

Om kurvers areal

Harald Hanche-Olsen

Institutt for matematiske fag
Norges teknisk-naturvitenskapelige universitet
NO-7491 Trondheim
hanche@math.ntnu.no

Tittelen på denne artikkelen virker nok som en åpenbar selvmotsigelse. Kurver har ingen bredde, og derfor ikke noe areal. Mange var nok både overrasket og sjokkert da man oppdaget at det eksisterer kurver med positivt areal.

Men først må jeg si noen ord om selve kurvebegrepet, og om hvorfor kurver (som regel) ikke har noe areal – eller rettere, hvorfor arealet av en kurve er null.

Kurver, deres lengde og areal

Vi tenker gjerne på en kurve i planet som en *delmengde* av planet, men det er ofte mer praktisk å betrakte parametriske kurver. En *parametrisk kurve* i planet er en kontinuerlig funksjon fra et intervall til planet. Vi skal betrakte kurver med lukket og begrenset parameterintervall, slik at kurven har et startpunkt og et endepunkt.

Vi kaster oss ut på dypt vann med en gang:

Setning 1 *Arealet av en kurve med endelig lengde er null.*

Denne setningen gir knapt mening om vi ikke vet hvordan lengden av en kurve er definert, men vi kommer til det. Setningen vil vise seg å være en direkte følge av et litt mer komplisert resultat:

Lemma 2 *Arealet av en ε -omegn om en brukket linje med lengde L er ikke større enn $2\varepsilon L + \pi\varepsilon^2$.*

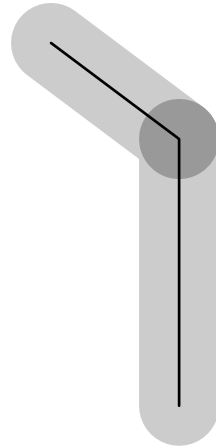
En brukket linje er en kurve som er satt sammen av et endelig antall rette linjestykker. I dette tilfellet er det ikke noe problem å definere lengden: Lengden av et linjestykke er avstanden mellom endepunktene, og lengden av et brukket linjestykke er summen av delenes lengder.

Med ε -omegnen om en kurve mener vi mengder av punkter som har avstand $\leq \varepsilon$ til minst ett punkt på kurven. Et raskt blikk på figur 1 viser at ε -omegnen om et linjestykke med lengde L har areal $2\varepsilon L + \pi\varepsilon^2$: Omegnen består nemlig av et $L \times 2\varepsilon$ -rektangel sammen med to halvsirkelskiver med radius ε .



Figur 1

Setter vi nå sammen to linjestykker som i figur 2 med lengder L_1 og L_2 , kan vi skrive arealet av ε -omegnen om unionen av de to som



Figur 2

$$\begin{aligned} A &\leq 2\varepsilon L_1 + \pi\varepsilon^2 + 2\varepsilon L_2 + \pi\varepsilon^2 - \pi\varepsilon^2 \\ &= 2\varepsilon(L_1 + L_2) + \pi\varepsilon^2, \end{aligned}$$

der vi har trukket fra arealet av en sirkelskive fordi dette arealet er regnet med to ganger. (Hvis vinkelen mellom de to ikke er 180° , blir ulikheten ekte.) Dette viser Lemma 2 for sammensetningen av to linjestykker. En åpenbar induksjon etter samme prinsipper viser resultatet også for generelle brukne linjestykker, så Lemma 2 er bevist.

Vi kan nå også bevise Setning 1: Anta kurven γ , med parameterintervall $[0, 1]$, er gitt. Vi bruker det faktum at γ er *uniformt kontinuert*: Gitt $\varepsilon > 0$ finnes en $\delta > 0$ slik at hver gang $|s - t| < \delta$ er $|\gamma(s) - \gamma(t)| < \varepsilon$. Så vi deler opp intervallet $[0, 1]$ i småbiter: La $0 = t_0 < t_1 < \dots < t_n = 1$ med $|t_j - t_{j-1}| < \delta$ for $j = 1, \dots, n$. Nå ligger hele kurven γ innenfor en ε -omegn om det brukne linjestykket med knekkpunkter $\gamma(t_0), \gamma(t_1), \dots, \gamma(t_n)$. Så arealet av γ kan ikke være større enn arealet av denne ε -omegnen, eller – i følge Lemma 2 –

$$\text{areal}(\gamma) \leq \varepsilon \sum_{j=1}^n |\gamma(t_j) - \gamma(t_{j-1})| + \pi\varepsilon^2.$$

Hvis vi nå er villig til å tro på at den rette linje virkelig er den korteste vei mellom to punkter, så må lengden L av kurven γ være minst så stor som summen ovenfor, slik at

$$\text{areal}(\gamma) \leq \varepsilon L + \pi\varepsilon^2.$$

Siden $\varepsilon > 0$ var vilkårlig, kan vi nå la $\varepsilon \rightarrow 0$ for å fullføre beviset for Setning 1.

