variabelbytesmetoden oumbärlig, så det är inte helt rättvisande att avfärda den som ”bara bokföring” trots allt.

Och även i enklare fall är den användbar, ifall man är det minsta osäker. T.ex. om du inte riktigt har självförtroende nog att integrera direkt i det föregående exemplet, så kan du utföra variabelbytet istället (till priset av att du får skriva lite mer, men du behöver å andra sidan inte hålla riktigt lika mycket i huvudet samtidigt):

$$\int \underbrace{\cos(x^3)}_{\cos t}\,\underbrace{3x^2\,dx}_{dt} = \begin{bmatrix} t = x^3 \\ \frac{dt}{dx} = 3x^2 \\ dt = 3x^2\,dx \end{bmatrix} = \int \cos t\,dt = \sin t + C = \sin(x^3) + C.$$

- Variabelbyte är kanske ännu mer användbart när man använder det ”baklänges”, i situationer där man *inte* har någon inre derivata att känna igen i det uttryck man ska integrera.

  Såhär funkar det: ovan hade vi likheten

$$\int \underbrace{f(g(x))}_{f(t)}\,\underbrace{g'(x)\,dx}_{dt} = \begin{bmatrix} t = g(x) \\ \frac{dt}{dx} = g'(x) \\ dt = g'(x)\,dx \end{bmatrix} = \int f(t)\,dt,$$

och den kan såklart lika gärna läsas åt andra hållet:

$$\int f(t)\,dt = \begin{bmatrix} t = g(x) \\ \frac{dt}{dx} = g'(x) \\ dt = g'(x)\,dx \end{bmatrix} = \int \underbrace{f(g(x))}_{f(t)}\,\underbrace{g'(x)\,dx}_{dt}.$$

Låt oss här byta namn på variablerna, så att x och t får ombytta roller (så att det liknar den situation man brukar ha, med variabeln x från början och en ny variabel t efter variabelbytet):

$$\boxed{\int f(x)\,dx = \begin{bmatrix} x = g(t) \\ \frac{dx}{dt} = g'(t) \\ dx = g'(t)\,dt \end{bmatrix} = \int \underbrace{f(g(t))}_{f(x)}\,\underbrace{g'(t)\,dt}_{dx}.}$$

Om vi gör denna uträkning som den står, ”från vänster till höger”, *inför* vi alltså en inre derivata som inte fanns där från början. Även om det i formeln ser ut som att man får något krångligare i högerledet kan en sådan här manöver ofta göra att man kommer vidare i uträkningen.

Exempel.

$$\int \frac{dx}{1+e^x} = \begin{bmatrix} x = \ln t \quad (t > 0) \\ \frac{dx}{dt} = \frac{1}{t} \\ dx = \frac{dt}{t} \end{bmatrix} = \int \frac{dt}{(1+t)\,t}$$

$$= \int \left(\frac{1}{t} - \frac{1}{1+t}\right) dt = \ln t - \ln(1+t) + C = x - \ln(1+e^x) + C.$$

Här fick vi efter variabelbytet en rationell funktion, och kunde fortsätta med s.k. partialbråksuppdelning (som gås igenom ordentligt nästa gång, på Föreläsning 10).

Notera vad som hände i allra sista steget, när vi gick tillbaka till den den ursprungliga variabeln x. Termen $\ln t$ kunde direkt ersättas med bara x, för vårt variabelbyte var ju helt enkelt $x = \ln t$. Men för att uttrycka den andra termen $\ln(1+t)$ i variabeln x var vi tvungna att **lösa ut** t ur variabelbytet: $x = \ln t \iff t = e^x$, vilket gav att $\ln(1+t) = \ln(1+e^x)$.

För att denna "baklängesversion" av variabelbyte ska fungera krävs alltså i allmänhet att man gör ett **inverterbart** variabelbyte. Någon sådan restriktion finns inte i den första versionen av variabelbyte, där man känner igen en inre derivata som redan står där.

En sak som är bekväm här är att **satsen om derivata av invers funktion** gör att man får samma samband mellan dx och dt oavsett om man deriverar x som funktion av t eller om man deriverar t som funktion av x. Vi hade alltså även kunnat göra såhär i exemplet ovan:

$$\int \frac{dx}{1+e^x} = \begin{bmatrix} t = e^x \\ \frac{dt}{dx} = e^x = t \\ dx = \frac{dt}{t} \end{bmatrix} = \int \frac{dt}{(1+t)\,t} = \cdots$$

Men tänk på att inte blanda gamla och nya variabler i en och samma integral, utan se till att du efter variabelbytet direkt har **enbart den nya variabeln**. Alltså helst inga sådana här onödiga mellansteg när du redovisar (eftersom det ser lite **fult** ut med x och t i samma integral):

$$\int \frac{dx}{1+e^x} = \left[t = e^x \text{ etc.}\right] = \int \frac{dt}{(1+t)e^x} = \int \frac{dt}{(1+t)\,t} = \cdots$$

Men framför allt, se till att du åtminstone har gått över till enbart den nya variabeln innan du försöker *räkna vidare*, för annars kommer du att få ett helt felaktigt resultat!

- Som sades ovan behöver man inte bekymra sig för att variabelbytet $t = x^2$ inte är inverterbart i ett fall som detta, när man känner igen en inre derivata som man kan utnyttja tillsammans med dx:

$$\int x e^{x^2}\,dx = \frac{1}{2}\int \underbrace{e^{x^2}}_{e^t} \cdot \underbrace{2x\,dx}_{dt} = \begin{bmatrix} t = x^2 \\ \frac{dt}{dx} = 2x \\ dt = 2x\,dx \end{bmatrix} = \frac{1}{2}\int e^t\,dt = \frac{1}{2}e^t + C = \frac{1}{2}e^{x^2} + C.$$

Men se upp lite! Att här använda den "bakvända" varianten på följande vis vore inte riktigt rätt:

$$\int x e^{x^2}\,dx = \begin{bmatrix} x = \sqrt{t} \\ \frac{dx}{dt} = \frac{1}{2\sqrt{t}} \\ dx = \frac{dt}{2\sqrt{t}} \end{bmatrix} = \int \sqrt{t}\,e^t\,\frac{dt}{2\sqrt{t}} = \frac{1}{2}\int e^t\,dt = \frac{1}{2}e^t + C = \frac{1}{2}e^{x^2} + C.$$

Vad är problemet med detta, resultatet blev ju samma? Jo, $x = \sqrt{t}$ är ju alltid ≥ 0, så man har bara behandlat ickenegativa x, och det framgår alltså inte av denna uträkning att den primitiva funktion man hittat är korrekt även för $x < 0$. Fallet $x < 0$ skulle kräva en separat uträkning med $x = -\sqrt{t}$ istället.

Lektion 9

- Beräkning av primitiv funktion: P5.3, P5.4, P5.5, P5.6, P5.8, B5.8a (jfr. Exempel 5.11).
- Förklara skenbar paradox: P5.10.
- Hantering av absolutbelopp: P5.34*.
- Funktionsundersökning av primitiv funktion: P5.38, P5.37*.

[Tillbaka till föreläsnings- och lektionsplan.]

