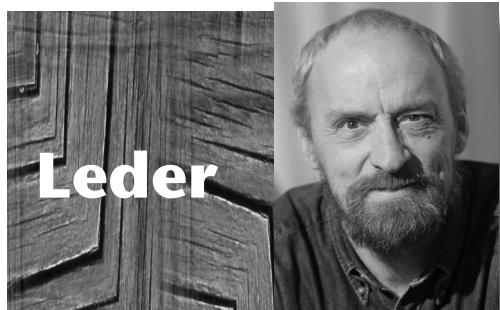


mat

MATILDE

TEMA:
Biologi og matematik





Leder

Af Freddy Bugge Christiansen

Temaet i dette nummer af Matilde er biologi, eller rettere sagt nogle eksempler på biologisk indsigt, som er opnået gennem anvendelse af matematisk argumentation og matematiske resultater. Matematikere er gennem tiden blevet inspireret af biologiske problemstillinger og har derved produceret matematiske objekter, som er blevet centrale dele af fagets kultur. Det bedste eksempel er nok Fibonaccis overvejelser over kaninavl. Resultatet var Fibonaccitalfølgen - det første eksempel på en populationsmodel der resulterer i eksponentiel vækst, oven i købet med en gylden vækstrate. Ud over at bidrage til matematikstuderendes imponerende resultater ved sessionens intelligensprøver, så vedbliver Fibonaccitalfølgen at glæde og begejstre matematikere. Den umiddelbare effekt af Fibonaccis arbejde fandtes dog ikke inden for studiet af levende organismer. Hans øvrige arbejde havde en mere blivende og generel effekt på Europas hverdag.

Matematikeres overvejelser inden for biologi har også rystet verden. Mest kendt er vel Thomas Robert Malthus' indflydelse på Darwins formulering af principippet om naturlig selektion i evolutionslæren. Mindre kendt er det vel, at Malthus med sit essay om ubegrænset vækst rystede sin samtid, idet han konkluderede:

The power of population is so superior to the power of the earth to produce subsistence for man, that premature death must in some shape or other visit the human race. The vices of mankind are active and able ministers of depopulation. They are the precursors in the great army of destruction, and often finish the dreadful work themselves.

But should they fail in this war of extermination, sickly seasons, epidemics, pestilence, and plague advance in terrific array, and sweep off their thousands and tens of thousands. Should success be still incomplete, gigantic inevitable famine stalks in the rear, and with one mighty blow levels the population with the food of the world.

Thomas Robert Malthus 1798

Ikke underligt at han udgav sit essay anonymt. Mere underligt er det, at hver generation synes overrasket over hvad eksponentiel vækst indebærer for menneskepopulationer.

Ronald Aylmer Fisher (1890–1962) er den næste store matematiker indenfor evolutionslæren. Han kastede sig over studiet af mendelgenetik og udviklede i 1918 det moderne grundlag for den biometrisk arv - en da klassisk disciplin baseret på statistisk beskrivelse af lighed mellem beslægtede. Studiet af genetik krævede statistik, som Fisher så udviklede, og Anders Hald beskrev ham i 1998 som "a genius who almost single-handedly created the foundations for modern statistical science." Selvom matematikere ikke samstemmende inkluderer statistik i deres fag, så findes der dog et forhold, som kan nævnes selv ved bedre middage.

En israelsk kollega sagde engang: "I began my studies in economics, but to distance myself from him, I switched to population genetics just to find that here he is also a great man." Han henviste til sin vejleder Samuel Karlin (1924–2007), ved en konference arrangeret af Karlin. Blandt matematikere er han kendt for sine bidrag til total positivitet, og selvom han har haft stor indflydelse på populationsgenetikken, er hans største bidrag til biologi sikkert det sandsynlighedsteoretiske grundlag for BLAST (Basic Local Alignment Search Tool) - en algoritme, som er central ved sammenligninger af DNA-sekvenser og dermed inden for genetik, molekylær biologi og bioinformatik.

Biologi har bidraget til matematik, men hvor dybt og betydningsfuldt dette bidrag har været, skal en biolog nok vare sig for at kommentere på. Men at matematikere og matematik bidrager og har bidraget til biologi, er hævet over enhver tvivl. Vil man indskrives på et biologiststudium i Danmark, kræver det gymnasie-matematik på højt niveau, fordi matematisk formulerede biologiske modeller er centrale i mange biologiske fags teoretiske grundlag. Evolutionslæren vejer tungt, navnlig på grund af populationsgenetikken. En anden linie tilbage til Malthus er populationsøkologien og epidemiologien, hvis fagområder han diskuterer i citatet - faktisk var forståelsen af pestepidemiernes regelmæssige tilbagekomst en motivation i hans arbejde. Et andet område med matematisk-fysiske modeller er fysiologien. De nævnte fagområder er dækket af indlæg i dette nummer af Matilde, og læg specielt mærke til, at et af indlæggene behandler matematisk-teoretiske bånd mellem evolutionslæren og fysiologien. Biologer kan altså også være med i matematiske kerneaktiviteter - men inden for biologi selvfølgelig.

**MATILDE – NYHEDSBREV FOR
DANSK MATEMATISK FORENING
medlem af
EUROPEAN MATHEMATICAL SOCIETY**

Nummer 34– April 2008

Redaktion:

**Bent Ørsted, Aau
(ansvarshavende og
TEMAREDAKTØR)
Freddy Bugge Christiansen
Aau (temaredaktør)**

**Carsten Lunde Petersen,
Ruc
Jørn Børling Olsson, Ku
Poul Hjorth, Dtu
Mikael Rørdam, Ku
Carl Winsløw, Ku**

ADRESSE:

**MATILDE
INSTITUT FOR MATEMATISKE FAG
KØBENHAVNS UNIVERSITET
UNIVERSITETSPARKEN 5
2100 KØBENHAVN Ø**

FAX: 3532 0704

e-post:

matilde@mathematics.dk

URL:

www.matilde.mathematics.dk

ISSN: 1399-5901

**Matilde udkommer 4 gange
om året**

**Indlæg til næste nummer
skal være redaktionen i
hænde senest
fredag 15. august 2008**

Tema:

Sandsynlighedsregning

Indhold:

| | |
|--|----|
| <i>Freddy Bugge Christiansen</i> Leder | 2 |
| Tema: Matematikog biologi | |
| <i>Johnny T. Ottesen</i> Matematisk modellering og kardiovaskulær fysiologi | 4 |
| <i>Carsten Wiuf</i> Simple matematiske modeller til beskrivelse af komplekse biologiske systemer | 12 |
| <i>Freddy Bugge Christiansen</i> Evolutionær tilpasning..... | 16 |
| <i>Viggo Andreasen</i> Matematiske modeller af influenzas epidemiologi og evolution .. | 20 |
| <i>Michael Pedersen</i> DTU initiativet i Matematisk Bioscience..... | 26 |
| <i>Lars Gårding</i> Lämmelgåtan | 28 |
| <i>Christian Berg</i> Beretning for 2007 om arbejdet i Nationalkomiteen for matematik | |
| <i>Johan Forsell og Knud Søgaard</i> En bro mellem forskning og skole for voksne matematiklæring | 31 |
| <i>Ulf Persson</i> Selberg i Normat | 32 |
| <i>Uddannelsesfronten</i> <i>Morten Blomhøj og Tinne Hoff Kjeldsen</i> BASE – et kursus i matematisk modellering på den naturvidenskabelige basisuddannelse ved RUC | 34 |
| <i>Matematikernyt</i> | 35 |
| Aftermath..... | 40 |
| | 41 |

Matematisk modellering og kardiovaskulær fysiologi

af Johnny T. Ottesen
RUC, johnny@ruc.dk

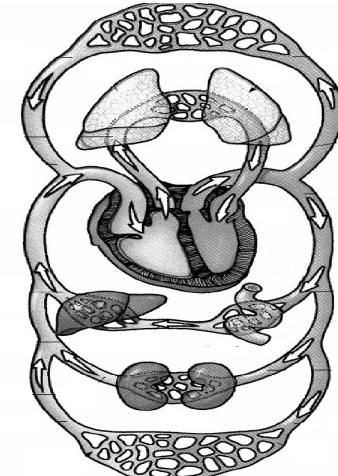
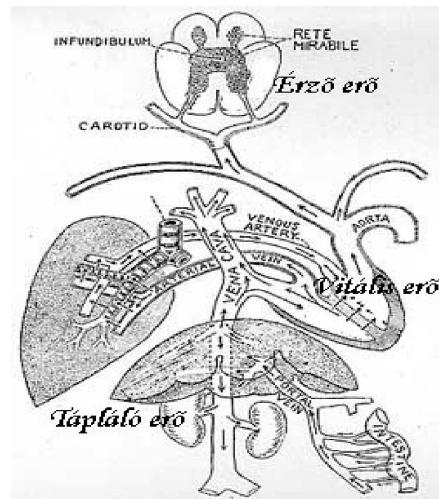


I denne artikel vil jeg illustrere, hvordan matematisk modellering på forskellig vis kan bidrage inden for biologi og specielt inden for kardiovaskulær fysiologi dels ved hjælp af simple historiske eksempler og dels ved hjælp af mere komplekse eksempler hentet fra egen forskning. Listen af eksempler og tilhørende pointer er på ingen måde udømmende og pointerne er elementære, dog uden at være trivielle. At pointerne ikke er trivielle illustreres af den noget naive og noget ureflekterede simuleringsmani, der ellers til tider præger matematisk modellering foretaget af personer uden matematisk kunnen.

Eksmpel 1 - Opdagelse af blodkredsløbet - det matematiske mikroskop.

I 1615 skrev den engelske læge William Harvey en revolutionerende sætning i sin dagbog: "Jeg betvivler den opfattelse af blodbanernes anatomi som har hersket i mere end 2000 år. For mig at se, er eneste rimelige mulighed at blodet recirkulerer i et lukket kredsløb med hjertet som pumpe." Før 1628 var det dogmatiske paradigme i grove træk at blodet blev produceret i leveren og transporteret i veneerne til hjertet, hvor der skete en opblanding med ilt (fra lungerne). Det resulterende "vitale blod" blev derefter transporteret i arterierne til de forskellige organer, hvor det blev konsumeret. Dvs. man troede at alt blod blev produceret, transporteret, modificeret og forbrugt i et ikke cirkulært system (grafteoretisk i et åbent træ). Grunde til at dette synspunkt herskede, var blandt andet at det i en vis udstrækning var understøttet af umiddelbare observationer på dissekerede lig fra slagmarker, hvor kapillærerne som i dag vides at forbinde de mindste arterier med de mindste veneer ikke kunne ses med det blotte øje. Dette blev først muligt med opfindelsen af lysmikroskopet 59 år efter at Harvey skrev sin revolutionerende sætning. Han turde imidlertid ikke offentliggøre sin kontroversielle konklusion straks. Først i 1628 turde fremføre sin banebrydende påstand, hvilke blev gjort ved en forelæsning og med en matematisk model - en konsekvensberegning vil nogle sige - som det primære argument: Hjertet pumper ca. 70 milliliter blod ud per slag og det slår i gennemsnit 72 gange per minut under normale omstændigheder, altså skulle mere end 7000 liter blod løbe gennem hjertet hver dag! Det er en grotesk mængde, ingen kunne forestille sig en mekanisme, der var i stand til at producere eller

forbruge så store mængder blod dagligt. I sammendragning er vores daglige væskeindtag under 5 liter og vores boldvolumen 5-6 liter, så der var ingen tvivl om at dette ikke var tænkligt. Eneste løsning på problemet var, hvis blodet blev recirkuleret. Selv om Harveys argument var slænende, var der en del, der forsøgte at bandlyse ham for hans påstand. Der gik 10-20 år før hans koncept med blodkredsløbet blev alment accepteret og de sidste kritikere blev først tavse med lysmikroskopet opfindelse, hvorved de usynlige kapillærer blev synlige for det menneskelige øje. Harvey havde gjort *det usynlige synligt* og *det utilgængelige tilgængeligt*.