Vi har ennå ikke *definert* lengden av en kurve. Men vi har brukt ulikheten

$$\sum_{j=1}^n |\gamma(t_j) - \gamma(t_{j-1})| \leq L.$$

Faktisk defineres lengden L av γ som minste øvre grense for summen ovenfor, der $t_0 < t_1 < \dots < t_n$ gjennomløper alle partisjoner av parameterintervallet til kurven. Slik blir lengden *additiv* i den forstand at lengden til en kurve sammensatt av flere deler er summen av lengdene til delene, og lengden til et rett linjestykke er avstanden mellom endepunktene. Endelig gir dette det korteste mulige lengdemål som tilfredsstiller disse to kravene og samtidig er slik at en rett linje alltid er korteste vei mellom to punkter.

Skal vi finne en kurve med positivt areal, må kurven altså være uendelig lang. Det virker kanskje ikke som noe stort problem, men uendelig lengde er langt fra nok. Nå er det ikke rom i denne artikkelen for å gi en like presis definisjon av areal som vi har gitt for lengde, så vi må støtte oss på en mer naiv tolkning av arealbegrepet. Arealet må være *additivt* i den forstand at $\text{areal}(A \cup B) = \text{areal}(A) + \text{areal}(B)$ om A og B er disjunkte. Det viser seg at det også må være *tellbart additivt* i den forstand at en tilsvarende addisjonsformel også gjelder for tellbart uendelige unioner av parvis disjunkte mengder. Nå leder dette til en motsigelse om vi antar at arealet er definert for alle delmengder av planet, så vi må begrense oss til de såkalt *målbare* mengdene. Jeg vil nøye meg med å påstå at alle mengder vi betrakter i denne artikkelen er målbare. (Teorien for målbare mengder er en del av mål- og integrasjonsteorien.)

Georg Cantor (1845–1918)

Cantor var født i St. Petersburg, og vokste opp der inntil familien flyttet til Tyskland da han var 11 år. Han studerte under Weierstrass, Kummer og Kronecker i Berlin, og ble siden ansatt ved universitetet i Halle.

Cantor kan kanskje betraktes som mengdelærens grunnlegger. I 1877 viste han [1] at det finnes en en-til-en avbildning, eller bijeksjon, av enhetsintervallet $[0, 1]$ på enhetskvadratet $[0, 1] \times [0, 1]$.

Cantors bevis bygger på teorien for *kjedebryk*. For å forklare kjedebryk kan vi ta utgangspunkt i *Euklids algoritme* for å bestemme forholdet mellom to linjestykker a og b : Først bruker vi b til å måle a . La oss si det går opp n hele ganger med en rest r , eller $a = nb + r$ i moderne notasjon, der n er et heltall og $0 \leq r < b$. Hvis $r > 0$ bruker vi nå r til å måle b . Prosedyren gjentas inntil vi ikke har noen rest igjen. Hvis så ikke skjer, er a og b inkommensurable.

La oss illustrere med et eksempel: Anta a er 157 enheter og b er 68 enheter. Vi finner $157 = 2 \cdot 68 + 21$, eller mer passende $157/68 = 2 + 21/68$. Neste trinn i Euklids algoritme består i å snu den siste brøken på hodet og gjøre det samme igjen: $68/21 = 3 + 5/21$. Og igjen: $21/5 = 4 + 1/5$. Disse resultatene brukes så til å bygge opp en kjedebryk:

$$\frac{157}{68} = 2 + \frac{21}{68} = 2 + \frac{1}{3 + \frac{5}{21}} = 2 + \frac{1}{3 + \frac{1}{4 + \frac{1}{5}}}$$

Prosessen stopper her, for om vi snur den nederste brøken opp ned får vi $\frac{5}{1}$, som går opp. Ett-tallet i denne siste brøken er for øvrig største felles divisor for utgangstallene 157 og 68. Det er ikke så vanskelig å vise at Euklids algoritme alltid terminerer, og altså gir en endelig kjedebrøk, så lenge vi starter med et rasjonalt tall.

Beviset til Cantor bygger derimot på at ethvert irrasjonalt tall i $[0, 1]$ kan skrives entydig som en *uendelig* kjedebrøk

$$t = \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_3 + \dots}}}$$

med positive, heltallige koeffisienter n_1, n_2, \dots . Disse koeffisientene kan man enkelt finne ved hjelp av Euklids algoritme: Skriv $t = 1/(n_1 + r_1)$ med $n_1 \geq 1$ heltallig og $0 < r_1 < 1$, gjenta prosedyren med r_1 , og så videre.