Föreläsning 10. Integration av rationella funktioner

Innehåll

- Avsnitt 5.3.
- Metod för att beräkna $\boxed{\int \frac{p(x)}{q(x)}\,dx}$ där p och q är polynom:
  - Om det behövs: **Polynomdividera** (för att få en integral där p har lägre grad än q).
  - Om det behövs: **Faktorisera nämnaren** i reella faktorer av grad ett och två, och **partialbråksuppdela** (motsatsen till att sätta på gemensamt bråkstreck).
  - Integrera partialbråken (enligt kända rutiner).

  Om man hamnar i detta läge, alltså att man har en rationell funktion att integrera, bör man **känna att det är lugnt**, för man behöver inte själv få någon snilleblixt, utan **det är bara att följa metoden**. I senare avsnitt, t.ex. vid integration av trigonometriska funktioner, söker vi ofta ett variabelbyte som leder tillbaka till denna välbekanta situation med en rationell funktion.

  Ibland kan man spara arbete om man råkar se någon genväg, t.ex.

$$\int \frac{x^3}{x^4+1}\,dx = \tfrac{1}{4}\ln(x^4+1)+C$$

  där situationen förenklades av att det (sånär som på en konstant faktor) var nämnarens derivata som stod i täljaren. Så håll gärna ögonen öppna efter sådant, men din ”defaultinställning” bör ändå vara att bara följa metoden.

Kommentarer

- **Polynomdivision** (liksom division av heltal) görs vanligen med **liggande stolen** eller **trappan** eller någon liknande uppställning. Tyvärr framgår det inte så tydligt av sådana procedurer vad det är man egentligen gör. Och ibland ser man somliga som lyckas utföra själva proceduren rätt, men sedan drar fel slutsats och t.ex. skriver upp svaret med kvoten och resten omkastade, eller andra konstigheter.

  Nedanstående exempel visar ett **alternativt sätt att utföra polynomdivision**, där det blir lite mer att skriva, men där det å andra sidan inte ingår något hokus-pokus som man måste memorera, utan bara vanlig enkel algebra, så att det syns precis vad som händer, och så att svaret automatiskt kommer på rätt form.

  Exempel. Nedan står en division som inte går jämnt ut. Men vi kan lägga till och dra ifrån termer som gör att *en del* av divisionen går jämnt ut åtminstone:

$$\begin{aligned}
&\frac{5x^4+3x^2}{x^2+2x+5} \\
&= \frac{5x^2\cdot x^2+3x^2}{x^2+2x+5} && \text{(bryt ut högsta termen nere från högsta termen uppe)} \\
&= \frac{5x^2\cdot\big(x^2+(2x+5)-(2x+5)\big)+3x^2}{x^2+2x+5} && \text{(lägg till \& dra ifrån för att matcha nämnaren)} \\
&= \frac{5x^2\cdot(x^2+2x+5)}{x^2+2x+5}+\frac{-5x^2\cdot(2x+5)+3x^2}{x^2+2x+5} && \text{(multiplicera in } 5x^2\text{, dela upp i två bråk)} \\
&= 5x^2+\frac{-10x^3-22x^2}{x^2+2x+5} && \text{(förkorta det första bråket, snygga till det andra).}
\end{aligned}$$

  Termen $5x^2$, som kom från den del av divisionen som gick jämnt ut (alltså där vi kunde förkorta), är nu färdig och kommer bara att ”hänga med” i fortsättningen av uträkningen. I nästa steg utför vi

[Tillbaka till föreläsnings- och lektionsplan.]

samma procedur på det bråk som återstår, dvs. på den del av divisionen som inte gick jämnt ut:

$$
\begin{aligned}
\frac{5x^4+3x^2}{x^2+2x+5} = \cdots &= 5x^2 + \frac{-10x^3-22x^2}{x^2+2x+5} \\
&= 5x^2 + \frac{-10x\cdot x^2-22x^2}{x^2+2x+5} \\
&= 5x^2 + \frac{-10x\cdot\Big(x^2+(2x+5)-(2x+5)\Big)-22x^2}{x^2+2x+5} \\
&= 5x^2 + \frac{-10x\cdot(x^2+2x+5)}{x^2+2x+5} + \frac{10x\cdot(2x+5)-22x^2}{x^2+2x+5} \\
&= 5x^2 + (-10x) + \frac{-2x^2+50x}{x^2+2x+5}.
\end{aligned}
$$

Och så gör vi samma sak igen med den återstående delen:

$$
\begin{aligned}
\frac{5x^4+3x^2}{x^2+2x+5} = \cdots &= 5x^2 + (-10x) + \frac{-2x^2+50x}{x^2+2x+5} \\
&= 5x^2 - 10x + \frac{-2\cdot x^2+50x}{x^2+2x+5} \\
&= 5x^2 - 10x + \frac{-2\cdot\Big(x^2+(2x+5)-(2x+5)\Big)+50x}{x^2+2x+5} \\
&= 5x^2 - 10x + \frac{-2\cdot(x^2+2x+5)}{x^2+2x+5} + \frac{2\cdot(2x+5)+50x}{x^2+2x+5} \\
&= 5x^2 - 10x + (-2) + \frac{54x+10}{x^2+2x+5}.
\end{aligned}
$$

Nu går det inte att fortsätta längre, eftersom täljaren i den återstående divisionen har **lägre gradtal än nämnaren**; då kan man ju inte bryta ut högsta termen nere från högsta termen uppe och fortfarande få ett *polynom*. Denna kvarvarande täljare är den **rest** som blir över ifall inte hela divisionen råkar gå jämnt ut. Resultatet av polynomdivisionen är därmed

$$
\frac{\overbrace{5x^4+3x^2}^{\text{täljare}}}{\underbrace{x^2+2x+5}_{\text{nämnare}}} = \underbrace{5x^2-10x-2}_{\text{kvot}} + \frac{\overbrace{54x+10}^{\text{rest}}}{\underbrace{x^2+2x+5}_{\text{nämnare}}}.
$$

Eller ekvivalent, om man multiplicerar med nämnaren på båda sidor:

$$
\underbrace{5x^4+3x^2}_{\text{täljare}} = \underbrace{(5x^2-10x-2)}_{\text{kvot}}\cdot\underbrace{(x^2+2x+5)}_{\text{nämnare}} + \underbrace{(54x+10)}_{\text{rest}}.
$$

Observera att man här enkelt kan **kontrollera att man räknat rätt** genom att multiplicera ut högerledet och se om det stämmer med vad som står i vänsterledet.

Prova gärna att utföra denna division med liggande stolen, så ser du förhoppningsvis att det är precis samma uträkningar som du gör, bara uppställda på ett annat sätt:

kvoten växer fram här uppe

$$
\begin{array}{r l|l l}
 & 5x^2+\cdots & & \\
\hline
\text{täljare från start} \longrightarrow & 5x^4+3x^2 & x^2+2x+5 & \longleftarrow \text{nämnare} \\
\cline{3-3}
 & -5x^2(x^2+2x+5) & & \\
\cline{2-2}
\text{ny täljare i nästa steg} \longrightarrow & \quad -10x^3-22x^2 & & \\
 & \qquad\ddots & &
\end{array}
$$

resten blir över här nere

[Tillbaka till föreläsnings- och lektionsplan.]