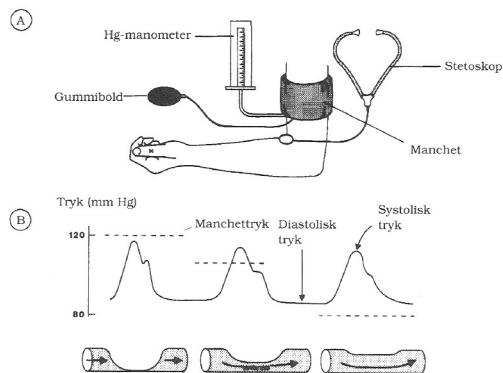


Matematiske modeller kan bruges til at falsificere koncepter og hypoteser og dermed give afsæt for nybrud og indsigt. Det skal i parentes bemærkes at Harvey også benyttede modellen til **at udvikle og strukturere sin videnskabelige tankegang**. Ud over brugen af matematiske modeller til at tolke enhver måling er falsificering nok den mest værdifulde brug af modeller inden for videnskaben.

Eksempel 2 - Måling af blodtryk.

Efter Harveys opdagelse, at blodet cirkulerer i et lukket kredsløb med hjertepumpen som drivkraft, begyndte man for alvor at spekulere over blodkredsløbets funktion og virkemåde. I 1733 målte Stephen Hales blodtrykket i en hest ved at fiksere en liggende hest på jorden og derefter bore et åbent glasrør ind i hestens halsarterie. Blodet dannede herefter en søjle i glasrøret og søjlens højde var således et udtryk for blodtrykket. Denne makaber målemetode blev aldrig standard og i dag måles blodtrykket hyppigst ved hjælp af en armmanchet hos egen læge. Armmanchetten er et fleksibelt 'armbånd' med et lufttæt kammer, en hård yderside og en blød side ind mod armen. Denne manchet spændes stramt om overarmen og der pumpes derefter luft ind i manchetten. Da ydersiden er hård sker udvidelsen mod armen. Antages armen på nær vene og arterier at være usammentrykkelig, resulterer dette i en sammenpresning af årerne i armen når manchetten fyldes, så det samlede volumen i arterier og manchet er konstant. Antages luftkammeret at ændre sig adiabatisk, kendes tryk-volumenrelationen således. Hvis man samtidig benytter det parsimoniske princip

(også kaldet Occam's razor; *keep the model as simple as possible*), vil en fornuftig antagelse om relationen mellem tryk P og volumen V for blodårerne være at trykændring og volumenændring er proportionale. Dette leder da til at trykændringerne i manchet og blodårer er proportionale. Proportionalitetsfaktoren afhænger af de respektive blodheder. Vælges manchetten til at være meget stivere end årerne, er de to trykændringer approksimativt



$$\Delta V_a + \Delta V_c = 0$$

$$\Delta V_a = c_a \Delta(P_a - P_c)$$

$$P_c V_c^\gamma = \text{constant}$$

$$\Delta P_c = \frac{c_a}{c_a + c_c} \Delta P_a \quad \text{with } c_c = \frac{V_c}{\gamma P_c}$$

$$\text{If } c_c \ll c_a \text{ then } \Delta P_c \approx \Delta P_a$$

lig hinanden. Trykket målt i manchetten giver derfor blodtrykket. Metoden baseres på at lægen pumper manchetten op, så årerne er helt sammenklemte. Derefter lukkes luft ud indtil trykket kommer under det systoliske tryk (maksimum tryk i arterien). Dette bemærker lægen ved at vedkommende i sit stetoskop kan høre 'klumper' af blod suse igennem åren, nemlig de 'klumper' hvor blodtrykket er over manchettrykket (i de tidsintervaller hvor trykket er under høres der ingen suselyd i stetoskopet). Nu lukkes yderligere luft ud indtil trykket passerer det diastoliske tryk (minimumstrykket i arterien). Da vil blodet strømme kontinuert og give anledning til konstant susen i stetoskopet). Metoden bygger på de to antagelser, at trykket i veneerne er meget mindre end det diastoliske tryk, således at disse er sammentrykket hele tiden samtidig med, at målingen foregår over et så kort tidsrum, at anden væske i armen ikke siver væk medens målingen varer. Er alt vel, har patienten et blodtryk, der typisk er omkring de 120 over 80 mmHg, hvilket blot er det systoliske blodtryk over det diastoliske blodtryk. Er blodtrykkene meget højere er det tegn på at patienten har problemer med helbredet. Den omvendte konklusion at hvis trykket er normal, så er alt vel, er desværre ikke gældende.

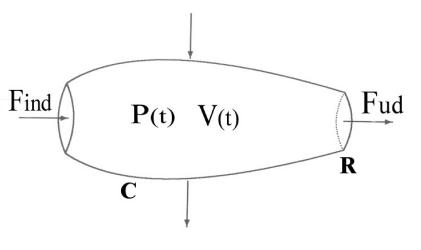
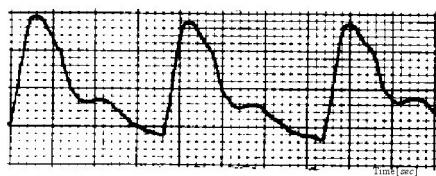


Modellering er superb for design og udvikling af målemetoder og systemer, hvilket er af stor betyd-

ning for industri og samfund. Det skal bemærkes at modellering er en kreativ proces, hvor entydighed ikke altid er til stede, modelløren skal træffe valg. Formålet med modelleringen er styrende og kriterier for succes kommer ikke fra det internt faglige matematiske univers, men fra det område der modelleres.

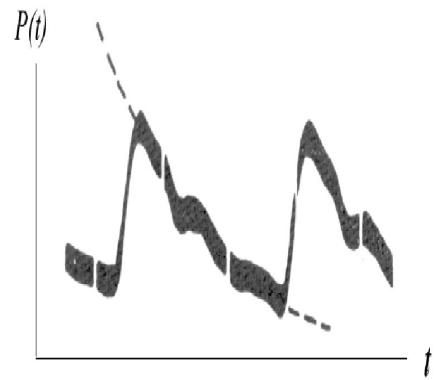
Eksempel 3 - Hvad bestemmer pulsprofilen i aorta?

Lad os begrænse os til spørgsmålet om hvilke faktorer der bestemmer den aftagende del af pulsprofilen i aorta, da dette kan besvares analytisk. Betragt aorta og hoveddelen af arterierne som en kæmpestør åre, et kompartiment, hvor tryk- og volumenaændringer er proportionale, hvor der er en modstand mod udströmning og hvor udströmning (F_{ud}) sker til et lavtryksområde (venerne), med et tryk vi sætter til nul. Da vi kun betragter en tidsperiode uden indströmning fra hjertet ($F_{ind}=0$), vil massebevarelse give en relation for den tidsafledte af volumenaændringen.



Samles disse informationer fås en velkendt differentialligning med eksponentialefunktionen som løsning. Sammenlignes modelresultat og data ses disse at stemme vældig godt overens på nær i et mindre område, hvor den såkaldte dikrote hævning forekommer. Årsagen til denne afvigelse er at modellen ikke inkluderer refleksion af trykbølgen forårsaget af forgreninger og den koniske indsnævrings af aorta.

$$\left\{ \begin{array}{l} \Delta V'(t) = -F_{out} \\ F_{out}(t) = \frac{1}{R} P(t) \\ P(t) = \frac{1}{C} \Delta V(t) \\ \Rightarrow \Delta V'(t) = -\frac{1}{CR} \Delta V(t) \\ \Rightarrow \Delta V(t) = V_s \exp\left\{-\frac{t}{CR}\right\} \text{ and } P(t) = P_0 \exp\left\{-\frac{t}{CR}\right\} \end{array} \right.$$



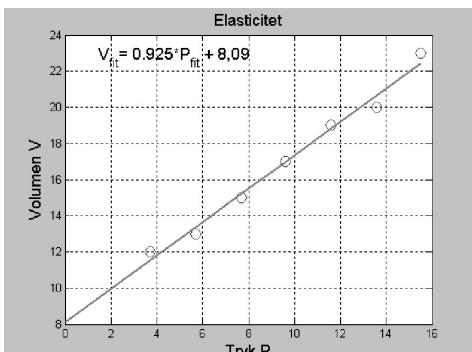
Matematisk modellering er et eminent redskab til at erhverve indsigt og forståelse i videnskabelige såvel som i samfundsmaessige problemstillinger. I bedste Sherlock Holmes stil guides forskeren til stringente deduktive konklusioner.

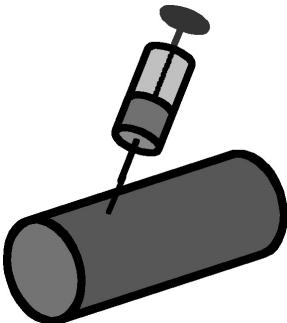
Eksempel 4 – Estimering af parametreværdier.

Som en eksperimentalfysiker engang sagde, "alle tror på eksperimenterne undtagen eksperimentalfysikeren selv." Dette er nok at sætte sagen lidt på spidsen og det kunne iøvrigt nok også siges om modelløren. Mere interessant er det at fremhæve det faktum, at *ingen måling kan foretages uden brug af matematiske modeller*. Det er ikke altid lige gennemsigtigt hvor og hvordan modellerne kommer ind, men helt sikkert er det, at de er der.

In vitro kan blødheden af en åre måles ved at tage et stykke ud og blokere enderne. Derefter fyldes stepvist væske ind i årestykket. Måles sammenhørende værdier af tryk og volumen kan disse afbildes i et tryk-volumen koordinatsystem. Hvis hypotesen om lineær sammenhæng er holdbar, skal punkterne ligge med tilstrækkelig tilnærmelse på en ret linie, hvis hældning da er et udtryk for blødheden. En anden metode ville være at måle blødheden in vivo f.eks. for aorta. En god tilnærmelse fås ved at tage hjertets slagvolumen og dividere med differensen mellem systolisk og diastolisk tryk. Modstanden mod udströmning kunne ligeledes estimeres ved brug af definitionen tryk per strömning. Et estimat for aorta vil da fås ved at benytte middelværdier: Modstanden er approksimativt givet ved gennemsnittet af systolisk og diastolisk tryk divideret med gennemsnitsströmningen, som er lig med gennemsnitsstrømmen ud ad hjertet. Sidstnævnte er hjertets slagfrekvens ganget med slagvolumen. Bemærk at produktet af modstand og blødhed giver halveringstiden for den tilnærmelsesvis eksponentielt aftagende del af

pulsprofilen som vist i eksempel 3.





Herved fås en uafhængig kontrol af de estimerede størrelser. Metoderne bruges i dag klinisk, men enhver matematikfagkyndig ville formentlig benytte bedre metoder hvis nøjagtigheden har betydning. Dette kunne være klassiske optimeringsalgoritmer, Kalman filtreringsmetoder kendt fra ingenørverdenen, eller andre statistiske metoder. Som en matematisk kuriositet kan det nævnes, at i forbindelse med parameterestimering er der tale om inverse problemstillinger hvor eksistens og entydighed ofte er ikke trivielle spørgsmål (bl.a. ill-posedness- og regularization spørgsmål adresseres her) ligesom spørgsmål om stabilitet og robusthed er ikke trivielle. Disse og andre spørgsmål er for tiden højaktuelle, da målinger på mennesker er underlagt stærke restriktioner, hvorfor der er begrænset adgang til data og dette samtidig med at der er stor personspecifik variation over tid.

Matematisk modellering er eneste måde at definere begreber på således at disse er entydigt defineret. Det er en selvstændig pointe, at **eneste måde hvorpå parametre kan estimeres er ved hjælp af matematiske modeller** (i tilknytning til eksperimenter eller observationer).

Eksempel 5. Modellering af hele det kropslige kredsløb.