Et underholdende eksempel er det gyldne snitt $G = \frac{1}{2}(\sqrt{5} - 1)$. Siden $G^2 + G = 1$ kan vi skrive $G = 1/(1 + G)$. Etter Euklids algoritme skal vi nå gjøre det samme med G i nevneren, og finner i grensen en kjedebrøk for G der alle koeffisientene n_k er 1.

Det er nå likefrem å gi en bijeksjon mellom de irrasjonale tallene i $[0, 1]$ og punktene i $[0, 1] \times [0, 1]$ med irrasjonale koordinater, simpelthen ved å skrive opp to nye kjedebrøker for x og y , med koeffisienter n_1, n_3, \dots for x og n_2, n_4, \dots for y . Dette gir så den ønskede bijeksjonen $t \mapsto (x, y)$. Vi kan også enkelt finne en bijeksjon mellom de irrasjonale tallene i $[0, 1]$ og hele $[0, 1]$: Plukk en tellbar uendelig mengde irrasjonale tall i $[0, 1]$, og avbild denne en-til-en på unionen av seg selv og de rasjonale tallene i $[0, 1]$. (Dette minner om historien om Hilberts hotell – se også [3].) La resten av de irrasjonale tallene i $[0, 1]$ avbildes på seg selv. Dermed gir konstruksjonen foran en bijeksjon mellom enhetsintervallet og enhetskvadratet.¹

Samme bevisteknikk viser at $[0, 1]$ kan avbildes bijektivt også på enhetskuben $[0, 1]^3$, og faktisk på enhver endelig hyperkube $[0, 1]^n$.

I et brev til sin venn og kollega Dedekind skrev Cantor følgende om dette resultatet: *Jeg ser det, men jeg tror det ikke!*

Cantor var utmerket godt klar over, og bemerket i [1], at hans bijeksjoner ikke var *kontinuerlige*. Kort tid etter at Cantor hadde publisert sitt resultat viste da også Netto [5] at ingen bijeksjon fra enhetsintervallet på enhetskvadratet *kan* være kontinuerlig. For i så fall måtte også den inverse avbildningen være kontinuerlig, men det er åpenbare topologiske forskjeller mellom enhetsintervallet og enhetskvadratet som gjør dette umulig. For eksempel kan vi dele enhetsintervallet i to komponenter ved å fjerne ett punkt, og det går ikke med enhetskvadratet. Netto påpekte at dette argumentet også kan utvides til å vise at det ikke finnes en kontinuerlig bijeksjon mellom to kuber $[0, 1]^m$ og $[0, 1]^n$ med $m \neq n$. For eksempel kan $[0, 1]^2$ deles i to komponenter med en kurve, mens $[0, 1]^n$ ikke kan det, når $n \geq 3$.

¹I moderne litteratur finner man oftere et bevis som fordeler tallsifrene i t på tallsifrene i x og y , men dette leder til en vanskelighet med de tallene som ikke har en entydig desimalutvikling, og denne vanskeligheten er litt mer vrien å løse.

Dette er starten på det vi kan kalle topologisk dimensjonsteori, men det vil føre for langt å diskutere den her. Vi skal nå se hva som kan skje dersom vi beholder kravet om kontinuitet, men ikke insisterer på at avbildningen skal være en-til-en.

Guiseppe Peano (1858–1932)

Peano studerte ved universitetet i Torino, og ble siden professor der. Han viste blant annet eksistensen av løsninger til differensialligningen $y' = f(x, y)$ der f er kontinuerlig, og påpekte at løsningen til initialverdi-problemet $y' = 3y^{2/3}$, $y(0) = 0$ ikke er entydig. Han er vel kjent for sin aksiomatisering av aritmetikken.

I 1890 viste han [7] at det finnes en kurve som fyller hele enhetskvadratet.

Peanos konstruksjon minner om teknikken for å etablere en bijeksjon mellom $[0, 1]$ og $[0, 1] \times [0, 1]$, ved bruk av desimalutvikling i stedet for kjedebrøk. Men i stedet uttrykker han tall i *tretallsystemet*, altså med sifrene 0, 1 og 2. Heller enn å gjenta Peanos beskrivelse av konstruksjonen her, illustrerer vi den med et eksempel: La $t = 0,02120112010012122111$ (uttrykt i tretallsystemet) og skriv opp følgende tabell:

$$\begin{array}{cccccccccccccccc} t = & 0, & 0 & 2 & 1 & 2 & 0 & \underline{1} & 1 & 2 & 0 & \underline{1} & 0 & 0 & 1 & 2 & 1 & 2 & 2 & \underline{1} & 1 & \underline{1} \\ & & & & & & & \cdot & & & & \cdot & & & & & & & & & & & \cdot \\ x = & 0, & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 2 & 1 & & & & & & & & & & & \end{array}$$

I første rad i tabellen har vi skrevet opp sifrene i t og understreket alle enere i *like* posisjon (2, 4, 6 etc.). Annen rad i tabellen inneholder enten en prikk (\cdot) eller en strek ($-$) i hver *odde* posisjon (1, 3, 5 etc). Hvilken av de to avhenger av antallet understrekede enere til venstre: En prikk om antall understrekede enere er et *liketall*, og en strek om antall enere er et *oddetall*. Sifrene i x kopieres direkte fra sifrene i t med odde posisjon, der hvor det står en prikk. Når det står en strek, komplementeres sifferet etter regelen $0 \mapsto 2$, $1 \mapsto 1$, $2 \mapsto 0$.