- Den typ av partialbråk som är besvärligast att integrera är
$$\frac{1}{(x^2+ax+b)^n},$$
där $n \geq 2$ och x^2+ax+b har icke-reella nollställen (dvs. inte kan faktoriseras i reella förstagradsfaktorer). Ett sätt att göra detta beskrivs i boken på s. 260–262. Här är en **alternativ metod**, som jag tycker är lite smidigare. Man kan alltid börja med att kvadratkomplettera och sätta $y = x + \frac{a}{2}$, för att få en integral av typen
$$\int \frac{dy}{(y^2+c^2)^n}, \qquad c > 0,$$
och sedan sätta $y = ct$, för att få en konstant gånger integralen
$$\boxed{\int \frac{dt}{(t^2+1)^n}},$$
så det räcker att veta hur man beräknar denna typ av integral.

För att spara bläck inför vi förkortningen
$$\boxed{T = T(t) = t^2+1}$$
så det vi vill beräkna är alltså
$$\boxed{\int \frac{1}{T^n}\,dt}, \qquad n \geq 2.$$
Notera att
$$\frac{d}{dt}\left(\frac{t}{T^n}\right) = \frac{d}{dt}\left(t \cdot \frac{1}{T^n}\right) = 1 \cdot \frac{1}{T^n} + t \cdot \frac{-n}{T^{n+1}} \cdot 2t = \frac{1}{T^n} - \frac{2nt^2}{T^{n+1}} = \frac{1}{T^n} - \frac{2n(T-1)}{T^{n+1}} = \frac{2n}{T^{n+1}} - \frac{2n-1}{T^n}.$$
Vi kan bokföra detta resultat som en ny "standardprimitiv"[8]:
$$\boxed{\int \left(\frac{2n}{T^{n+1}} - \frac{2n-1}{T^n}\right) dt = \frac{t}{T^n} + C.}$$
Eller utskrivet med $n = 1, 2, 3, 4, \ldots$ insatt:
$$\boxed{\begin{gathered}
\int \left(\frac{2}{T^2} - \frac{1}{T^1}\right) dt = \frac{t}{T} + C \\
\int \left(\frac{4}{T^3} - \frac{3}{T^2}\right) dt = \frac{t}{T^2} + C \\
\int \left(\frac{6}{T^4} - \frac{5}{T^3}\right) dt = \frac{t}{T^3} + C \\
\int \left(\frac{8}{T^5} - \frac{7}{T^4}\right) dt = \frac{t}{T^4} + C \\
\vdots
\end{gathered}}$$
Och vi vet såklart också att
$$\boxed{\int \frac{1}{T}\,dt = \int \frac{dt}{1+t^2} = \arctan t + C.}$$
Exempel. Låt oss se hur vi med hjälp av detta kan beräkna
$$\int \frac{dt}{(t^2+1)^3} = \int \frac{dt}{T^3}.$$

[8] Det är knappast lönt att memorera denna formel, utan man bara återhärleder den vid behov.

[Tillbaka till föreläsnings- och lektionsplan.]

Knepet är att skriva $\dfrac{1}{T^3}$ som en linjärkombination av uttrycken

$$\frac{4}{T^3}-\frac{3}{T^2}, \qquad \frac{2}{T^2}-\frac{1}{T^1}, \qquad \frac{1}{T},$$

vars primitiva funktioner är kända enligt ovan. Detta kan göras med några ”lägga till och dra ifrån”-knep. Först får man

$$\frac{1}{T^3}=\frac{1}{4}\frac{4}{T^3}=\frac{1}{4}\left(\frac{4}{T^3}-\frac{3}{T^2}+\frac{3}{T^2}\right)=\frac{1}{4}\left(\frac{4}{T^3}-\frac{3}{T^2}\right)+\frac{3}{4}\frac{1}{T^2},$$

och sedan likadant med

$$\frac{1}{T^2}=\frac{1}{2}\frac{2}{T^2}=\frac{1}{2}\left(\frac{2}{T^2}-\frac{1}{T^1}+\frac{1}{T^1}\right)=\frac{1}{2}\left(\frac{2}{T^2}-\frac{1}{T^1}\right)+\frac{1}{2}\frac{1}{T},$$

vilket tillsammans ger

$$\begin{aligned}\frac{1}{T^3}&=\frac{1}{4}\left(\frac{4}{T^3}-\frac{3}{T^2}\right)+\frac{3}{4}\left(\frac{1}{2}\left(\frac{2}{T^2}-\frac{1}{T^1}\right)+\frac{1}{2}\frac{1}{T}\right)\\&=\frac{1}{4}\left(\frac{4}{T^3}-\frac{3}{T^2}\right)+\frac{3}{8}\left(\frac{2}{T^2}-\frac{1}{T^1}\right)+\frac{3}{8}\frac{1}{T}\end{aligned}$$

Med lite vana gör man hela denna omskrivning i ett enda svep, och sedan integrerar man term för term:

$$\boxed{\begin{aligned}\int\frac{dt}{(t^2+1)^3}=\int\frac{dt}{T^3}&=\int\left(\frac{1}{4}\left(\frac{4}{T^3}-\frac{3}{T^2}\right)+\frac{3}{4}\cdot\frac{1}{2}\left(\frac{2}{T^2}-\frac{1}{T^1}\right)+\frac{3}{8}\frac{1}{T}\right)dt\\&=\frac{1}{4}\int\left(\frac{4}{T^3}-\frac{3}{T^2}\right)dt+\frac{3}{8}\int\left(\frac{2}{T^2}-\frac{1}{T^1}\right)dt+\frac{3}{8}\int\frac{dt}{T}\\&=\frac{1}{4}\frac{t}{T^2}+\frac{3}{8}\frac{t}{T}+\frac{3}{8}\arctan t+C\\&=\frac{1}{4}\frac{t}{(t^2+1)^2}+\frac{3}{8}\frac{t}{t^2+1}+\frac{3}{8}\arctan t+C\\&=\frac{3t^3+5t}{8(t^2+1)^2}+\frac{3}{8}\arctan t+C.\end{aligned}}$$

Exempel. Denna metod är ännu effektivare om man har en hel kedja

$$\frac{a_n}{T^n}+\frac{a_{n-1}}{T^{n-1}}+\cdots+\frac{a_2}{T^2}+\frac{a_1}{T^1},$$

vilket är vad man typiskt får som del av en partialbråksuppdelning ifall faktorn T^n finns i nämnaren. Istället för att först räkna ut $\int\frac{dt}{T^2}$ med hjälp av att man vet $\int\frac{dt}{T}$, sedan $\int\frac{dt}{T^3}$ med hjälp av att man

[Tillbaka till föreläsnings- och lektionsplan.]

vet $\int \frac{dt}{T^2}$, osv., så kan man ta hela kedjan på en gång genom att lägga till och dra ifrån, såhär:

$$
\begin{aligned}
\int \frac{t^6-t^2+1}{(t^2+1)^4}\,dt = \big[\mathrm{PBU}\big] &= \int \left(\frac{1}{(t^2+1)^4} + \frac{2}{(t^2+1)^3} + \frac{-3}{(t^2+1)^2} + \frac{1}{t^2+1} \right) dt \\
&= \int \left(\frac{1}{T^4} + \frac{2}{T^3} + \frac{-3}{T^2} + \frac{1}{T} \right) dt \\
&= \int \left(\frac{1}{6}\left(\frac{6}{T^4} - \frac{5}{T^3}\right) + \frac{\frac{5}{6}+2}{T^3} + \frac{-3}{T^2} + \frac{1}{T} \right) dt \\
&= \frac{1}{6}\frac{t}{T^3} + \int \left(\frac{17}{6}\cdot\frac{1}{4}\left(\frac{4}{T^3} - \frac{3}{T^2}\right) + \frac{\frac{17}{8}-3}{T^2} + \frac{1}{T} \right) dt \\
&= \frac{1}{6}\frac{t}{T^3} + \frac{17}{24}\frac{t}{T^2} + \int \left(\frac{-7}{8}\cdot\frac{1}{2}\left(\frac{2}{T^2} - \frac{1}{T^1}\right) + \frac{\frac{-7}{16}+1}{T} \right) dt \\
&= \frac{1}{6}\frac{t}{T^3} + \frac{17}{24}\frac{t}{T^2} - \frac{7}{16}\frac{t}{T} + \frac{9}{16}\arctan t + C \\
&= \frac{1}{6}\frac{t}{(t^2+1)^3} + \frac{17}{24}\frac{t}{(t^2+1)^2} - \frac{7}{16}\frac{t}{t^2+1} + \frac{9}{16}\arctan t + C \\
&= \frac{-21t^5-8t^3+21t}{48(1+t^2)^3} + \frac{9}{16}\arctan t + C.
\end{aligned}
$$