Blodkredsløbet er alt for komplekst til at kunne modelleres i stor detalje på alle skalaer. Benyttes kompartmentmodellering i stil med ovenstående, kan et brugbart kompromis dog ofte opnås. Antallet af kompartments afhænger naturligvis af, hvad formålet er ligesom beskrivelse af delsystemerne afhænger af dette. Man kan oven i dette tilføje de kontrolmekanismer der regulerer blodkredsløbet osv. Alle delmodellerne kan da (i principippet) studeres og valideres enkeltvis inden de sammensættes til en mere omfattende og kompleks model. Det skal bemærkes, at dette i mange tilfælde udgør 'state of the art' på området, f.eks. når det handler om at studere fornævnte kontrolmekanismer og deres interaktion.

Medens moderne naturvidenskab - og ikke mindst moderne biologi - er god til at separere systemer i komponenter, der er simple nok til, at deres struktur og funktion håbefuldtilt kan studeres isoleret, så er **matematisk modellering eneste bud på hvordan disse dele på kontrolleret vis kan sættes sammen igen.** På denne vis **tilbyder matematisk modellering en enestående mulighed for at forstå komplekse systemer.**

Eksempel 6 – Modellering gør det utilgængelige tilgængeligt.

De følgende eksempler er hentet fra min egen forskning. I eksperimenter kan man umiddelbart skelne oscillationer i blodtrykket hidrørende fra hjerteslagene fra de der skyldes respiration, men effekter af de forskellige mekanismer kan ikke studeres separat, de er så at sige ikke tilgængelige separat. Det er sådan, at man trækker vejret ved at sænke (indånding) henholdsvis løfte (udånding) mellemgulvet (diafragma). Herved ændres på de ydre tryk i forhold til blodårerne, hvorved der trykkes (pumpes) på disse årer. I øvre torso stiger trykket, når det falder i nedre torso og omvendt.



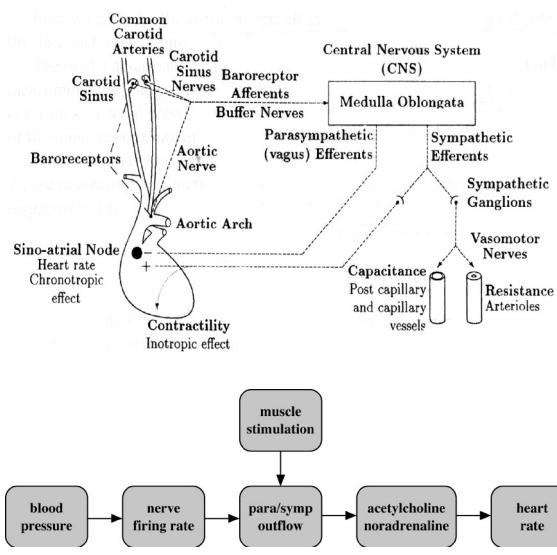
I en model af blodkredsløbet er det relativt nemt at tilføje denne respiratoriske effekt. Og resultatet er overraskende. Det er måske ikke så overraskende at dybe vejrtrækninger understøtter hjertet

som en ekstra (lille) pumpemekanisme og måske heller ikke, at denne pumpemekanisme favoriserer visse dele af kroppen med op til 100%. Men det er yderst overraskende, at selv når middeltrykket holdes konstant mens respirationsdybde og base-level-værdi ændres, så kan der forekomme op til en fordobling af blodstrømningen gennem nedre torso uden at de øvrige perifere regioner ændrer blodstrømning væsentligt. Altså forekommer der en relativ redistribution af strømning som en direkte konsekvens af respirationens mekaniske kobling til blodkredsløbet helt uden inddragelse af det autonome kontrolsystem. Uden sidstnævnte vil man fejlagtigt tilfælde de respiratoriske autonome kontrolmekanismer en for stor vægt i forbindelse med en evt. parameterestimering. Det er i øvrigt vigtigt på dette sted at understrege det elementære faktum at *ændres en model, så vil de tilhørende parameterværdier (normalt) også ændres.* En erkendelse der kun sjeldent rækker udover matematikfagkyndiges cirkler! Det er endvidere interessant, at åbnes og analyseres den nævnte model, fås Starling's hjertelov fra 1914 som konsekvens (nemlig at hjertes slagvolumen som funktion af fyldningstryk giver en karakteristisk familie af kurver parametreret ved respirationstrykket.)

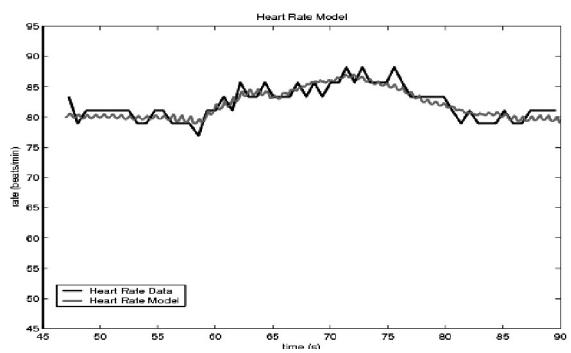
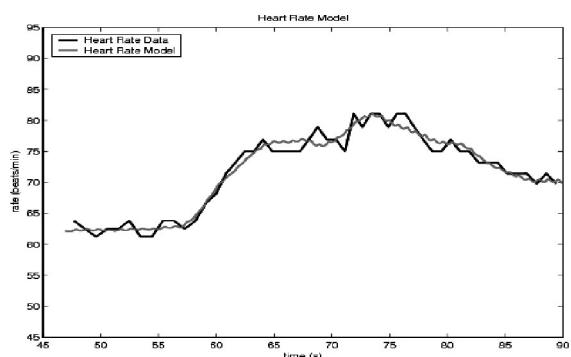
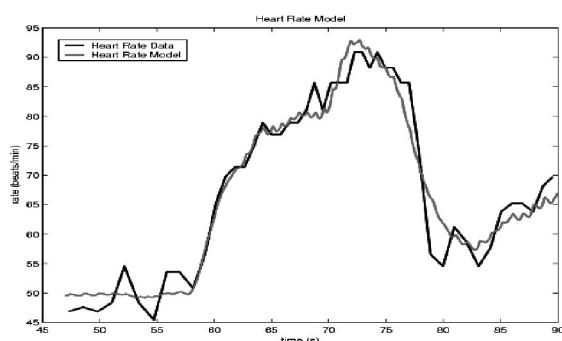
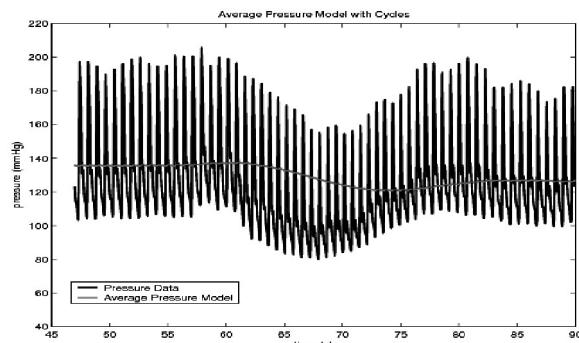
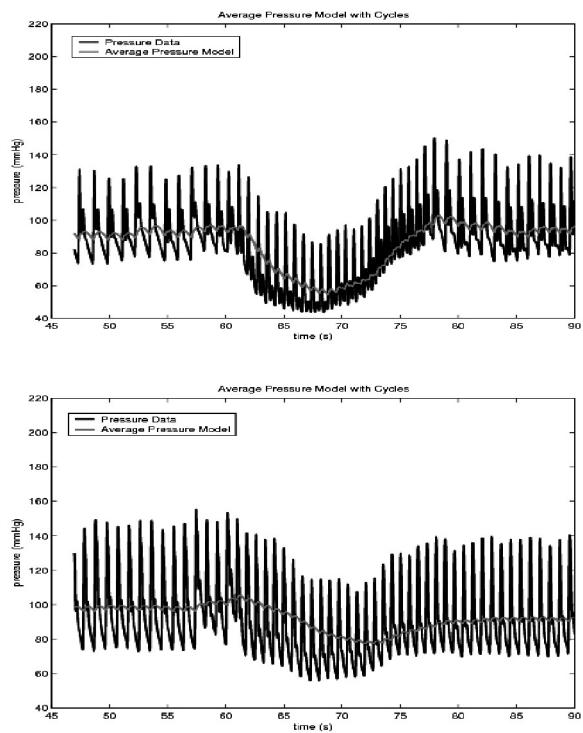
Dele der ikke kan isoleres og studeres enkeltvist i naturen, kan ved hjælp af matematiske modeller studeres enkeltvist og adgang til sådanne systemers parameterværdier kan dermed erhverves. Det **matematiske mikroskop** kommer igen på banen denne gang ved at **gøre det utilgængelige tilgængeligt og klargøre sammenhængskomponenter, somellers var skjulte.**

Eksempel 7 – Modeller foreslår bio-markører til diagnosticering.

Opstilles en kompartmentmodel af hele det kardiovaskulære- og respiratoriske system med diverse kontrolmekanismer, kan man f.eks. studere, hvad der sker med blodtryk og blodstrømning når en person aktivt ændrer stilling fra siddende til stående eller fra passivt liggende på et bord, der kan vippes, til delvist oprejst stilling. Her kan man måle blodstryk og blodstrømning nogle få udvalgte steder. Har man tiltro til modellen får en indsigt i de autonome kontrollmekanismer (samt andre lokale kontrollmekanismer, som autoregulering og myogen kontrol etc.) En model af baroreceptorkontrol mekanismen, der bl.a. virker styrende for hjertets slagfrekvens på kort tids skala, er skitseret nedenfor.



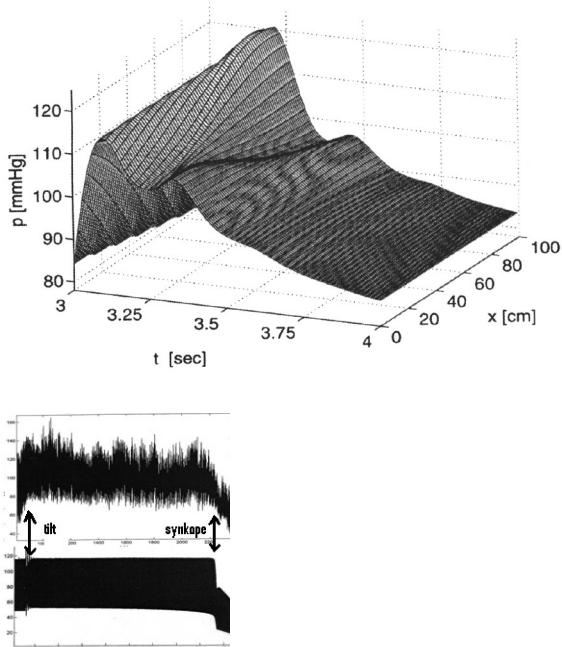
Nedenfor ses resultatet af tre typiske objekter, en ung rask person, en ældre rask person og en ældre person med forhøjet blodtryk for det scenarie, hvor objekterne aktivt



skifter position fra siddende til stående. Venstre øjle er blodtrykket versus tid og højre øjle er hjertets slagfrekvens versus tid. Overensstemmelse mellem kurverne er slænde, hvilket er bemærkelsesværdigt, da der kun er brugt to frihedsgrader mere end et maksimum likelihood estimat kræver (dvs. 16 parametre i forhold til 14.) Tiltroen til en så flot overensstemmelse er stor, hvorfor man fristes til at have tiltro til modellens mellemliggende lag. Der tilbydes således en indsigt i det sympatiske og det parasympatiske kontrollsystem, som udviser store forskelle i de tre nævnte tilfælde og dermed foreslår bio-markører, dvs. let tilgængelige mål for systemets / patientens (patologiske) tilstand. Disse er efterfølgende blevet undersøgt mellem forskellige patientgrupper med positivt udfald.