Sifrene i y kopieres tilsvarende fra sifrene i t med like posisjon, men denne gang direkte eller komplementert avhengig av antall enere i odde posisjoner i t . Tabellen nedenfor gir utregningen av y for vårt eksempel.

$$\begin{array}{cccccccccccccccc} t = & 0, & 0 & 2 & \underline{1} & 2 & 0 & \underline{1} & \underline{1} & 2 & 0 & \underline{1} & 0 & 0 & \underline{1} & 2 & \underline{1} & 2 & 2 & \underline{1} & \underline{1} & \underline{1} \\ & & & & \cdot & - & - & \cdot & \cdot & \cdot & - & \cdot & \cdot & - & \cdot & \cdot & - & & & & & & \\ y = & 0, & 2 & 0 & 1 & 2 & 1 & 0 & 0 & 2 & 1 & 1 & & & & & & & & & & & \end{array}$$

Det geniale med Peanos konstruksjon er at den virkelig gir en veldefinert avbildning fra $[0, 1]$. Vi illustrerer dette med et eksempel (husk at vi hele tiden arbeider i tretallsystemet): $t = 0,012 = 0,011\bar{2}$.²

Representasjonen $t = 0,012$ gir $x = 0,00\bar{2}$ og $y = 0,1$.

Representasjonen $t = 0,011\bar{2}$ gir $x = 0,01$ og $y = 0,1$.

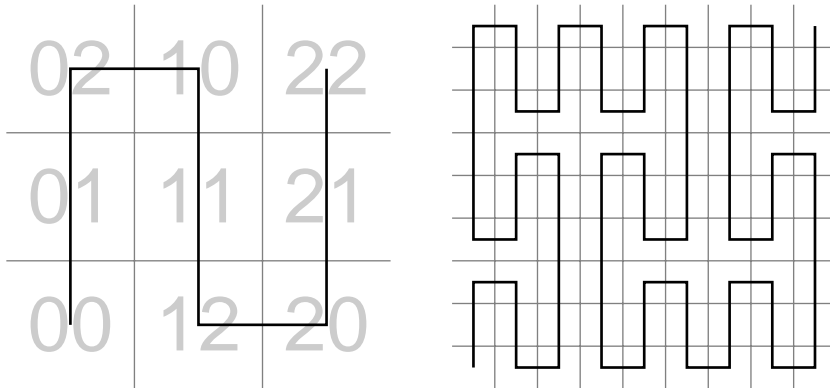
Disse resultatene er identiske, og det er da heller ikke vanskelig – om enn noe arbeidskrevende – å overbevise seg om at tilsvarende skjer for alle andre $t \in [0, 1]$ som har to representasjoner i tretallsystemet – de trinært rasjonale tall.

²Vi setter strek over en gruppe av siffer for å vise at den gjentas i det uendelige.

Når vi kjenner de første $2n$ sifrene i t , kjenner vi også de første n sifrene i x og y . Det er ikke vanskelig å vise at dette, sammen med det faktum at $t \mapsto (x, y)$ er veldefinert, impliserer at avbildningen er kontinuerlig.

Den er selvsagt ikke bijektiv, siden dette ville stride mot Nettos resultat. Som en øvelse oppfordres leseren til å finne alle $t \in [0, 1]$ med $(x(t), y(t)) = (0, 0, 1, 1)$.

Første trinn mot en geometrisk forståelse av Peanos kurve er å merke seg at vi kan plassere $(x(t), y(t))$ i ett av ni kvadrater ved å se på de første to sifrene etter kommaet i representasjonen av t . Spesielt tilbringer kurven den første niendedelen av parameterintervallet i kvadratet merket 00, den neste i 01, og så videre, som antydnet til venstre i figur 3. Legg spesielt merke til at kvadratene i den midterste søylen (med en ener som første siffer) besøkes ovenfra og ned, mens de to andre søylene (med 0 og 2 som første siffer) besøkes nedenfra og opp.