Lektion 10

- En gammal bekanting som varit med tidigare: B5.25. (Jfr. P4.9f.)
- Varning för klassiskt nybörjarmisstag: B5.27.
  (”Du skall icke dividera med inre derivatan.”)
- Partialbråksuppdelning: P5.12, P5.13, P5.14.
- Beräkning av primitiv till rationell funktion: P5.15, P5.16, P5.36*.

Lektion 11

- Fler rationella funktioner att integrera: P5.17abcefg.
- Den besvärligaste typen av partialbråk: P5.17h*.
- Partiell integration leder till rationell integrand: P5.20.
- Diverse integraler: P5.21.
- Bestäm konstanter så att primitiven blir rationell (inga ln- eller arctan-termer): P5.35*.

Föreläsning 11. Integration av trigonometriska funktioner och rotuttryck

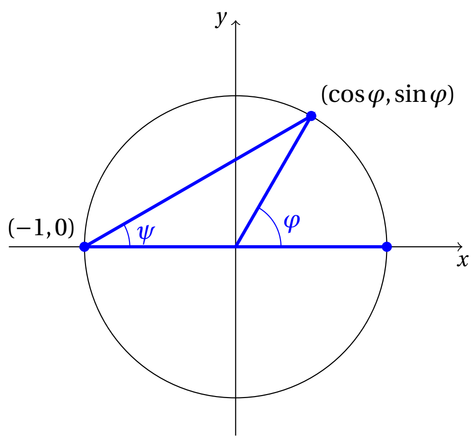
Innehåll

- Avsnitt 5.4 och 5.5.

[Tillbaka till föreläsnings- och lektionsplan.]

Kommentarer

- När man gör variabelbytet $t = \tan(\varphi/2)$ behöver man formlerna för $\cos\varphi$ och $\sin\varphi$ uttryckta i variabeln t:
$$\cos\varphi = \frac{1-t^2}{1+t^2}, \qquad \sin\varphi = \frac{2t}{1+t^2}.$$
Härledningen i boken (Exempel 5.35 och övning 5.21) använder formlerna för dubbla vinkeln ihop med trig-ettan:
$$\sin(\varphi) = \frac{\sin(2\cdot\frac{\varphi}{2})}{1} = \frac{2\sin\frac{\varphi}{2}\cos\frac{\varphi}{2}}{\cos^2\frac{\varphi}{2}+\sin^2\frac{\varphi}{2}} = \frac{2\cdot\frac{\sin\frac{\varphi}{2}}{\cos\frac{\varphi}{2}}}{1+\left(\frac{\sin\frac{\varphi}{2}}{\cos\frac{\varphi}{2}}\right)^2} = \frac{2\tan\frac{\varphi}{2}}{1+\tan^2\frac{\varphi}{2}} = \frac{2t}{1+t^2},$$
och analogt för $\cos\varphi$ via $\cos(2\cdot\frac{\varphi}{2}) = \cos^2\frac{\varphi}{2} - \sin^2\frac{\varphi}{2}$.
Här är en mer geometrisk härledning, med hjälp av enhetscirkeln:



Randvinkeln ψ är hälften av medelpunktsvinkeln φ, enligt randvinkelsatsen:
$$\psi = \frac{\varphi}{2}.$$
Linjen som går mellan punkterna $(-1,0)$ och $(\cos\varphi, \sin\varphi)$ har ekvationen $y = t(x+1)$, med riktningskoefficienten
$$t = \tan\psi = \tan(\varphi/2).$$
Ekvationssystemet
$$x^2 + y^2 = 1, \qquad y = t(x+1)$$
talar om var denna linje skär cirkeln. Lös ut $x = \frac{y}{t} - 1$ från den andra ekvationen och sätt in i den första; detta ger
$$\left(\frac{y}{t}-1\right)^2 + y^2 = 1 \iff y^2\left(\frac{1}{t^2}+1\right) - \frac{2y}{t} = 0 \iff y = 0 \quad \text{eller} \quad y = \frac{2t}{1+t^2}.$$
Dessa värden insatta i ekvationen $y = t(x+1)$ ger de motsvarande x-värdena, $x = -1$ resp. $x = \frac{1-t^2}{1+t^2}$. Så de två skärningspunkterna är $(x, y) = (-1, 0)$ (såklart) samt
$$(x, y) = (\cos\varphi, \sin\varphi) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right).$$
Alltså: $\cos\varphi = \frac{1-t^2}{1+t^2}$ och $\sin\varphi = \frac{2t}{1+t^2}$, som önskat.

[Tillbaka till föreläsnings- och lektionsplan.]

- Sidokommentar: Formlerna $(x, y) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$ ger en parametrisering av enhetscirkeln med rationella funktioner, där rationella värden på parametern t motsvarar punkter på enhetscirkeln med rationella koordinater x och y, vilket i sin tur (efter multiplikation med den gemensamma nämnaren) motsvarar **pythagoreiska** (eller **egyptiska**) trianglar, dvs. rätvinkliga trianglar vars sidlängder är heltal.

  Exempelvis $t = 1/2$ ger $(x, y) = (3/5, 4/5)$, och att denna punkt ligger på enhetcirkeln är ekvivalent med att punkten $(3, 4)$ ligger på avståndet 5 från origo, dvs. att triangeln med sidorna 3, 4 och 5 är rätvinklig.

Lektion 12

- Enkel rimlighetskontroll: P5.1, P5.2.
- Beräkning av trigonometriska primitiver: P5.23abcdef, P5.25a, P5.24, P5.23i*, P5.27*.
- Beräkning av primitiv till rotuttryck: P5.28, P5.29abcfh, P5.30, P5.33c, P5.40*.

Föreläsning 12. Riemannintegral (bestämd integral)

Föreläsning 13. Riemannintegral, forts.

Innehåll

- Avsnitt 6.1–6.4.
- Definition av Riemannintegralen $\int_a^b f(x)\,dx$.
- Räkneregler och andra egenskaper.
- Analysens huvudsats. Insättningsformeln.