Cirka 6% af alle hospitalsindlæggelser skyldes brud eller hjernerystelse som følge af svimmelhed eller besvismesse (synkope) og 40% af alle danskere oplever synkope mindst en gang i livet. Fænomenet skyldes, at der ikke kommer ilt nok til hjernen og at der ikke transportereres karbondioxid nok væk fra hjernen, fordi blodtilførelsen falder. Dette fremprovokeres ofte når der skiftes fra liggende til stående stilling og kan derfor undersøges ved tilt-eksperimenter som nævnt i forrige eksempel. En distribueret model (hvor Navier-Stokes ligninger, visko-elasticitet og kontinuitetsligningen er nogle af grundelementerne) giver et analytisk udtryk for trykbølgers udbredelseshastighed. Når en person rejser sig op eller vippes op på et tiltbord, så trækker gravitationen blod ned i benene, hvorfor forsyningen til hjernen daler. Systemets kontrolmekanismer skal da helst hurtigt rette op på dette 'forsyningssvigt'.

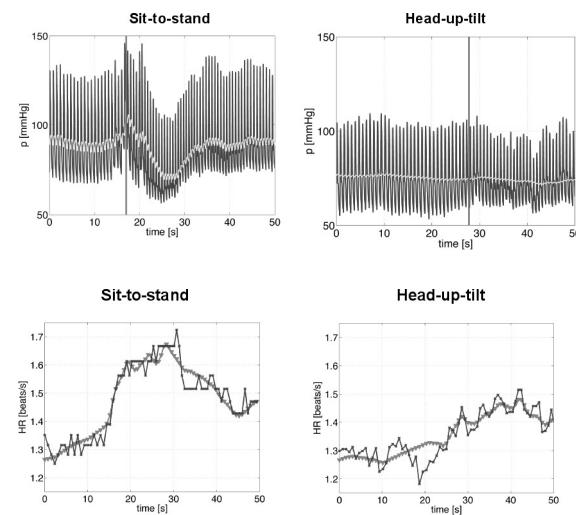


En af de kontrolmekanismer, der virker på en mellemlang tidsskala (af størrelsesorden minutter), er reguleringen af årernes blødhed og deres modstand mod gennemstrømning via regulering af årernes radius. En ændring af disse 'parametre' ses i ændringen af trykbølgeudbredelseshastigheden, der netop afhænger af disse. Herved fås direkte indsigt i disse kontrolmekanismer. Ved tilt-eksperimenter kan personers kontrolsystem altså undersøges og i modellen kan man se, under hvilke omstændigheder synkope vil opstå. Dette kan i et fremtidigt perspektiv være ret afgørende for dialysepatienter, der ofte må stoppe behandlinger i tide formentlig pga. samme mekanismes 'kollaps'.

Matematisk modellering tilbyder bio-markører, som kan anvendes ved diagnosticering og behandling på såvel individ- som populationsniveau.

Det er selvfølgelig fristende at sammenligne sit-to-stand modellen og head-up-tilt modellen. Det viser sig at begge datasæt kan fittes vældig fint ved i samme model kun at

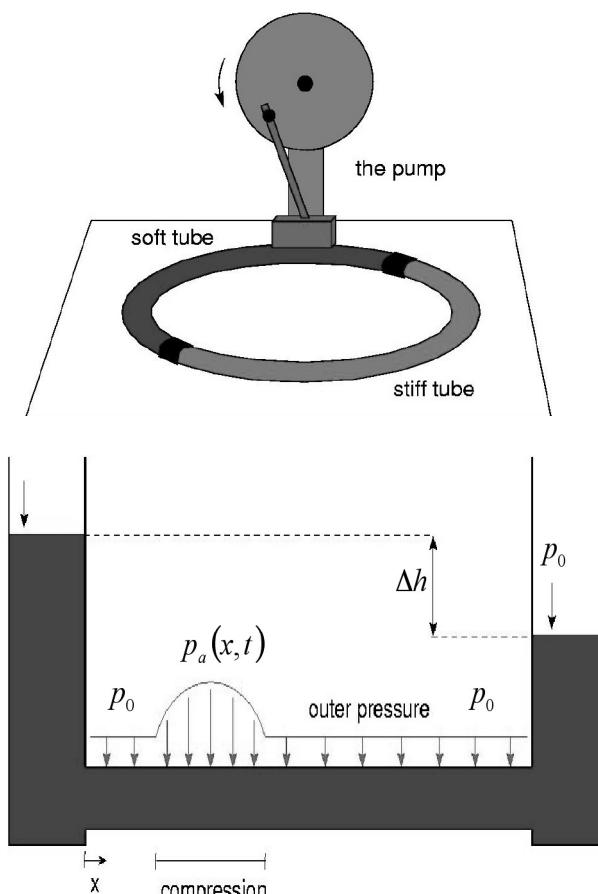
variere signifikant på to parameter. De to figurer til venstre er trykkene, som er input til modellen, og de to figurer til højre viser sammenligninger mellem de respektive modelherteslagfrekvenser og de tilsvarende datasæt.



Eksempel 8 – universalitet kan undersøges ved modellering.

En matematisk model af gummislanger og væske kan give vigtig viden om biomedicinske systemer og vise at *matematik er livsvigtig*.

Tag to gennemsigtige gummislanger af forskellig elasticitet og sæt dem sammen så der dannes en ring. Før ringen lukkes, fyldes der vand i slangerne (tilsæt evt. et par dråber farvet olie, så man kan se strømningen). Klem nu rytmisk på det bløde rør og se, hvad der sker.

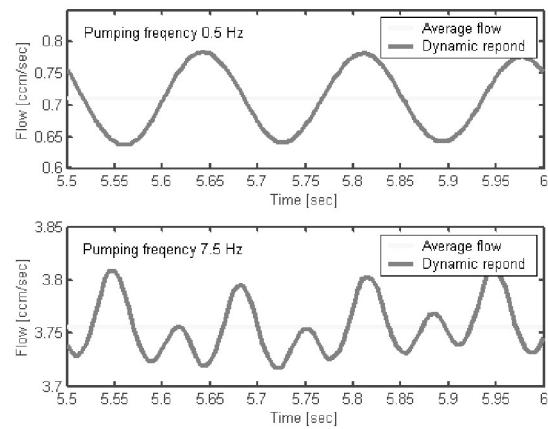
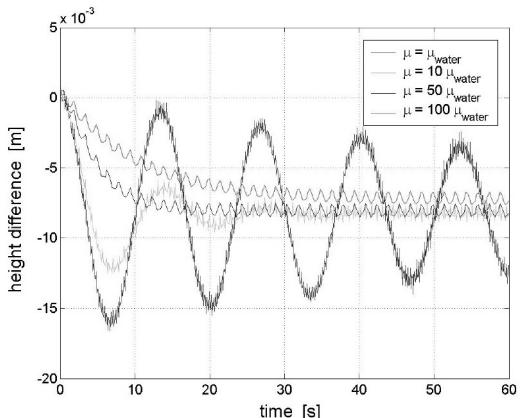


Væsken skvulper ikke bare lidt frem og tilbage, som de fleste ville tro. Den begynder at cirkulere i ringen. Fænomenet har været kendt i århundreder, men ingen kunne forklare, hvordan væsken kan strømme rundt i et kredsløb, når der ikke er en ventil, som forhindrer det i at løbe tilbage igen i pauserne mellem kompressionerne.

Man ser et lignende fænomen, hvis man tager en gummislange som forbinder to vandkar. Når man klemmer gummislangen rytmisk sammen, vil vandet netto begynde at strømme den ene vej indtil gravitationen afbalancerer den kraft pumpmekanismen skaber. Resultatet er, at vandstanden i det ene kar står højere end i det andet. Ændres pumpefrekvensen så varierer højdeforskellen tilsvarende og ved visse frekvenser skifter den eleverede vandhøjde fra det ene kar til det andet. Der sker lignende skift hvis pumpestedet flyttes langs slangen eller sammentrykningen af slangen sker mere voldsomt. Dette slangesystem udviser en væsentlig forskel i forhold til det ringformede slangesystem omtalt ovenfor. Skiftet i eleveret vandhøjde fra det ene kar til det andet kan ske på to fundamentalt forskellige måder: Via et resonanspunkt eller via en glat overgang kaldet et kontinuitetspunkt.

Når man giver en person hjertemassage ved at presse brystkassen hårdt sammen i rytmiske tryk, udfører man principielt det samme "eksperiment" dog indeholder dette system klapper. Her er virkningen livsvigtig. Derfor er det vigtigt at vide, hvad der sker. Forsøg udført af basis- og matematikstuderende på RUC har vist, at væsken faktisk kan bringes til både at stoppe og skifte retning i gummislangesystemet.

Over de sidste 10 år har jeg udviklet matematisk modeller, som simulerer og delvist forklarer klappeløse pumpmekanismer. Når pumpefrekvensen øges i det ringformede system ændres strømningshastigheden, og ved bestemte frekvenser (f.eks. tre pumpetryk pr. sekund) standser strømningen og hvis frekvensen yderligere øges, vendes strømretningen og væsken strømmer derefter den anden vej rundt. Dette fænomen gentager sig ved yderligere voksende frekvenser. Fænomenet er overraskende, for det kan ikke forstås som et lineært respons eller ud fra en simpel ikke-lineær teori. Klappeløs strømning er en universal skalainvariant effekt der skyldes at modstand mod væskestrømning i elastiske slangesystemer (samt at eftergivelighed af slangen og træghed mod acceleration af væsken) er ikke-lineær og radiusafhængig samt at radius afhænger af trykket, der jo varierer i tid og rum. Kompleksitetsrigdommen i fænomenet overrasker selv for garvede hydrodynamikere og fænomenet byder stadig på uløste problemer.



Fænomenet klappeløs strømning er vigtig i mange forsknings- og anvendelsesmæssige sammenhænge: I 3.-4. uge af fostertilværelsen pumper hjertet blod, men det har endnu ikke udviklet hjerteklapper. Mere dramatisk er den type "klappeløs pumpning", som finder sted når man giver hjertemassage. Her er opdagelsen i bogstavelig forstand livsvigtig for den viser, at effekten / resultatet af den potentielle pumpning af blodet kan afhænge af frekvensen, der masseres med. Gøres det forkert, vil blodet standse helt op – eller forsøge at løbe den forkerte vej i årerne. Ved hjælp af modeller kan man studere optimering af hjertemassage, hvilket selvsagt ikke kan lade sig gøre eksperimentelt uden etiske problemer.

Klappeløs strømning har også betydning for de små blodkar i øjet. Fourier spektret fra direkte observationer af kapillærradius i øjet indeholder mange frekvenser, hvoraf visse frekvenser statistisk set er signifikant tilstede for raske individer, mens disse er fraværende for personer med diabetes endda allerede i en meget tidlig fase af sygdommen. Dette er benyttet til tidlig diabetes-diagnosetisering gennem non-invasive målinger på kapillærer i øjet men rejser desuden spørgsmålet om oscillationernes funktion.

Fænomenet klappeløs strømning optræder også når musklene benyttes. Muskler trækker sig sammen når de stimuleres, hvorved de mange små blodårer der løber gennem den aktive muskel sammentrykkes.

Den menneskelige vejtrækning foregår ved at man periodisk sænker diafragma (mellemgulvet), som beskrevet i eksempel 6, hvorved et periodisk undertryk henholdsvis overtryk i brystregionen (og modsat i mellemgulvet) etableres. Dette har stor betydning bl.a. i forbindelse med arbejds- og idrætsmæssige præstationer.

Endelig har fænomenet stor betydning i forbindelse med mikro- og nanopumper. På lille skala kan man ikke konstruere konventionelle pumpmekanismer, men da det ofte er partikler med ladning eller spin, der skal ledes gennem en kanal, kan principippet benyttes ved at påtvinge nanokanalen et eksternt elektrisk eller magnetisk felt. Herved fremkommer en klappeløs pumpmekanisme der kan benyttes til at regulere strømninger på mikro- og nanoskala (de styrende ligninger er her lidt anderledes).

En sidste noget spekulativ anvendelse har med verdenshavene at gøre. Månenes periodiske bane om jorden udsætter vandet i verdenshavene for en periodisk gravitationel påvirkning. Det er meget tænkligt at der

herved fremkommer en klappeløs pumpemekanisme, der foruden at skabe tidevandseffekter er med til at styre strømningen i verdenshavene.