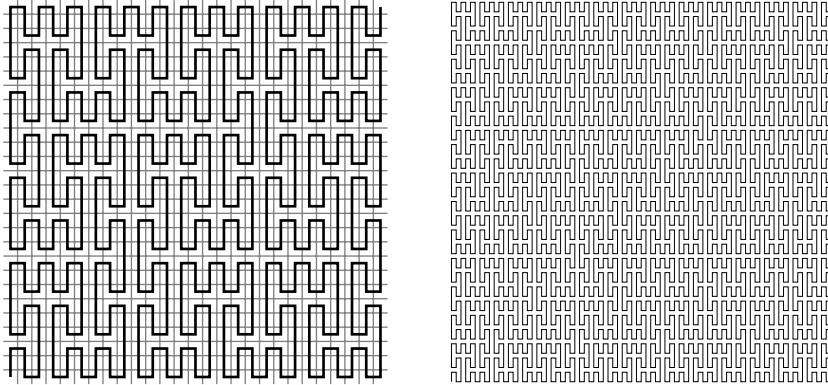
Figur 3: Første og annen tilnærming til Peano-kurven.

Neste trinn er å studere hva som skjer innenfor hvert underkvadrat. Det viser seg at hvert underkvadrat er som det store, men av og til med x -koordinaten invertert, av og til med y -koordinaten invertert, og av og til begge deler. Mer presist, for $t \in [0, 1]$:

$$\begin{aligned}
 (x(\frac{1}{9}(0+t)), y(\frac{1}{9}(0+t))) &= (\frac{1}{3}(0+x(t)), \frac{1}{3}(0+y(t))), \\
 (x(\frac{1}{9}(1+t)), y(\frac{1}{9}(1+t))) &= (\frac{1}{3}(1+x(t)), \frac{1}{3}(1-y(t))), \\
 (x(\frac{1}{9}(2+t)), y(\frac{1}{9}(2+t))) &= (\frac{1}{3}(2+x(t)), \frac{1}{3}(0+y(t))), \\
 (x(\frac{1}{9}(3+t)), y(\frac{1}{9}(3+t))) &= (\frac{1}{3}(3-x(t)), \frac{1}{3}(1+y(t))), \\
 (x(\frac{1}{9}(4+t)), y(\frac{1}{9}(4+t))) &= (\frac{1}{3}(2-x(t)), \frac{1}{3}(2-y(t))), \\
 (x(\frac{1}{9}(5+t)), y(\frac{1}{9}(5+t))) &= (\frac{1}{3}(1-x(t)), \frac{1}{3}(1+y(t))), \\
 (x(\frac{1}{9}(6+t)), y(\frac{1}{9}(6+t))) &= (\frac{1}{3}(0+x(t)), \frac{1}{3}(2+y(t))), \\
 (x(\frac{1}{9}(7+t)), y(\frac{1}{9}(7+t))) &= (\frac{1}{3}(1+x(t)), \frac{1}{3}(3-y(t))), \\
 (x(\frac{1}{9}(8+t)), y(\frac{1}{9}(8+t))) &= (\frac{1}{3}(2+x(t)), \frac{1}{3}(2+y(t))).
 \end{aligned}$$

Kurven til høyre i figur 3 er laget ved å avbilde den venstre kurven inn i hver av de ni underkvadratene etter formlene over og sette sammen kurvebitene. Neste

tilnærming er å dele hver av de 81 småkvadratene inn i ni kvadrater hver, som gir de 729 kvadratene og kurven til venstre i figur 4. Slik kan man fortsette å lage stadig nye tilnærminger til Peano-kurven: I n -te tilnærming er enhetsintervallet delt inn i 9^n like store delintervaller, som avbildes inn i $3^n \times 3^n$ delkvadrater. Disse tilnærmingene konvergerer uniformt mot Peano-kurven. Dette gir et nytt bevis for kontinuitet.



Figur 4: Tredje og fjerde tilnærming til Peano-kurven.