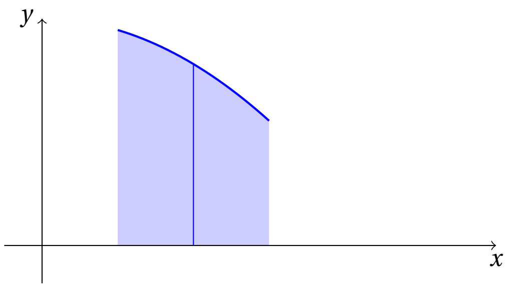
Kommentarer

- Boken säger såhär i kapitlet om primitiva funktioner (s. 238), apropå notationen $\int f(x)\,dx$:

  > ”Uttrycket dx kommer att få sin förklaring när vi kommer till avsnittet om integraler.”

  Det står lite om dx i derivatakapitlet (s. 180), men såvitt jag kan se säger de inte så mycket om den saken i integralkapitlet, så det kanske kan vara värt en kommentar här.

  Skrivsättet $\int y\,dx$ infördes av Leibniz på 1600-talet, och integraltecknet $\int$ är ett stiliserat S som i ”summa”. För att beräkna arean av området under en kurva, betrakta området som bestående av oändligt många lodräta linjesegment, ett för varje x i intervallet ifråga:



[Tillbaka till föreläsnings- och lektionsplan.]

I vanliga fall skulle man väl säga att arean av ett sådant linjesegment är noll, men Leibniz säger att man istället ska betrakta det som en rektangel med höjden y och en oändligt liten (”infinitesimal”) bredd dx, där bokstaven d står för att dx är en ”differens” mellan två ”oändligt närliggande” x-värden. Linjesegmentets area $dA = y\,dx$ är då oändligt liten (men inte noll), och områdets totala area A är ”summan” av alla dessa oändligt många oändligt små areabidrag dA:

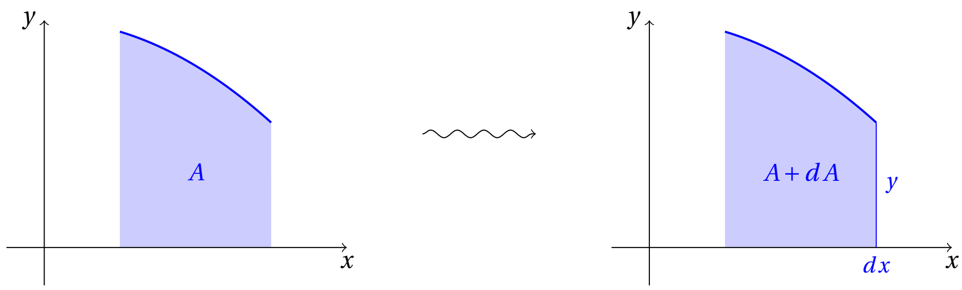
$$A = \int dA = \int y\,dx.$$

Funktionsskrivsättet $y = f(x)$ infördes senare, av Euler på 1700-talet. Skrivsättet

$$\int_a^b f(x)\,dx$$

för att markera integrationsintervallet $[a, b]$ infördes *mycket* senare, av Fourier i 1800-talets början. Fram tills dess hade man specificerat gränserna i ord eller låtit dem framgå av sammanhanget.

Man skilde inte heller på bestämd integral (”summa”) och obestämd integral (primitiv funktion), för det var ju ”uppenbart” samma sak. Ty anta att man börjar med en area A, och ”lägger till bara ett enda linjesegment med höjd y och bredd dx” i högra kanten:



Hur mycket har arean då ändrats? Jo, ändringen dA består ju bara av den infinitesimala area $y\,dx$ som vi har lagt till, och då får vi ju

$$dA = y\,dx \qquad \Longrightarrow \qquad y = \frac{dA}{dx},$$

dvs. kurvan som vi började med är derivatan av areafunktionen som mäter arean upp till och med punkten x, dvs. areafunktionen är en primitiv funktion till den som vi började med. (Och det var ju också detta som var det fiffiga: genom att lätt beräkna primitiv funktion kunde man lösa areaberäkningsproblem som tidigare varit svåra.)

Det var först på 1800-talet som matematiker på allvar började oroa sig för hur allt det här egentligen hänger ihop logiskt, och hur man skulle kunna förklara det **enbart med hjälp av det reella talsystemet**, som ju **inte** innehåller några sådana där mystiska tal som är ”oändligt små men inte noll”.

- För att bättre förstå analysens huvudsats (integration och derivering som inversa operationer) kan det kanske hjälpa att jämföra med summering och subtraktion som inversa operationer.

  T.ex. när man spelar minigolf protokollför somliga resultaten bana för bana och lägger sedan ihop allt på slutet:

Bana	1	2	3	4	⋯	18	S:a
Linus	3	7	1	4		7	84
Linnea	2	4	5	6		1	62

Andra föredrar att lägga ihop resultaten vartefter, så att man kan se den nuvarande ställningen hela tiden:

[Tillbaka till föreläsnings- och lektionsplan.]

Bana	1	2	3	4	$\cdots$	18	S:a
Linus	3	10	11	15		84	84
Linnea	2	6	11	17		62	62

Detta sätt att bokföra innehåller uppenbart exakt samma information som det första sättet, för man kan ju t.ex. beräkna differensen 17 − 11 för att se att Linnea slog en sexa på bana 4.

Mer abstrakt, om man för en given talföljd

$$a_1, a_2, a_3, a_4, \ldots$$

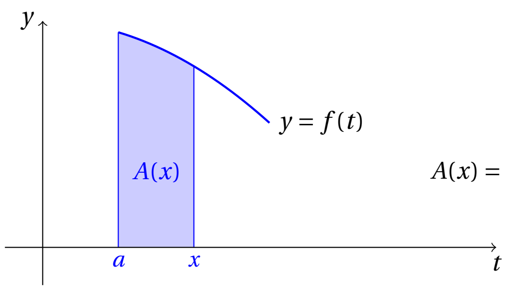
beräknar följden av ackumulerade summor

$$\begin{aligned}
s_1 &= a_1 \\
s_2 &= a_1 + a_2 \\
s_3 &= a_1 + a_2 + a_3 \\
s_4 &= a_1 + a_2 + a_3 + a_4 \\
&\vdots \\
s_n &= a_1 + a_2 + a_3 + a_4 + \cdots + a_n = \sum_{k=1}^{n} a_k \\
&\vdots
\end{aligned}$$

så kan man **återskapa** ursprungsföljden $(a_k)_{k=1}^{\infty}$ från den ihop-ackumulerade följden $(s_n)_{n=1}^{\infty}$ **genom att subtrahera** successiva värden:

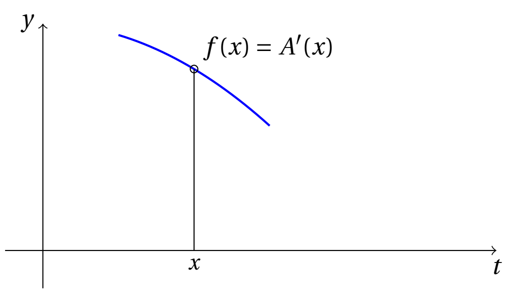
$$\begin{aligned}
a_1 &= s_1, \\
a_n &= s_n - s_{n-1}, \qquad n = 2, 3, 4, \ldots
\end{aligned}$$

Om man istället för en talföljd har en funktion $f(t)$, dvs. "en strid ström av data", ett värde för varje $t \in \mathbf{R}$, så motsvaras de ackumulerade summorna av att man integrerar ihop f:s alla värden från någon given startpunkt (säg $t = a$) fram till $t = x$:



$$A(x) = \int_a^x f(t)\,dt$$

Och då kan man **återskapa** ursprungsfunktionen f (under förutsättning att den är kontinuerlig) **genom att derivera** den ihop-ackumulerade funktionen A:



[Tillbaka till föreläsnings- och lektionsplan.]