Kun matematik byder på en eksakt universalitet:

At tilsyneladende forskellige systemer strukturelt er ens fremgår af om de matematiske strukturer i modellerne af disse systemer er ækvivalente. Videns og indsigt kan herved overføres fra område til område. Endelig giver matematiske modeller mulighed for at undersøge modellen på dens egne præmisser, dvs. **man kan eksperimentere med modellen og lave undersøgelser, som det ikke er mulig at udføre på det 'virkelige' system der modelleres.**

Eksempel 9 – Modellering og intern faglig matematik

Det er klart for enhver, der har arbejdet intensivt med modellering, at der indgår alverdens matematik og ofte på meget overraskende vis. Her vil jeg blot nævne nogle af de klassisk matematiske problemstillinger, som er involveret i forbindelse med ovenstående.

I forbindelse med kompartmentmodellering er det ofte kvalitativ teori for dynamiske systemer der er på banen. Alene ud fra de grundlæggende strukturer kan man vise en lang række generelle sætninger. Periodiske løsninger, stabilitet, bifurkationer og grænsecykler er her gennemgående fænomener der studeres.

I forbindelse med parameterestimering er en forudgående undersøgelse af disses identificerbarhed vigtig, dvs. en undersøgelse af hvorvidt det er meningsfyldt at estimerer disse parametre. Selv for simple lineære differentialligninger, er spørgsmålet om identificerbarhed af parametre ikke triviel. Dette skyldes at man her betragter det inverse problem, som ikke nødvendigvis behøver være lineært. Tag blot et eksempel med koncentrationen af et medikament (x) i blodet, der er tilført kroppen via en mængde (u). Opfattes kroppen som et kompartiment hvor medikamentet løbende nedbrydes eller udvaskes, kan systemet beskrives ved

$$\begin{aligned}x'(t) &= -p_1 x(t) + p_2 u(t) \\y(t) &= p_3 x(t)\end{aligned}$$

Her betegner p_1 en tidskonstant for nedbrydningen/udvaskningen, p_2 en omregningskonstant fra mængde tilført medikament per tidsenhed til medikament per volumen blod per tid, og p_3 er en proportionalitetskonstant der ganget på koncentrationen af medikament i kroppen giver målingen y , kaldet den observable. Her er ikke alle tre parametre identificerbare uden yderligere observatiorer. Kun p_3 og produktet $p_1 p_2$ er observerbare.

Endelig i forbindelse med de distribuerede modeller betragtes ofte Navier-Stokes ligninger og diverse approksimationer defineret på grafer. Kanterne i graferne repræsenterer de elastiske årer, hvorfor konstitutivt tilstands ligninger yderligere er nødvendige. Dette giver under visse approksimationer ofte anledning til et system af

hyperboliske differentialligninger og relaterede Riemann problemer. Eksistens og entydighed af løsninger på forskellige rum er svært centralt, og for nylig blev global eksistens af svage løsninger vist ved hjælp af Glimm's random choice skema.

Matematisk modellering benytter ikke blot spændende matematik, men er også med til at **udvikle nye internt faglige resultater og generere nye matematiske discipliner.**

Til slut vil jeg gerne påpege, at mange studerende tænder på matematisk modellering, medens andre tænder på ren abstrakt matematik, hvilket bør udnyttes pædagogisk i højere grad end det er tilfældet i dag. Matematik er verdens største undervisningsfag og legitimiteten herfor er at matematik kan anvendes. For at leve op til det moderne samfunds krav, bør de der foretrakker den rene abstrakte tilgang til faget derfor også udvikle kompetencer indenfor matematisk modellering ligesom de der anvender matematik bør have en solid indsigt i grundlæggende matematiske strukturer og tankegange. Lad mig slutteligt citerer Professor Vagn Lundsgaard Hansen, "For at være i stand til formålsrettet at udvikle matematiske modeller skal man kunne matematik, kunne bringe matematik i anvendelse og kunne vurdere matematikanvendelsen. Disse kompetencer er langt fra sammenfaldende. Og det er ikke nogen simpel sag at organisere vekselspillet imellem matematik i egen ret og matematik som redskab i anvendelsessammenhænge optimalt."

Simple matematiske modeller til beskrivelse af komplekse biologiske systemer

af Carsten Wiuf
Center for Bioinformatik
(BiRC) ved Aarhus Universitet,
wiuf@birc.au.dk



Biologien er i disse år i rivende udvikling blandt andet på grund af de mange tekniske fremskridt der gør det muligt simultant at måle og registrere tusinder af variable. For eksempel er det i dag muligt at bestemme aktiviteten af alle gener i en celle på en gang, hvilke af en celles tusinder af proteiner der vekselvirker med hinanden eller afkode millioner af variable DNA-positioner i det humane genom ved et enkelt eksperiment.

Den slags målinger giver os mulighed for at studere celler, væv og organismer i et nyt perspektiv – et helheds- eller systemperspektiv, hvor tendenser og mønstre træder frem som ikke er synlige, hvis kun en lille del af systemet observeres. Således kan man for eksempel se, at generne i en hjernecelle er markant anderledes udtrykt end generne i en levercelle, og at udtrykkene ændrer sig gennem en organismes levetid. Udover at være interessant i sig selv, åbner sådanne observationer for nye muligheder inden for fx lægevidenskaben, idet vi ud fra et systemperspektiv kan sammenligne syge celler med raske og belyse sygdommes årsager og undersøge effekten af medicin i et nyt lys.

Ikke kun den molekulære biologi er i rivende udvikling, men også andre grene af biologien oplever tilsvarende forandringer i form af øgede datamængder, øget information og viden. Dette stiller naturligvis krav til måden at registrere og indsamle data på, samt til hvordan vi efterfølgende behandler og analyserer data med henblik på spørgsmål vi finder videnskabeligt eller samfundsmæssigt relevante. Derfor kan man parallelt med udviklingen inden for biologien spore en udvikling inden for statistik, datalogi og matematik, hvor man søger at udvikle modeller, metoder og teknikker til håndtering og analyse af store datamængder; dette kræver indsigt i biologisk teori. Det samspil eller krydsfelt, der dermed opstår mellem biologi, statistik, datalogi og matematik kaldes for bioinformatik.

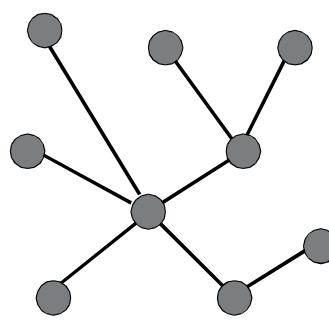
Faget bioinformatik har udviklet sig fra i slutningen af 60erne at være enkelpersoners sammenligninger af små DNA- og proteinsekvenser til i dag at være centre med mange ansatte med forskellige baggrunde og eksperter. I Danmark er der større centre (>20 ansatte) på Danmarks Tekniske Universitet, Københavns Universitet og Aarhus Universitet. Jeg er selv ansat

på Center for Bioinformatik ved Aarhus Universitet og her er folk med baggrund i matematik, statistik, datalogi, biologi, kemi, molekylærbiologi og ingeniørvidenskab. Dette giver et meget spændende arbejdsmiljø, hvor den fælles interesse er at forstå biologiske problemstillinger gennem teoretiske overvejelser.

Selvom alle de nævnte fagområder gennem årtier har opereret med begrebet *systemteori* i forskellige former, er begrebet dukket op på ny i forøget styrke som en følge af de muligheder, der har vist sig inden for biologien i dag. Her vil jeg beskrive et biologisk system kaldet *interaktomet* og vise hvordan vi med simple matematiske modeller kan beskrive hvordan interaktomet gennem evolution har formet sig over tid. Artiklen er bygget op så jeg først beskriver biologien, dernæst motiverer valget af matematisk model og diskuterer en statistisk analyse af biologiske data vha. modellen. Endelig fortolker jeg de statistiske resultater til matematiske udsagn om modellen.

Interaktomet: Biologisk fortolkning og evolution

Mængden af alle de forskellige proteiner i en organisme samt de fysiske bindinger, der er mellem par af proteiner kaldes for interaktomet; se Figur 1. Proteiner er produkter af gener, og et gen udøver til dels sin funktion gennem dets protein og de vekselvirkninger proteinet kan have med andre proteiner. Det enkelte protein påvirker cellens processer ved at binde til andre proteiner, gener og molekyler og interaktomet repræsenterer dermed den del af cellens processer, der er bestemt af bindinger mellem proteiner.



Figur 1. Interaktomet illustreret som en graf med proteiner som knuder og fysiske vekselvirkninger som kanter. I eksemplet er der $N=9$ knuder og $M=8$ kanter. Graden af en knude er antallet af kanter der forbinder den med andre knuder.



Der er derfor stor interesse for at forstå interaktomet og de evolutionære processer der har skabt det. Et eksperimentelt bestemt interaktom kaldes et *proteininteraktionsnetværk* (PIN). I dag er PIN-data fra en række organismer tilgængelige, dog er PIN-data fra mennesket endnu sparsomme, da de anvendte eksperimentelle teknikker stadig er dyre og tidskrævende. Desværre er PIN-data også ufuldstændige i den forstand at kun kendte proteiner kan medtages i et eksperiment. I Tabel 1 er størrelsen på typiske PIN-datasæt angivet med et bud på hvor fuldstændigt datasættet er.

| Organisme | Knuder | Kanter | Total | Procent |
|--------------------------------|--------|--------|--------|---------|
| <i>S. cerevisiae</i> (gær) | 4.959 | 17.226 | 5.500 | 90 |
| <i>D. melanogaster</i> (flue) | 7.451 | 22.636 | 12.900 | 58 |
| <i>C. elegans</i> (orm) | 2.638 | 3.970 | 22.000 | 12 |
| <i>H. pylori</i> (bakterie) | 675 | 1.096 | 1.500 | 45 |
| <i>P. falciparum</i> (parasit) | 1.271 | 2.642 | 5.300 | 24 |

Tabel 1. Antallet af knuder og kanter i typiske PIN-datasæt.

Total: Estimeret interaktomstørrelse.

Procent: Antallet af knuder (proteiner) i % af det estimerede antal af proteiner i organismen.

Data er taget fra en offentlig tilgængelig database DIP (Database of Interacting Proteins).

En organisms interaktom udvikler sig evolutionært over tid. I sin yderste konsekvens forestiller vi os at interaktomet er startet som et enkelt protein og vokset i størrelse over tid. Da vi ikke til fulde forstår livets udvikling i dets tidlige faser, er interaktomets udvikling fra et enkelt protein at betragte som en bekvem abstraktion. Til gengæld findes en række andre biologiske processer, der ændrer interaktomet, og som vi forstår langt bedre. Den vigtigste er genduplikation, en proces hvorved et eksisterende gen bliver kopieret og indsats i genomet, så der lige efter duplikationen er to kopier af det samme gen. Disse to gener er identiske og producerer derfor det samme protein. Interaktomet ændrer sig altså ikke umiddelbart som følge af en genduplikation. Over relativt kort tid (i evolutionær forstand) udvikler de to gener sig; her er der forskellige muligheder: 1) Det ene gen ødelægges af ska-

delige mutationer og ophører med at være aktivt, 2) Det ene gen får en helt ny funktion vha. gavnlige mutationer eller 3) De to gener deler det oprindelige gens funktioner mellem sig. Den sidste mulighed menes at være langt den vigtigste. De ændrede funktioner giver anledning til ændringer i de proteiner generne producerer og dermed også til ændringer i de bindinger proteinerne har til andre proteiner. Dvs. over tid sætter genduplikation sig spor i interaktomet; se Figur 2.

Genduplikation er evolutionært en meget vigtig proces, idet den giver et gen/protein mulighed for at specialisere sig (se punkt 3) – fx ser man ofte proteiner der kun er udtrykt i en bestemt type væv, men som er tæt beslægtede med andre proteiner med lignende funktioner i andet væv. Derudover findes der en række andre processer, der specielt er vigtige i bakterieevolution og som vi samlet vil betegne *tilknytningsprocesser*. De dækker integration af gener fra vira og andre bakterier og fusion af eksisterende gener i genomet, hvorved organismen bliver i stand til at udvikle gener, og dermed proteiner med helt nye funktioner. Den relative hyppighed af og vigtighed mellem de forskellige typer processer er i det store hele ukendt.