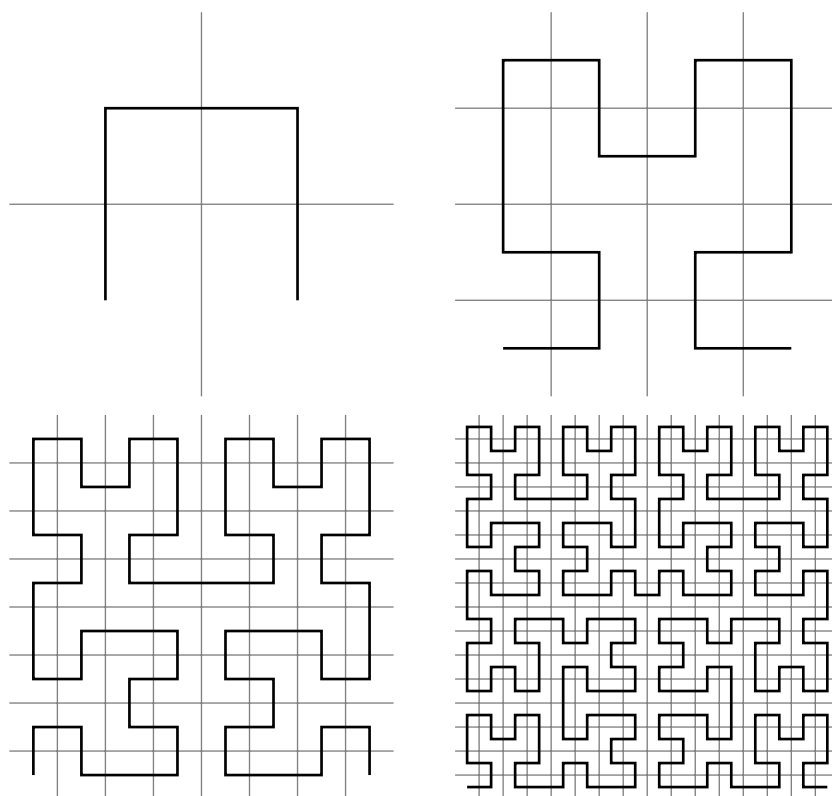
David Hilbert (1862–1943)

Hilbert studerte i Königsberg og ble siden professor der, men tilbrakte mesteparten av sin karriere som professor i Göttingen. Hans bidrag til matematikken var dype og spente svært vidt, og hans 23 problemer som han presenterte for matematikernes verdenskongress i Paris hadde en stor innflytelse på matematikken i det tjuende århundre. Hilberts bidrag til vår lille historie er knapt nok som en fotnote å regne.

Han må nok ha resonert omtrent som over før han fant en enklere geometrisk konstruksjon av en kurve som fyller enhetskvadratet. Hans konstruksjon i [4] er den som i dag vanligvis assosieres med Peanos navn.

Hilbert delte inn enhetskvadratet i fire deler og lot første tilnærming besøke de fire delene som vist øverst til venstre i figur 5. Deretter delte han inn hver av de fire småkvadratene i fire og satte inn kopier av første tilnærming, som vist øverst til høyre i figuren. Ved å fortsette slik og gå til grensen fikk han en kurve med de samme egenskaper som Peanos kurve.

Hilberts kurve kan beskrives ved å uttrykke parameteren i totallsystemet og gjøre manipulasjoner tilsvarende de Peano gjorde, men bokholderiet blir adskillig mer komplisert fordi to av småkvadratene i Hilberts konstruksjon må ha x - og y -koordinatene byttet om.



Figur 5: Første til fjerde tilnærming til Hilbert-kurven.

William F. Osgood (1864–1943)

Osgood studerte først ved Harvard, senere i Göttingen og Erlangen under Kleins veiledning. Senere returnerte han til Harvard hvor han tilbragte resten av sin karriere. Han arbeidet med funksjonsteori, differensialligninger og variasjonsregning og romfylende kurver.

Peano-kurven og Hilberts variant kan virke spennende nok, men fordi de fyller hele enhetskvadratet og ikke er en-til-en tilfredsstillende kanskje ikke *alle* våre forventninger om hvordan en kurve skal se ut.

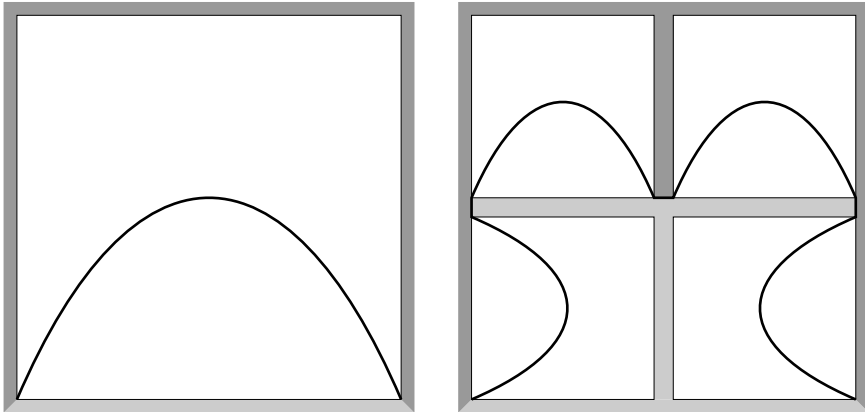
En *enkel* kurve er en (parametrisk) kurve som også er en-til-en, med mulig unntak for endepunktene, som kan falle sammen. (Hvis de gjør det kaller vi gjerne kurven en *Jordan-kurve* til ære for Camille Jordan, som først påpekte behovet for et bevis for at slike kurver deler planet i to. Ellers kalles kurven en Jordan-bue.)

Nettos resultat sier at ingen enkel kurve kan fylle enhetskvadratet. Det vakte derfor berettiget oppsikt da Osgood [6] i 1903 konstruerte en enkel kurve med positivt areal.