Notera att man i den nutida notationen, där man anger integrationsgränserna med $\int_a^x$, behöver använda ett annat variabelnamn (exempelvis t) i integralen, så att inte symbolen x används i dubbla betydelser. Detta är alltså av samma anledning som man kallar summationsindexet för k och inte n i formeln $s_n = \sum_{k=1}^n a_k$. Att skriva $s_n = \sum_{n=1}^n a_n$ skulle se konstigt ut, eller hur? Och $A(x) = \int_a^x f(x)\,dx$ ser lika konstigt ut.

- Observera att vid variabelbyte i bestämd integral måste man **byta gränserna** också:

$$\int_a^b f(g(x))\,g'(x)\,dx = \left[\begin{array}{l} \quad t = g(x) \\ \frac{dt}{dx} = g'(x) \\ dt = g'(x)\,dx \\ x = a \implies t = g(a) \\ x = b \implies t = g(b) \end{array}\right] = \int_{g(a)}^{g(b)} f(t)\,dt = \Big[F(t)\Big]_{g(a)}^{g(b)} = F(g(b)) - F(g(a)).$$

Det är **inte** korrekt att skriva

$$\cdots = \int_a^b f(t)\,dt = \cdots$$

i mittenledet här!

Som synes blir slutresultatet samma som om man först räknar ut den primitiva funktionen $F(g(x))$, uttryckt i den ursprungliga variabeln x, så som vi gjorde i avsnittet om primitiva funktioner, och därefter sätter in gränserna för x:

$$\cdots = \Big[F(g(x))\Big]_a^b = F(g(b)) - F(g(a)).$$

Men när man håller på med bestämda integraler är det ju onödigt att gå tillbaka till den gamla variabeln på detta vis; sätt bara in de nya gränserna för t i $F(t)$ så är det klart.

- I fallet ovan, där man har $g'(x)$ tillsammans med dx från början, så måste inte g vara inverterbar (jfr. diskussionen i kommentarerna till Föreläsning 9). Det kan alltså mycket väl hända att t.ex. $g(a) = g(b)$, och det betyder bara att hela integralen blir noll.

Men om man använder den ”bakvända” metoden där man inte har någon inre derivata från början så måste variabelbytet vara **inverterbart** (åtminstone på integrationsintervallet ifråga):

$$\int_a^b f(x)\,dx = \left[\begin{array}{l} \quad x = g(t)\ \big(\text{eller } t = g^{-1}(x)\big) \\ \frac{dx}{dt} = g'(t)\ \big(\text{eller } \frac{dt}{dx} = (g^{-1})'(x) = \frac{1}{g'(g^{-1}(x))} = \frac{1}{g'(t)}\big) \\ dx = g'(t)\,dt \\ x = a \implies t = g^{-1}(a) \\ x = b \implies t = g^{-1}(b) \end{array}\right] = \int_{g^{-1}(a)}^{g^{-1}(b)} f(t)\,g'(t)\,dt = \cdots$$

Annars kan man råka ut för sådana här orimligheter:

$$\underbrace{\int_0^\pi \sin^2 x\,dx}_{\text{uppenbart} > 0 \text{ ju}} = \left[\begin{array}{l} \quad t = \sin x \quad (\text{Obs! Ej inv.bart på } [0,\pi].) \\ \frac{dt}{dx} = \cos x = \sqrt{1-\sin^2 x} = \sqrt{1-t^2} \\ dx = \dfrac{dt}{\sqrt{1-t^2}} \\ x = 0 \implies t = \sin 0 = 0 \\ x = \pi \implies t = \sin \pi = 0 \end{array}\right] \overset{?}{=} \underbrace{\int_0^0 t^2 \frac{dt}{\sqrt{1-t^2}}}_{=0\ (?!?)}.$$

Problemet här är bl.a. att $\cos x = \sqrt{1-\sin^2 x}$ bara gäller i den vänstra halvan av integrationsintervallet (fram till $\pi/2$); i den högra halvan gäller istället $\cos x = -\sqrt{1-\sin^2 x}$, med ett minustecken framför roten. Följande är dock sant (om än inte så användbart, det finns bättre sätt att räkna ut denna integral):

$$\int_0^{\pi} \sin^2 x\,dx = \int_0^{\pi/2} \sin^2 x\,dx + \int_{\pi/2}^{\pi} \sin^2 x\,dx$$
$$= \int_0^1 t^2 \frac{dt}{\sqrt{1-t^2}} + \int_1^0 t^2 \frac{dt}{-\sqrt{1-t^2}} = 2\int_0^1 t^2 \frac{dt}{\sqrt{1-t^2}},$$

där variabelbytet $t = \sin x$ har gjorts i delintegralerna var för sig; i den vänstra halvan av intervallet har vi då $x = \arcsin t$, men i den högra har vi $x = \frac{\pi}{2} - \arcsin t$.

Vi kan väl här också passa på att "bevisa" att **alla** integraler $\int_a^b f(x)\,dx$ faktiskt är noll! Det är ju bara att sätta $t = (x-a)(x-b)$ så fås $\int_0^0 (\cdots)\,dt = 0$, eller hur? Om denna revolutionerande upptäckt vore sann skulle alla analysstudenters liv bli mycket enklare.

- Här är en relaterad fälla som man lätt kan gå i om man inte ser upp. Säg att vi vill beräkna (den uppenbart positiva) integralen

$$\int_0^{\pi} \underbrace{\frac{1}{1+\sin^2 x}}_{f(x)}\,dx.$$

Låt oss ta fram primitiv funktion först, med variabelbytet $x = \arctan t$, $dx = dt/(1+t^2)$:

$$\int \frac{1}{1+\sin^2 x}\,dx = \int \frac{1}{1+\sin^2(\arctan t)} \frac{dt}{1+t^2} = \int \frac{1}{1+\left(\dfrac{t}{\sqrt{1+t^2}}\right)^2} \frac{dt}{1+t^2}$$
$$= \int \frac{dt}{1+2t^2} = \int \frac{dt}{1+\left(\sqrt{2}\,t\right)^2} = \frac{1}{\sqrt{2}} \arctan(\sqrt{2}\,t) + C = \underbrace{\frac{1}{\sqrt{2}} \arctan(\sqrt{2}\tan x)}_{F(x)} + C.$$

Insättning av gränserna i denna primitiv ger dock det uppenbart orimliga resultatet

$$0 < \int_0^{\pi} \frac{1}{1+\sin^2 x}\,dx \overset{?}{=} \left[\frac{1}{\sqrt{2}} \arctan(\sqrt{2}\tan x)\right]_0^{\pi} = 0 - 0 = 0.$$

Problemet är att F inte är en primitiv funktion till f på **hela** integrationsintervallet $[0,\pi]$, vilket är en förutsättning för att insättningsformeln ska gälla! I själva verket är funktionen F odefinierad för $x = \pi/2$, där den har en singularitet i form av ett språng (vänstergränsvärdet är $\frac{\pi}{2\sqrt{2}}$ medan högergränsvärdet är $-\frac{\pi}{2\sqrt{2}}$).