En matematisk model af interaktomets evolution

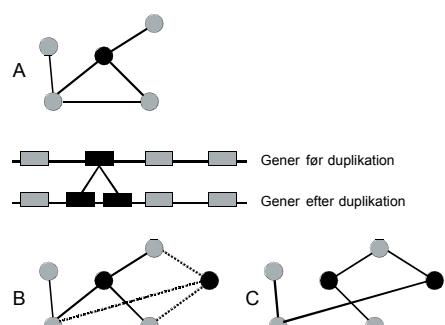
Første del af den matematiske model består i at repræsentere interaktomet ved hjælp af en graf som vist i Figur 1. Grafer (eller netværk) finder i dag hyppig anvendelse i mange naturvidenskabelige fag blandt andet fordi der er udviklet en rig og sund matematisk teori for grafer. De er desuden nemme at visualisere og derfor naturligt appellerende.

De grafer vi betragter ændrer sig over evolutionær tid på en måde vi ikke på forhånd kender, og vi beskriver dem derfor i sandsynlighedsteoretiske eller stokastiske termmer. Lad $G_t = (V_t, E_t)$ betegne en graf med t knuder, hvor V_t er mængden af knuder og E_t er mængden af kanter i grafen. Et netværk bygges op således:

- Start med en graf $G_s = (V_s, E_s)$ af størrelse s . Vi vælger $s=1$, dvs. en knude uden kanter
- Antag grafen i skridt t er $G_t = (V_t, E_t)$. I næste skridt, $t+1$, gør følgende:
- Duplikation: Vælg med sandsynlighed $0 \leq \alpha \leq 1$ en knude tilfældigt i grafen og kopier knuden med alle dens kanter; se Figur 2. Dernæst modificeres den nye

og den gamle knude vha. parametrene p og q , som beskrevet i figur 2

- Tilknytning. Vælg med sandsynlighed $1-\alpha$ en knude tilfældigt i grafen. Forbind en ny knude til den valgte knude ved hjælp af en enkelt kant
- Fortsæt indtil netværket har den ønskede størrelse.



Figur 2. Figuren viser genomet før og efter en genduplikation og det tilhørende proteinnetværk (A). Lige efter duplikationen af det sorte gen har de to sorte gener de samme interaktioner (B), men i løbet af evolutionær kort tid differentieres disse (C). I den matematiske model antages at begge kanter til et gråt protein bevares med sandsynlighed p og at en kant skabes mellem de to sorte proteiner med sandsynlighed q . Hvis kun en kant bevares, vælges med lige sandsynlighed. I eksemplet er der sandsynlighed $p(1-p)^2(1-q)/4$ for at observere netværket i (C), hvis det sorte gen duplikeres.

Dernæst udvælges knuder så antallet passer med et observeret PIN-datasæt

Der er naturligvis mulighed for variationer og for at udvikle modellen yderligere. Den beskrevne model er specificeret ved tre parametre (α, p, q), og da det er en evolutionær model vil grafen i skridt $t+1$ kun afhænge af grafen i det foregående skridt t – dvs. vores grafgenerende proces er en Markov-kæde. Desuden er enhver modifikation af grafen lokal – dvs. den involverer kun få knuder (den nye og den gamle, plus den gammels naboer). Disse egenskaber sætter os i stand til at udsige interessante udsagn om de grafer vi kan generere.

To eksempler: *Helicobacter pylori* og *Plasmodium falciparum*

Jeg og min gruppe har analyseret PIN-datasæt fra *H. pylori* (675 knuder) og *P. falciparum* (1.271 knuder). I øjeblikket er vi ved at analysere et PIN-datasæt fra gær (*Saccharomyces cerevisiae*) med ca. 4.500 knuder, hvilket er på grænsen af hvad vi i øjeblikket beregningsmæssigt kan klare inden for en overkommelig tid.

H. pylori er en lille bakterie der kan forårsage mavesår og derfor har været genstand for en del forskning. Der findes medicin mod *H. pylori*, men flere resistente bakteriestammer er kendte og den er desuden mistænkt for at spille en rolle i udviklingen af cancer. *P. falciparum* er en encellet parasit, der forårsager malaria i mennesker. Siden man i 1976 fik held til at dyrke *P. falciparum* i laboratorier er malariamedicinen blevet betragteligt forbedret, men sygdommen kræver stadig flere millioner menneskeliv om året på verdensplan. Resultaterne fra vores analyser er ganske interessante og opsummerede i Tabel 2.

Tabel 2. Vist i tabellen er estimerer for p, q og α . Kanter er det estimerede antal kanter i det fulde (uobserverede) interaktom. Antallet af knuder og kanter i de to PIN-datasæt

$$n_{t+1}(k) = \left(1 - \frac{1 + kp}{t}\right)n_t(k) + \frac{1 + (k-1)p}{t}n_t(k-1) + 2 \sum_{j=k-1}^t \binom{j}{k-1} \psi^k (1-\psi)^{j-k+1} \frac{n_t(j)}{t}$$

| | p | q | α | Kanter |
|----------------------|------|------|----------|--------|
| <i>H. pylori</i> | 0,56 | 0,05 | 0,78 | 5.636 |
| <i>P. falciparum</i> | 0,51 | 0,05 | 0,93 | 43.835 |

fremgår af Tabel 1. Estimaterne er opnået vha. avancerede statistiske metoder (primært Bayesian Statistics).

For det første ser vi at de to parametre der bestemmer genduplikation er meget ens i de to datasæt. Sandsynligheden (q) for at den nye og den kopierede (gamle) knude er forbundet er meget lille, hvilket er konsistent med hvad andre har fundet ved analyse af andre typer datasæt. Det peger også på at genduplikationens rolle er at skabe nye funktioner og ikke blot at modificere de gamle.

Derimod er sandsynligheden (p) for at både den nye og den gamle knude bevarer en kant til en knude ca. fifty-fifty og den ligger tæt på den grænse, hvor netværket under nogen omstændigheder ophører med at være stabilt (se næste afsnit). Endelig ser vi at parameteren α , der angiver hvor hyppigt genduplikation sker i forhold til tilknytninger, er meget forskellig i de to datasæt. *H. pylori* er som nævnt en bakterie, og bakterier udveksler ofte genetisk materiale mellem hinanden ved hjælp af de processer, der her er modelleret som tilknytninger. Derimod er *P. falciparum* en encellet eukaryot (den er grundlæggende opbygget som vores celler) og i eukaryoter er den slags processer langt sjældnere. Dvs. resultaterne stemmer overens med vores forventning.

Endvidere er de to organismers interaktomer også meget forskellige i størrelse, hvilket naturligt hænger sammen med hvor mange proteiner de hver især har (Tabel 1). Antallet af mulige kanter vokser kvadratisk med antallet af knuder, $N(N-1)/2$, og ratioen $(2M)/[N(N-1)]$ af faktiske kanter M til antallet af mulige kanter kan opfattes som udtryk for netværkets forbundenhed. I tilfældet *H. pylori* er ratioen ca. $5 \cdot 10^{-3}$ mens den er ca. $3 \cdot 10^{-3}$ for *P. falciparum*, dvs. bakterien er mere forbundet end eukaryoten. Også dette er i tråd med vores forventning, idet eukaryoten har mange varierede funktioner, der varetages af forskellige proteiner, mens bakteriens proteiner hyppigere er involveret i flere forskellige funktioner.

Matematiske resultater

Vi kan bevise en række udsagn om denne simple model. Endvidere bemærker vi at angivelsen af modellen svarer nøje til hvordan man kan simulere tilfældige grafer fra modellen, og at vi kan lære om modellen ved at simulere grafer og registrere deres egenskaber. Her vil vi kun interesser os for graden af en knude, dvs. hvor mange andre knuder den er knyttet til. Man kan opstille følgende rekursion for det forventede antal knuder $n_t(k)$ af grad k i grafen G :

hvor $\omega = (1+p)/2$ og vi for simpelheds skylds har antaget at $q=1$ og $\alpha=1$ (helt tilsvarende resultat kan opnås generelt blot med en del flere led).

Det første led kommer fra at betragte en knude A af grad k og beregne sandsynligheden for at den bevarer graden efter en duplikation; det sker, hvis ingen af dens naboer eller den selv udvælges (sandsynlighed $1-[1+k]/t$) eller hvis en af naboerne udvælges til duplikation, men ikke etablerer en kant til A (sandsynlighed $[1-p]k/t$). Tilsvarende fremkommer andet led ved at betragte en knude af grad $k-1$ og beregne sandsynligheden for at dens grad stiger med 1. Endelig angiver sidste led graden af den nye knude og knuden der duplikeres ('den gamle'): hvis den gamle knude har grad j får den nye og den gamle knude hver et binomialt fordelt antal kanter fra naboerne, samt en kant i mellem dem ($q=1$). Rekursionen er en konsekvens af Markov-egenskaben omtalt tidligere.

Rekursionen kan kun løses eksplisit i trivielle tilfælde, men det er i nogle tilfælde muligt at afgøre om hyppigheden af knuder af grad k stabiliseres som netværket



vokser, eller om der ikke indtræder en ligevægt i takt med at netværket bliver vilkårligt stort. For et vilkårligt parametervalg (α, p, q) har vi:

- Hvis $\alpha p < 0,5$ findes en ligevægt (ergodisk rekurrent).
- Hvis $\alpha = 1$ og $p < 0,533$ har et uendelig stort netværk *uen-delig* mange knuder af vilkårlig grad, men en ligevægt findes ikke nødvendigvis (rekurrent). Her er 0,533 den approksimative løsning til ligningen $2\log(\omega) + p = 0$.
- Hvis $\alpha = 1$ og $p > 0,562$ har et uendelig stort netværk *højst endelig* mange knuder af en vilkårlig grad, dog potentielt uendelig mange af grad nul (transient). Her er 0,562 den approksimative løsning til ligningen $1/(1-\omega) + 1/(p+\omega) - p - 2 = 0$.
 - Hvis $\alpha < 1$ har et uendelig stort netværk *uen-delig* mange knuder af vilkårlig grad, men en ligevægt findes ikke nødvendigvis (rekurrent).

Hvad der sker i det lille vindue mellem 0,533 og 0,562 ved vi ikke, ej heller om grænsen for en ligevægt er højere end 0,5. Simulationer er af lille hjælp her for selv hvis vi simulerede meget store netværk, ville vi ikke være i stand til at afgøre om graderne stabiliseres eller ej. Det første resultat hænger tæt sammen med den gennemsnitlige grad af en knude der er

$$\frac{2 - 2(1-q)\alpha}{1 - 2\alpha p}$$

forudsat $\alpha p < 0,5$ og ellers uendelig.

Biologisk er resultaterne interessante, idet de fortæller os om hvad vi skal forvente for et virkeligt netværk. Bringer evolutionen graden af en knude til en ligevægt? Ved hjælp af matematiske overvejelser bliver det således muligt at klassificere systemerne efter deres stabilitet. I de to tidligere eksempler finder vi således at de begge falder under betingelsen $\alpha < 1$, men også at parametrene for *H. pylori* ligger nær stabilitetsgrænsen (både α og p).

Afsluttende bemærkninger

Da modellen på flere måder er en tilnærmelse til virkeligheden, skal vi være påpasselige med at fremhæve resultaterne ukommenterede og man bør undersøge hvorvidt resultaterne er robuste over for forskellige modifikationer af modellen. Det er dog meget interessant at estimererne stemmer fint overens med vores forventning og viden fra andre typer data. Desværre er stadig kun få PIN-datasæt tilgængelige – med tiden kommer der forhåbentlig flere til. Nogle af de metodiske spørgsmål der trænger sig på er, hvorledes man kan analysere PIN-data fra tæt beslægtede arter, således at man tager højde for deres fælles evolution, og hvordan man kan analysere data fra arter der har samspillet gennem evolution. Et eksempel på sidstnævnte er parasitten *P. falciparum*s samspil med mennesket. I den forbindelse vil det være interessant at betragte farvede netværk, således at knuderne fx kan antage farve efter deres funktion. Det er sandsynligvis kun proteiner med visse funktioner i *P. falciparum*, der har tilpasset sig mere menneskelige omgivelser.