Osgoods konstruksjon er en variant av Peanos kurve, men hvor småkvadratene holdes adskilt for å hindre kurven i å besøke samme punkt flere ganger. Vi skal her anvende Osgoods idé på Hilberts kurve i stedet.

Hilberts kurve starter i $(0, 0)$ og ender i $(1, 0)$. Vår variant skal gjøre det samme. Vi antyder dette ved en «nullte tilnærming» som til venstre i figur 6. Tenk på buen i figuren som en vilkårlig valgt kurve i enhetskvadratet med endepunktene fiksert.

Den endelige kurven vil dele kvadratet i to: En del mot nederste side i kvadratet, og resten mot de tre andre sidene. Vi tenker oss at kvadratet er omgitt av tørt land på de tre sidene og vann på den nederste siden (også antydnet i figuren).



Figur 6: Nullte og første tilnærming til Hilbert–Osgood-kurven.

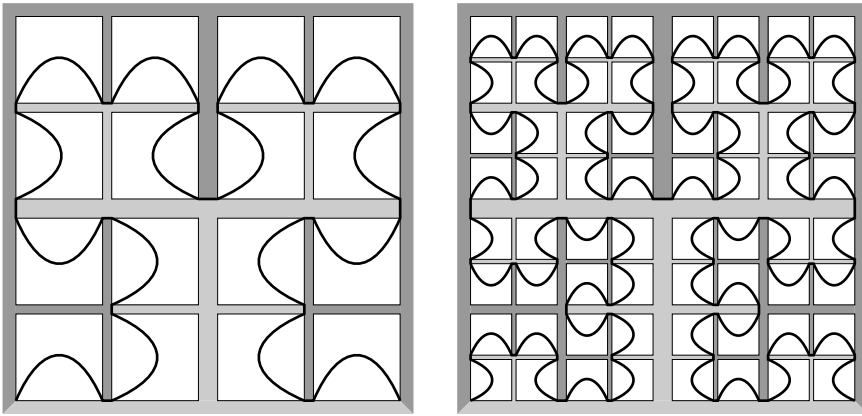
Vi deler nå kvadratet i fire deler, men skiller delene fra hverandre med en T-formet kanal og et I-formet dike som til høyre i figur 6. Hver av de fire delene er nå omgitt av land på tre sider og vann på den fjerde, eller omvendt. Vi plasserer en kopi av figuren til venstre i hver av de fire delene som vist, og forbinder kopiene med tre korte linjestykker langs grensene mellom land og vann. Slik får vi en kurve med sju deler. Vi delere parameterintervallet $[0, 1]$ inn i sju like store deler, og foreskriver at disse delene skal avbildes til de sju kurvene i figuren. Spesielt vil segment nummer 2, 4 og 6 avbildes til de tre linjestykkene omtalt ovenfor, og vi foreskriver at alle fremtidige tilnærminger til kurven, og derfor også grensen, skal ha samme verdier på disse tre segmentene.

I neste runde gjør vi det samme i hver av de fire småkvadratene – se figur 7 til venstre. Denne tilnærmingen består nå av $3 + 4 \cdot 7 = 31$ deler, hvorav $4 \cdot 3 = 12$ nye rettlinjede segmenter langs de nye grensene mellom land og vann, som skal fikseres i alle fremtidige tilnærminger.

Vi fortsetter slik i det uendelige og får en enkel kurve i grensen.

Hvis vi tenker oss at alle dikene og kanalene vi har konstruert underveis er åpne mengder, så vil alle punkter i kvadratet falle i en av tre disjunkte kategorier: Land, vann og punkter på kurven. Vi kan lett regne ut arealet av land og vann, og trekker vi dette fra arealet av enhetskvadratet, altså 1, står vi nødvendigvis igjen med arealet av kurven.

La oss for eksempel si at dikene og kanalene vi gravde i første tilnærming hadde bredde $\delta < 1$. Tilsammen har de areal $\varepsilon = 2\delta - \delta^2$. De fire kvadratene som er igjen har da samlet areal $1 - \varepsilon$. Hvis neste trinn er en geometrisk reduksjon av det første, slik at vi graver diker og kanaler med bredde δ ganger bredden av hvert kvadrat, vil de nye dikene og kanalene få samlet areal $(1 - \varepsilon)\varepsilon$, og de 16 småkvadratene



Figur 7: Annen og tredje tilnærming til Hilbert–Osgood-kurven.

som er igjen har samlet areal $(1 - \varepsilon)^2$. I n -te iterasjon graver vi diker og kanaler med samlet areal $(1 - \varepsilon)^{n-1}\varepsilon$ og står igjen med 4^n småkvadrater med samlet areal $(1 - \varepsilon)^n$. Det totale arealet av land og vann blir i grensen

$$\sum_{n=0}^{\infty} (1 - \varepsilon)^n \varepsilon = 1,$$

så kurven får areal 0 som forventet.