Detta innebär att man får trixa lite "för hand" för att skapa en primitiv som är giltig på hela intervallet:

$$F_2(x) = \begin{cases} F(x), & 0 \le x < \pi/2, \\ \frac{\pi}{2\sqrt{2}}, & x = \pi/2, \\ F(x) + 2 \cdot \frac{\pi}{2\sqrt{2}}, & \pi/2 < x \le \pi. \end{cases}$$

Insättning av gränserna i *denna* primitiv ger det rätta resultatet:

$$\int_0^{\pi} \frac{1}{1+\sin^2 x}\,dx = F_2(\pi) - F_2(0) = \left(F(\pi) + 2 \cdot \frac{\pi}{2\sqrt{2}}\right) - F(0) = 0 + \frac{\pi}{\sqrt{2}} - 0 = \frac{\pi}{\sqrt{2}}.$$

Sådana här fenomen kan ställa till med en del besvär för datoralgebrasystem också. På nätet kan man hitta flera artiklar i detta ämne av matematikern David J. Jeffrey, t.ex. **The Importance of Being Continuous** (*Mathematics Magazine, 1994*).

- Symmetriegenskaperna

$$\int_{-a}^{a} f(x)\,dx = 2\int_0^a f(x)\,dx \qquad \text{om } f \text{ är en jämn (och integrerbar) funktion}$$

[Tillbaka till föreläsnings- och lektionsplan.]

och

$$\int_{-a}^{a} f(x)\,dx = 0 \qquad \text{om } f \text{ är en udda (och integrerbar) funktion}$$

är ofta användbara, särskilt den senare (man behöver ju inte beräkna primitiv funktion, integralen blir noll direkt).

Ibland kan det även finnas ”dold symmetri” som man kan utnyttja. Ta t.ex.

$$\int_{-\pi}^{\pi} \underbrace{\frac{1}{1+e^{\sin x}}}_{f(x)}\,dx.$$

Integranden $f(x)$ är varken udda eller jämn här, men om man ritar grafen med dator så märker man att den ser rätt symmetrisk ut, som en konstant plus en udda funktion. Och så är det faktiskt, vilket man kan inse på följande sätt. Varje funktion vars definitionsmängd är symmetrisk kring origo kan delas upp i en summa av en udda och en jämn funktion, såhär:

$$f(x) = \underbrace{\frac{f(x)+f(-x)}{2}}_{\text{jämna delen av } f} + \underbrace{\frac{f(x)-f(-x)}{2}}_{\text{udda delen av } f} = j(x) + u(x).$$

I detta fall har vi

$$f(-x) = \frac{1}{1+e^{\sin(-x)}} = \frac{1}{1+e^{-\sin x}} = \frac{e^{\sin x}}{e^{\sin x}+1} = \frac{e^{\sin x}}{1+e^{\sin x}},$$

vilket gör att den jämna delen av f helt enkelt är

$$j(x) = \frac{f(x)+f(-x)}{2} = \frac{1}{2}\left(\frac{1}{1+e^{\sin x}} + \frac{e^{\sin x}}{1+e^{\sin x}}\right) = \frac{1}{2}\cdot\frac{1+e^{\sin x}}{1+e^{\sin x}} = \frac{1}{2},$$

alltså en konstant. Därmed har vi $f(x) = \frac{1}{2} + u(x)$ där u är en udda funktion, vilket ger

$$\int_{-\pi}^{\pi} \frac{1}{1+e^{\sin x}}\,dx = \int_{-\pi}^{\pi} \left(\tfrac{1}{2} + u(x)\right)dx = \int_{-\pi}^{\pi} \tfrac{1}{2}\,dx + \int_{-\pi}^{\pi} u(x)\,dx = \tfrac{1}{2}\cdot 2\pi + 0 = \pi,$$

dvs. vi kunde räkna ut integralen utan att hitta primitiv funktion (vilket torde vara ett ganska hopplöst företag i detta fall).

Den som tyckte att detta var kul kan t.ex. testa att räkna ut $\int_0^4 \frac{dx}{4+2^x}$ på liknande sätt. (Tips: Om man plottar kurvan $y = \frac{1}{4+2^x}$ med dator ser den ut att ha en udda symmetri kring punkten $(x, y) = (2, 1/8)$, eller hur?)

Lektion 13

- Integral av trappfunktion: B6.1.
- Uppskattning av integral med under- och övertrappor: P6.4.
- Beräkning av bestämd integral direkt ur definitionen: B6.2*.
- Beräkning av bestämda integraler med insättningsformeln: P6.3achk, P6.9dhjk, P6.11acfg, B6.11f.

  Observera hur man hanterar **integrationsgränserna** vid partiell integration och variabelbyte. (Sats 6.9 och 6.10, samt kommentarer ovan.)

[Tillbaka till föreläsnings- och lektionsplan.]

Lektion 14

- Uppskattning av integral: P6.1.
- Bestämd integral med absolutbelopp: P6.8ab.

  Ledning: Dela upp intervallet, för att göra dig kvitt beloppstecknen. Det är ingen bra idé att försöka räkna ut primitiv funktion så länge det finns absolutbelopp kvar i uttrycket…
- Ett par integraler som blir noll: B6.11e, P6.14*.

  Det går att räkna ut dessa integraler med insättningsformeln som vanligt. Men en intressantare utmaning är att försöka visa att de blir noll *utan* att beräkna primitiv funktion!
- Funktion som ej är integrerbar: P6.2*.
- Analysens huvudsats (ihop med kedjeregeln, i vissa fall): P6.6, B6.12, P6.13*.
- Blandade integraler: P6.12adf.
- Effektivvärde av spänning: P6.15*.

Föreläsning 14. Generaliserad Riemannintegral; uppskattning av summor med integraler

Innehåll

- Avsnitt 6.5 och 6.7.
- Generaliserad integral = integral där **integrationsintervallet är obegränsat** och/eller **funktionen som integreras är obegränsad**.

  Definieras som gränsvärde av vanlig integral, t.ex. $\int_a^\infty f(x)\,dx = \lim_{\omega\to\infty}\int_a^\omega f(x)\,dx$, ifall f är Riemann-integrerbar på varje intervall $[a,\omega]$ med $a<\omega$.

  Integralen sägs vara **konvergent** om detta gränsvärde existerar, **divergent** annars.
- Samband mellan summor och integraler, och uppskattningar som kan göras med hjälp av detta.

Kommentarer

- När man redovisar uträkningen av en generaliserad integral kan man i princip skriva såhär:

$$\int_a^\infty f(x)\,dx = \lim_{\omega\to\infty}\int_a^\omega f(x)\,dx = \lim_{\omega\to\infty}\Big[F(x)\Big]_a^\omega = \lim_{\omega\to\infty}\big(F(\omega)-F(a)\big) = (\cdots).$$

  Men precis som för gränsvärden rekommenderar jag att man först räknar ut $\int_a^\omega$ utan att blanda in några gränsvärden, och *sedan* gör gränsövergången, dvs. att man skriver såhär istället:

$$\int_a^\omega f(x)\,dx = \Big[F(x)\Big]_a^\omega = F(\omega)-F(a) \to (\cdots), \quad \text{då } \omega\to\infty.$$

$$\textbf{Svar:}\ \int_a^\infty f(x)\,dx = (\cdots).$$

  Detta av samma anledningar som förut (man slipper skriva "$\lim_{\omega\to\infty}$" femtielva gånger, etc.), plus att det här finns ytterligare några fällor som man vill undvika att gå i.