I denne artikel har jeg forsøgt at give et indblik i nogle af de problemstillinger, data og metoder moderne biologer og "anvendte" matematikere tumler med. Det er blevet spået at den anvendte matematik inden for biologi vil gennemgå en ekspllosion næsten af samme størrelse som

biologien selv har gennemgået. Det bliver næppe tilfældet, men interessant er det i hvert fald at opleve hvordan mange biologiske problemstillinger kan hjælpes godt på vej af matematik og at mange unge matematikere er begyndt at søge mod den anvendte matematik. Som "anvendt" matematiker hilser jeg naturligvis dette velkommen.

Efterskrift

Flere af mine studerende og postdocs har arbejdet med netværk og en speciel tak går til Oskar Hagberg, Oliver Ratmann og Michael Knudsen, som alle har ydet væsentlige bidrag. Jeg ønsker også at takke Michael Knudsen og Freddy Bugge Christiansen for kommentarer til manuskriptet og Enette Berndt Knudsen for hjælp med det stilistiske.

Litteratur i udvalg

Desværre er jeg ikke bekendt med populærvidenskabelige fremstillinger af emnet, dog er 'Linked' af Barabasi en glimrende populær introduktion til den omsiggrubende netværkstænkning. Den videnskabelige litteratur er rig og omfatter bidrag af matematisk, statistisk, fysisk og biologisk karakter. Nedenfor er udvalg med vægt på nogle af mine egne bidrag.

- A.-L. Barabasi (2003) *Linked*, Penguin Books.
 R. Durrett (2006) *Random Graph Dynamics*, Cambridge University Press.
 M. Lynch (2007) *The Origins of Genome Architecture*, Sinauer Press.
 S. Ohno (1970) *Evolution by Gene Duplication*, Springer Verlag.
 M. Knudsen og C. Wiuf (2008) A Markov chain approach to Randomly Grown Graphs, *Journal of Applied Mathematics* 2008: 190836.
 O. Ratmann *et al.* (2007) Using likelihood-free inference to compare evolutionary dynamics of the protein networks of *H. pylori* and *P. falciparum*. *PLoS Computational Biology* 3: e320.
 C. Wiuf *et al.* (2006) A likelihood approach to analysis of network data. *Proceedings of the National Academy of Science USA* 103: 7566-7570.
 DIP, <http://dip.doe-mbi.ucla.edu/> (proteininteraktionsdata).

Evolutionær tilpasning

af Freddy Bugge Christiansen
Biologisk Institut & BiRC, AU
freddy@biology.au.dk



Populationsgenetik støtter sig til et meget veludviklet teoretisk grundlag, som blev udviklet i begyndelsen af det tyvende århundrede. Anledningen var genopdagelsen af Mendels arvelove—eller rettere, at de blev anerkendt som den generelle basis for biologisk arv. Den umiddelbare ekspllosion i interessen for hvordan en populations samlede arvemasse videregives af fra forældre til afkom skyldes mangelen på en konsistent arvelære i Darwins evolutionslære—et problem, som bekymrede Darwin. Spørgsmålet var derfor om Mendels arvelove kunne lukke dette hul i evolutionslæren. Projektet lykkedes. Darwins evolutionslære transformeredes i løbet af 1930'erne til den såkaldte neodarwinistiske evolutionslære ved indarbejdelse af den mendelske genetik. Successen blev båret frem af kvantitative teorier formuleret af bl.a. de tre eminente genetikere R. A. Fisher, Sewall Wright og J. B. S. Haldane. Fisher producerede samtidig grundlæggende resultater inden for statistik. Selv matematikeren G. H. Hardy deltog og formulerede hvad der nu er kendt som Hardy–Weinberg loven, men resultatet også fundet af andre. For eksempel brugte Wright uafhængigt resultatet uden at formulere det som en lovmaessighed—det gik stærkt dengang.

Mendelsk genetik beskriver indirekte arv: Et individs egenskaber arves ikke, men de gener som påvirker egenskaben videregives til afkommet, som så har øget sandsynlighed for at udvikle en lignende egenskab. Virkningen af naturlig selektion er derfor dybest set, at de varianter af et gen, som findes i populationen, bliver sorteret. Selektion opstår når individer med forskellige egenskaber har forskellig sandsynlighed for at overleve og få afkom. Evolution ved naturlig selektion ændrer en popu-

lations udseende med hensyn til de arvelige egenskaber, som påvirker et individs antal afkom, men informationen, som videregives fra generation til generation findes i de gener, som videreregives. Uden selektion forventer vi ingen ændringer i populationens genetiske sammensætning. Mendels arvelove medfører derfor en grundlæggende konstans over tid i en populations arvelige variation. Denne variationsbevarelse var netop hvad Darwin manglede i evolutionslæren.

Genetisk Evolution

Det neodarwinistiske projekt sluttede ikke med gennembruddet i 1930rne. De dybere konsekvenser af genernes evolution viste sig at være mangfoldige, og forståelsen af genetisk evolution har lige siden været et centralt projekt i populationsgenetikken, navnlig den teoretiske populationsgenetik. Da jeg selv i 1960rne skiftede fra et studium i matematik og fysik til et i genetik var argumentet, at populationsgenetikken behøvede matematisk uddannede forskere. Dette var hverken første eller sidste gang et sådant behov opstod i en empirisk videnskab, men jeg var let at lokke på grund af min dybe interesse for biologi. Tidspunktet var heldigt: Slutningen af 1960rne og begyndelsen af 1970erne var en brydningstid, hvor den fulde konsekvens af den neodarwinistiske teori var ved at vise sig. Inden for dette nye univers vil jeg begrænse mig til et emne, som jeg brugte en del energi på, nemlig betydningen af genernes organisation i genomet for virknin gen af selektion.

Haldane beskrev den tidlige populationsgenetik som kugleposegenetik, fordi man studerede enkelte gener og

kombinerede resultaterne til forudsægelser om en organisms evolution—en massevirkningsmodel, hvor varianter af de enkelte gener forekommer uafhængigt af hinanden i individerne. Det gav resultater, som var i overensstemmelse med Darwins formodninger og virkede fint under den neodarwinistiske revolution. Fisher viste at selektion forårsaget af arvelig variation i sandsynligheden for overlevelse medførte, at populationen generation efter generation opnåede en længere gennemsnitlig overlevelse—den blev bedre og bedre tilpasset til det miljø, den lever i. Sådan selektion forårsaget af variation i ét gen giver en dynamik som i den klassiske fysiks læreforsøg: Det hele styres af en energi- eller lyapunovfunktion, i dette tilfælde populationens gennemsnitlige overlevelse. Fortolkningen er, at populationen gennem evolution ved naturlig selektion tilpasser sig det miljø hvori den lever.

Denne smukke dynamik bryder sammen så snart mere end ét gen varierer og producerer variation i individernes overlevelse. Uafhængighedsantagelsen i kugleposegenetikke redder dog næsten situationen. Den er, som vi skal se efterfølgende, konsistent med arvelovene såfremt vi kan antage, at den betragtede genetiske variation kun forårsager små forskelle i individernes overlevelse. Denne antagelse kaldes *svag selektion*, og den stemte med de forestillinger, som Fisher havde om virkningen af naturlig selektion. Evolutionslæren beskæftigede sig meget med kvantitative eller målbare egenskaber, fx kropshøjde, længde af næb og lignende. Beskrivelsen af sådanne egenskabers arvelighed var netop opnået ved at antage mange gener, hver med en lille effekt på egenskaben.



Simpel selektion: ét gen

Betræt et gen med to varianter A_0 og A_1 som i en given generation forekommer i hyppighederne p_0 og p_1 , $p_0 + p_1 = 1$. Antag tilfældig parring, så de tre genotyper A_0A_0 , A_0A_1 og A_1A_1 blandt nyfødte forekommer i hyppighederne p_0^2 , $2p_0p_1$ og p_1^2 . Sandsynlighederne for at de nyfødte bliver voksne og yngler er w_{00} , w_{01} og w_{11} ; disse kaldes genotypiske fitnesser. Antager vi at alle voksne individer forventes at få samme antal afkom, så bliver hyppigheden p'_1 af A_1 i næste generation

$$p'_1 = \frac{p_1^2 w_{11} + p_0 p_1 w_{01}}{W},$$

hvor W er den gennemsnitlige overlevelse i populationen:

$$W = p_0^2 w_{00} + 2p_0 p_1 w_{01} + p_1^2 w_{11}$$

(læg mærke til at $p_1 w_{11} + p_0 w_{01}$ er den gennemsnitlige overlevelse for A_1 -gener). I hver generation gælder at $W' \geq W$, hvor $W' = W$ hvis og kun hvis $p'_1 = p_1$.

Denne beskrivelse dækker en population hvor generationerne ikke overlapper. Fisher anvendte en kontinuert demografisk beskrivelse udviklet af H. T. J. Norton. Overlevelsandsynlighederne w erstattes her med malthusiske vækstrater m , så ændringen af genhyppighederne beskrives som

$$\begin{aligned}\frac{dp_1}{dt} &= p_1(p_1 m_{11} + p_0 m_{01} - M), \\ M &= p_0^2 m_{00} + 2p_0 p_1 m_{01} + p_1^2 m_{11}\end{aligned}$$

(læg mærke til at $p_1 m_{11} + p_0 m_{01}$ er den gennemsnitlige vækstrate for A_1 -gener). Disse ligninger viser sig at give en god beskrivelse for svag selektion, og resultatet er at

$$\begin{aligned}\frac{dM}{dt} &\geq 0, \quad \text{og} \\ \frac{dM}{dt} = 0 &\Leftrightarrow \frac{dp_1}{dt} = 0.\end{aligned}$$

Hastigheden hvormed M ændres er desuden proportional med den arvelige variation i fitness i populationen. Dette resultat kaldes Fishers fundamentalsætning for naturlig selektion.