Men dette er ikke eneste mulighet: I hvert trinn av prosessen kan vi bygge diker og grave kanaler så smale eller brede vi måtte ønske, så lenge det er plass. Det betyr at det samlede arealet ε_n av diker og kanaler i iterasjon nummer n kan velges fritt, så lenge $\varepsilon_n > 0$ og $\sum_n \varepsilon_n < 1$. Spesielt kan det totale arealet $\sum_{n=1}^{\infty} \varepsilon_n$ av land og vann velges så lite vi vil, og kurven kan dermed få areal så nær 1 vi måtte ønske.

Avsluttende bemerkninger

Det er verdt å merke seg at Hilbert–Osgood-kurven er en grense av kurver som hver har endelig lengde, og derfor areal null. Så hvorfor må ikke arealet av grensen være like grensen av arealene, og altså null? I avsnittet foran har jeg jo nettopp gjort dette med diker og kanaler. Men det er en viktig forskjell. I tilfellet med dikene og kanalene har vi en *voksende* følge av områder i planet: Dikene og kanalene på trinn $n + 1$ er ikke annet enn diker og kanaler fra trinn n , med et antall nye diker og kanaler lagt til. Og «grensen» er ikke annet enn unionen av denne voksende følgen av områder. Men når jeg tar grensen av kurver tar vi ikke unionen av de tilnærmede kurvene. I stedet har vi $\gamma(t) = \lim_{n \rightarrow \infty} \gamma_n(t)$ for hver t , og denne grenseoperasjonen gir mange punkter som ikke ligger på noen av de opprinnelige kurvene.

En (noe tvilsom) analogi er om vi betrakter delmengdene $E_n \subset [0, 1]$ gitt ved $E_n = \{2^{-n}k : k = 0, 1, \dots, 2^n\}$. Unionen av disse delmengdene er de dyadisk rasjonale tallene, som målteorien forteller oss har mål null, men hvis vi betrakter alle mulige *grenser* av følger (x_n) med $x_n \in E_n$, får vi hele $[0, 1]$.

En konklusjon synes å være at målteori og grenser går dårlig sammen. Så er det kanskje en desto større overraskelse at det finnes en kurve i enhetskvadratet som kan brukes til å måle arealet av enhver delmengde av kvadratet! (Innen rimelighetens grenser – metoden virker på *Borel-målbare* mengder.) Det er en viktig forskjell fra de kurvene vi har betraktet her: For denne typen målinger trenger vi en kurve parametrisert med det halvåpne intervallet $[0, 1)$. Påstanden er altså at arealet (strengt tatt målet) av en (Borel) delmengde A av kvadratet er det endimensjonale målet av mengden $\{t \in [0, 1) : \gamma(t) \in A\}$, bare kurven $\gamma: [0, 1) \rightarrow [0, 1] \times [0, 1]$ er valgt smart nok. Det er åpenbart at en slik kurve må ha positivt areal i alle sine deler (betrakt $\{\gamma(t) : t_0 < t < t_1\}$), men samtidig kan den ikke unngå noen åpen delmengde av kvadratet, slik Osgood-kurven gjør. Se [2] for detaljene, som krever en mye dypere forståelse for målteori enn jeg har lagt til grunn her.

Referanser

- [1] G. Cantor, Ein Beitrag zur Mannigfaltigkeitslehre. *J. reine angew. Math.* **84**, 242–258 (1878).
- [2] Piotr Hajlasz, Pawł Strzelecki, How to Measure Volume With a Thread. *Amer. Math. Monthly* **112**, 176–179 (2005).
- [3] Vagn Lundsgaard Hansen, Rundt om uendeligheden. *Normat* **52**, 105–117 (2004).
- [4] D. Hilbert, Über die stetige Abbildung einer Linie auf ein Flächenstück. *Math. Annalen* **38** 459–460 (1891).
- [5] E. Netto, Beitrag zur Mannigfaltigkeitslehre. *J. reine angew. Math.* **86**, 263–268 (1879).
- [6] W. F. Osgood, A Jordan curve of positive area. *Trans. Amer. Math. Soc.* **4** 107–112 (1903).
- [7] G. Peano, Sur une courbe, qui remplit toute une aire plane. *Math. Annalen*, **36** 157–160 (1890).

De biografiske omtalene av Cantor, Hilbert og Osgood er basert på materiale fra the History of Mathematics Archive på <http://www-groups.dcs.st-and.ac.uk/~history/>.

Det var en feil i den publiserte artikkelen som er rettet her: Jeg hadde utelatt faktoren 2 foran εL i Lemma 2. Takk til Hans-Georg Killingbergtrø for at han påpekte feilen.