  Exempel. Här är ett fel som man kan råka begå om man räknar med $\int_a^\infty$ istället för $\int_a^\omega$, i samband med partialbråksuppdelning:

$$\text{(Fel!)}\qquad \int_2^\infty \frac{2\,dx}{x^2-1} = \int_2^\infty \left(\frac{1}{x-1}-\frac{1}{x+1}\right)dx = \int_2^\infty \frac{dx}{x-1} - \int_2^\infty \frac{dx}{x+1}.$$

[Tillbaka till föreläsnings- och lektionsplan.]

Den rödmarkerade likheten stämmer inte! Integralen i vänsterledet är konvergent med värdet $\ln 3$ (se nedan), medan det som står i högerledet är differensen av två divergenta integraler, ett nonsensuttryck av typen ”$\infty - \infty$”. Det är tyvärr inte helt ovanligt att man får se sådana uträkningar när man rättar tentor, och i det läget kan det sedan fortsätta på lite olika sätt. Somliga drar felaktigt slutsatsen att integralen som de började med är divergent, och då blir det 0p direkt. Ett vanligare scenario är att man slår man ihop de två termerna igen innan man sätter in gränserna, så att det blir rätt svar till slut, men det blir åtminstone 1p avdrag ändå.

Om man istället hade räknat med en vanlig integral $\int_2^\omega$, så som jag rekommenderar, hade det naturligtvis varit fullkomligt harmlöst att dela upp integralen på ovanstående vis – om än en aning onödigt kanske. Jag själv skulle skriva något i den här stilen:

$$\begin{aligned}
\int_2^\omega \frac{2\,dx}{x^2-1} &= \int_2^\omega \left(\frac{1}{x-1} - \frac{1}{x+1}\right) dx = \Big[\ln|x-1| - \ln|x+1|\Big]_2^\omega \\
&= \ln(\omega-1) - \ln(\omega+1) - (\ln 1 - \ln 3) = \ln\frac{\omega-1}{\omega+1} + \ln 3 \\
&= \ln\frac{1-\frac{1}{\omega}}{1+\frac{1}{\omega}} + \ln 3 \to \ln\frac{1-0}{1+0} + \ln 3 = \ln 3 \quad \text{då } \omega \to \infty,
\end{aligned}$$

$$\text{dvs.}\quad \int_2^\infty \frac{2\,dx}{x^2-1} = \ln 3.$$

Notera att gränsvärden i samband med generaliserade integraler ska räknas ut precis som vi lärde oss i kapitlet om gränsvärden, dvs. genom omskrivningar tills man kommer till ett läge där standardgränsvärden och räknelagar kan användas. Man bör t.ex. **inte** helt plötsligt få för sig att skriva sådana här gräsligheter:

$$\text{(Fel!)} \qquad \ln(\omega-1) - \ln(\omega+1) \to \ln(\infty) - \ln(\infty) = 0.$$

Det råkar visserligen i detta fall vara sant att gränsvärdet är noll (som vi såg på det rätta sättet ovan), men en sådan där ”motivering” är ju rent nonsens, och kan i andra fall leda till helt fel slutsats.

- **Summor** är *begreppsmässigt* mycket enklare än **integraler** – man adderar ju bara ändligt många tal – men de kan vara betydligt svårare att räkna ut. För att ta ett ändå relativt enkelt exempel så är det en barnlek att beräkna integralen

  $$\int_0^x t^2\,dt = \tfrac{1}{3}x^3,$$

  men betrakta istället den liknande summan

  $$s_n = \sum_{k=1}^n k^2 = 1+4+9+16+25+\cdots+(n-1)^2+n^2.$$

  Den är inte lika enkel! Känner du till någon formel som ger dig s_n direkt, t.ex. för $n = 10^9$ utan att du behöver summera en miljard termer? Det är inte en aritmetisk summa, och inte heller en geometrisk summa, så det du har lärt dig i grundkursen är inte omedelbart tillämpligt. I själva verket ges summan av formeln

  $$s_n = \tfrac{1}{6}n(n+1)(2n+1) = \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n,$$

  som kan härledas på massor av olika sätt.[9] Det enklaste sättet, om man väl har formeln och bara vill visa att den stämmer, är att verifiera att subtraktion av successiva värden återskapar den följd

[9] Se t.ex. boken *Concrete Mathematics: A Foundation for Computer Science* av Graham, Knuth & Patashnik.

[Tillbaka till föreläsnings- och lektionsplan.]

$a_k = k^2$ som man har ackumulerat ihop:

$$\begin{aligned} s_1 &= \frac{1\cdot 2\cdot 3}{6} = 1^2 = a_1, \\ s_n - s_{n-1} &= \frac{n(n+1)(2n+1)}{6} - \frac{(n-1)\big((n-1)+1\big)\big(2(n-1)+1\big)}{6} \\ &= \frac{n}{6}\Big((n+1)(2n+1) - (n-1)(2n-1)\Big) \\ &= \frac{n}{6}\Big((2n^2+3n+1) - (2n^2-3n+1)\Big) = \frac{n}{6}\cdot 6n = n^2 = a_n, \qquad \text{för } n \geq 2. \end{aligned}$$

- Istället för det svåra problemet att beräkna summor exakt kan det ibland vara att föredra att *uppskatta* summors värde med hjälp av integraler som är enklare att räkna ut, och det är alltså detta som vi ska ägna oss åt här.

Lektion 15

- Undersökning av generaliserade integraler: P6.16ab, P6.17, B6.24abc, B6.25ab, B6.26cd.
- En integral som uppkommer i ett mekanikproblem: B6.33*.
- En integral relaterad till normalfördelning: B6.32*.

Lektion 16

- Summor och integraler: B6.16, B6.17, B6.19, P6.18.
- Och en svårare variant, där funktionen ifråga inte är monoton: P6.25*.

Föreläsning 15. Kompletteringar och repetition

[Tillbaka till föreläsnings- och lektionsplan.]

Facit till extrauppgiften på Lektion 2

(a) Falskt i allmänhet. Men sant ifall f är definierad godtyckligt nära a både till höger och till vänster, dvs. om det för varje $\delta > 0$ finns tal från definitionsmängden D_f både i intervallet $a-\delta < x < a$ och i intervallet $a < x < a+\delta$.

(b) Sant. Högergränsvärdet existerar och är lika med noll.

(c) Falskt. Vänstergränsvärdet existerar inte, ty definitionsmängden för $\sqrt{x}$ innehåller inga tal $x < 0$.

(d) Sant. Gränsvärdet existerar och är lika med noll; det sammanfaller med högergränsvärdet i detta fall, eftersom definitionsmängden är som den är. Detta är ett exempel på att påståendet i (a) kan vara falskt.

(e) Falskt. Man måste ju ta hänsyn till värdet $f(a)$ också!

(f) Falskt i allmänhet; jfr. (a) ovan. Vad som är sant är att f är kontinuerlig i punkten a om och endast om $\lim_{x\to a} f(x) = f(a)$ eller a är en isolerad punkt i D_f.

(g) Sant. Värdet $f(0) = \sqrt{0} = 0$ stämmer överens med gränsvärdet $\lim_{x\to 0} f(x) = \lim_{x\to 0} \sqrt{x} = 0$; jfr. (d) ovan. Detta är ett exempel på att påståendet i (f) kan vara falskt.

(h) Sant. Om det vore falskt skulle det ju motsäga satsen att alla elementära funktioner är kontinuerliga, eller hur?