Genetisk Kobling

Man vidste tidligt at de forskellige gener befandt sig på karakteristiske steder i cellens kromosomer, og at nabogener oftest videregives sammen—de viser kobling i transmissionen fra forældre til afkom. Man vidste også at kobling kan brydes ved overkrydsning mellem de to kromosomer i forbindelse med dannelsen af kønsceller (Figur 1), og at sandsynligheden for at en overkrydsning sker mellem to gener afhænger af afstanden mellem dem. Lad os i en given generation i populationen betragte hyppigheden, lad os sige x_{11} , af kønsceller, der indeholder varianterne A_1 og B_1 af generne betragtet i figuren. Sandsynligheden for at en kønscelle er en rekombinantgamet mellem de to gener er r , $0 \leq r \leq \frac{1}{2}$. Beregningen af hyppigheden i den følgende generation benytter sig af en opsplitning af populationens kønsceller, i parantalgameter og rekombinantgameter, og vi vil antage, at selektion ikke påvirker variationen i de to gener. Hyppigheden af A_1B_1 kromosomer blandt parantalgameterne er da x_{11} . Hyppigheden blandt rekombinantgameterne afhænger af hyppigheden hvormed et kromosom, der bærer A_1 , sidder sammen med et kromosom, der bærer B_1 i forældregenerationen. Vi behøver derfor kendskab til hyppigheden p_{A1} af kromosomer, der bærer A_1 , og til hyppigheden p_{B1} af kromosomer, der bærer B_1 i forældregenerationen. Uafhængighed mellem de to kromosomer i et forældre-individ er den simpleste antagelse, og den svarer til Hardys antagelse om tilfældig parring. Hyppigheden af A_1B_1 kromosomer blandt rekombinantgameterne er med denne antagelse $p_{A1}p_{B1}$, og hyppigheden af A_1B_1 kromosomer i afkomspopulationen er derfor

$$x'_{11} = (1 - r)x_{11} + rp_{A1}p_{B1}.$$

Overkrydsning tilfører derfor i hver generation en fraktionen r af gameter, hvor de to varianter forekommer uafhængigt af hinanden, og da hyppighederne af varianterne ikke ændres, får vi at

$$x_{11}^{(n)} - p_{A1}p_{B1} = (1 - r)^n(x_{11} - p_{A1}p_{B1}),$$

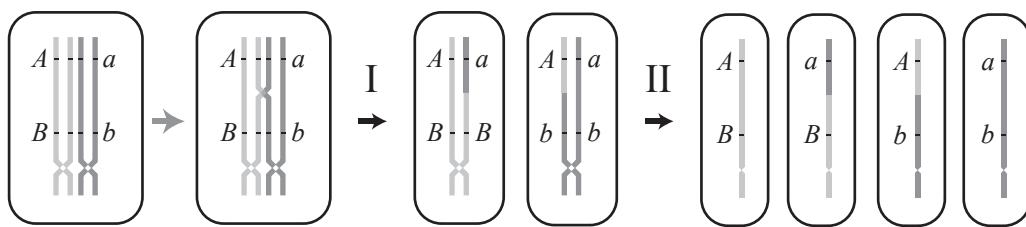
hvor hyppighederne af kromosomer, der bærer hhv. A_1 og B_1 i generation

n , stadig er p_{A1} og p_{B1} . De to genvarianter vil derfor ende med at forekomme uafhængigt i populationen idet

$$\lim_{n \rightarrow \infty} (x_{11}^{(n)} - p_{A1}p_{B1}) = 0.$$

Tilstanden hvor $x_{11} = p_{A1}p_{B1}$ kaldes koblingsligevægt, og antagelsen om svag selektion er netop, at populationen kan antages at være tæt på koblingsligevægt under den evolutionære process (Figur 2).

Selvom approksimationer, der baserer sig på antagelsen om svag selektion, er meget nyttige i mange evolutionære argumenter, så er det ikke hele historien. Samspillet mellem kobling og selektion er ikke påen. Betragt en population med kun A_0B_0 kromosomer, og overlevelsen i denne population beskrives ved sandsynligheden v_{00} . Kan en mutation B_1 i B genet invadere denne population? Det afhænger af overlevelsen v_{01} af A_0B_0/A_0B_1 individer, idet den sjældne mutant B_1 bogstavelig talt altid forekommer i et individ hvis andet kromosom bærer B_0 . Hyppigheden af B_1 vokser derfor når $v_{01} > v_{00}$. Ligeledes invaderer en mutation A_1 når $v_{10} > v_{00}$. Kan ingen af mutanterne invadere, når både v_{10} og v_{01} er mindre end den oprindelige genotypes overlevelse v_{00} ? Umiddelbart er svaret, at det afhænger af overlevelsen v_{11} af genotypen A_0B_0/A_1B_1 , men denne genotype dannes meget sjældent og burde derfor bagatelliseres, præsis som de genotyper, der blev negligeret i ovenstående argument. Den må dog forventes at blive dannet, men argumentet for det kræver en længere diskussion af balancen mellem produktion af varianter ved mutation og selektion mod dem. Antag derfor, at A_1B_1 kromosomet introduceres i populationen, og at $v_{01} < v_{00}$ og $v_{10} < v_{00}$. Når et individ af genotypen A_0B_0/A_1B_1 producerer kønsceller, vil kun fraktionen $1 - r$ af A_1B_1 kromosomerne overleve delingerne og videregives til næste generation. Betingelsen for at A_1B_1 hyppigheden i næste generation er større, $x'_{11} > x_{11}$, er derfor at $(1 - r)v_{11} > v_{00}$. Det er altså ikke nok at genotypen A_0B_0/A_1B_1 overlever bedre end A_0B_0/A_0B_0 , idet overkrydsning omdanner A_1B_1 kromosomerne til de unfavorable rekombinant-kromosomer A_1B_0 og A_0B_1 . Par af mutanter med lille effekt kan kun invadere en population, hvis de er me-



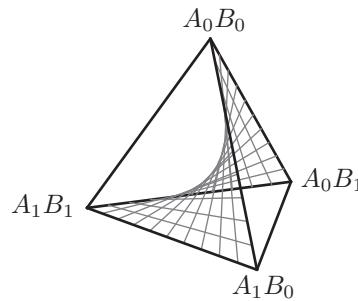
Figur 1: Kønscelledelingerne med en overkrydsning. To gener er markerede på et kromosompar i en celle klar til deling. Generne fra moderen er markeret med store bogstaver dem fra faderen med små. Kromosomerne fremstår som dobbeltstrengede, da kopieringen af deres DNA er afsluttet; de to strenge holdes sammen i et punkt. Ved overkrydsningen mellem de to gener brydes to kromosomstrenge på tilsvarende steder, og bruddene repareres, så de to afbrækkede stumper bytter plads. I første celledeling segregerer kromosomerne, og de resulterende celler gennemgår umiddelbart endnu en deling, hvor de to kromosomstrenge segregerer. Resultatet af overkrydsningen er to kønseller der ligner de, som dannede den oprindelige celle—de såkaldte parantalgameter—and to der afspejler overkrydsningen—de såkaldte rekombinantgameter.

get tæt koblede, og par af løst koblede gener skal give en dobbelt så stor sandsynlighed for overlevelse, hvis de skal invadere en population—and det er et grotesk krav. Resultatet er, at introduktion af A_1B_1 kromosomet med $v_{11} > v_{00}$ i lav hyppighed i populationen oftest medfører, at den gennemsnitlige overlevelse falder i de følgende generationer.

Dette eksempel kan virke specielt. Det er dog ikke patologisk, men både evolutionær relevant og en god illustration af den evolutionære virkning af genetisk kobling. Optimal tilpasning af populationen til det miljø, hvori den lever, er derfor ikke garanteret. Spørgsmålet i 1970erne var, om den evolutionære dynamik stadig var matematisk pæn, men allerede inden tiaret var omme viste matematikeren Ethan Akin, at det er den ikke. Ved en geometrisk analyse viste han, at i den metrik—den såkaldte shashahanimetrik—hvori Fishers resultater holder, dvs. hvori dynamikken kan beskrives ved et gradientfelt, kan dynamikken for selektion med genetisk overkrydsning ikke beskrives ved et gradientfelt. Endelig viste han, at resultatet er, at en ikke-trivial—hverken matematisk eller biologisk—delsmængde af modeller for selektion med genetisk overkrydsning vil tillade oscillerende dynamik, som ikke har et fortætningspunkt. Variation under selektion i flere gener kan derfor give en meget forvirret dynamik.

Akins resultat er relevant for intermediær kobling. For løst koblede gener svarer hans model til modellen for svag selektion, dvs. Fishers resultat gælder, når populationen ef-

ter kort tid er tæt på koblingsligevætsfladen (Figur 2). I den anden ekstrem—meget tæt kobling—gælder Fishers resultat også. For absolut kobling gælder det trivelt, idet de fire kromosomer A_0B_0 , A_0B_1 , A_1B_0 og A_1B_1 opfører sig som fire varianter af et kugleposegen. En lav overkrydsningshyppighed forårsager kun en trivel forstyrrelse af denne dynamik—lav hyppighed skal selvfølgelig forstås i relation til styrken af selektion. Selektion på mange gener spredt over et genom giver derfor en evolutionær dynamik af intermediær kompleksitet.

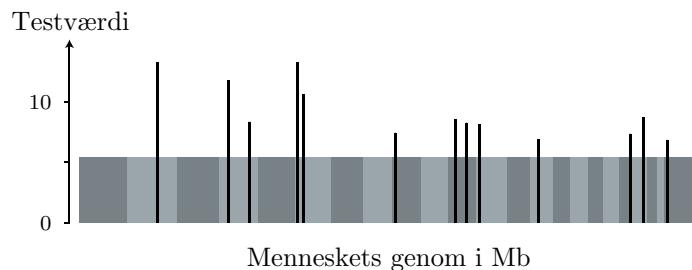


Figur 2. Hyppighederne af de fire kromosomer A_0B_0 , A_0B_1 , A_1B_0 og A_1B_1 summer til én. De kan derfor afbildes i et regulært tetraeder med højden én, så symmetrien mellem kromosomerne bevares. Den antydede flade er kromosomhyppigheder i koblingsligevægt, dvs. hyppighederne er $p_{A0}p_{B0}$, $p_{A0}p_{B1}$, $p_{A1}p_{B0}$ og $p_{A1}p_{B1}$, hvor $p_{A0} + p_{A1} = 1$ og $p_{B0} + p_{B1} = 1$. Med svag selektion vil hyppighederne hurtigt befinde sig tæt ved koblingsligevætsfladen, så effekten af selektionen i hver generation kan vurderes separat for de to gener.

Evolutionær Analyse

Evolutionær analyse af variation i kromosomerne DNA baserer sig ofte på, at vi inden for begrænsede længder af DNA-sekvenser kan antage absolut genetisk kobling, dvs. vi kan se bort fra overkrydsning. For længere strækninger må muligheden for overkrydsning indrages, hvilket umiddelbart giver en mere kompliceret analyse, men også åbner mulighed for ny insight i genomets evolution. Af speciel interesse i denne sammenhæng er fænomenet selektionsfejning. En ny mutation i et gen, hvis umiddelbare effekt giver en højere overlevelse, vil stige i hyppighed i populationen. Mutation er en sjælden begivenhed, så vi kan antage at det starter med ét stykke DNA med den nye mutation. Når mutantens hyppighed øges følger det omgivende DNA med, indtil det ved en overkrydsning frigøres. I det menneskelige genom svarer en sandsynlighed på 1% for en overkrydsning i et stykke kromosom til en fysisk afstand på omkring en million DNA-baser (en megabase eller 1Mb). Dette er bare et tal, men det kan måske blive mere konkret med oplysningen om, at et sådant stykke kan indeholde adskillige gener, selvom disse sidder meget spredt i genomet. Ventetiden til en overkrydsning inden for et sådant stykke DNA ved siden af mutanten er 100 generationer, hvilket for mennesker er i størrelsесordenen 3000 år. Denne overkrydsning kan ske hvor som helst i stykket, men som illustration lad os placere den midt i, så mutanten derefter er assosieret med 500kb af det oprindelige





Figur 3: Resultater fra et studium af genetisk forskelle mellem 12 regioner i Storbritanien. Data består af DNA-baserne i 500.000 positioner i genomerne i 17.000 individer. Abscissen er genomposition målt i antal baser med de 23 chromosomer lagt i rækkefølge efter hinanden, startende med kromosom 1 og sluttende med kromosom X. Testværdien er $-\log_{10} P$, hvor P er testsandsynligheden for det enkelte test for homogenitet mellem de 12 regioner. På grund af de mange test er testværdier under 5 uinteressante, og de mest interessante er vist med stolper. Genet for arvelig mælkesukkerintolerans er stolpen i kromosom 2. Nogle befolkningsgrupper har en lav tolerance over for mælkesukker og andre kan fordøje det. Forskellene i befolkningssammensætning i de 12 regioner giver det klare signal. Efter The Wellcome Trust Case Control Consortium, *Nature* 447: 661–678 (2007).

kromosom, og ventetiden til yderligere overkrydsning er nu 6000 år. Og så videre. For hver gang en overkrydsning sker, skal vi vente længere til den næste. Resultatet er, at efter mutanten har nået en høj hyppighed i populationen, vil den være associeret med et ikke ubetydelig stykke af det oprindelige DNA—givet selektionen er stærk nok.

De to genomer vi bærer på—et fra faderen og et fra moderen—er forskellige i omkring tre millioner basepar, eller én forskel per 1000 baser, men omkring vores mutation vil der være et interval, hvor alle eller næsten alle er ens, så der vil være et hul i den almindelige variation i genomet. Sådan et hul dannes, når selektionen fejrer variationen væk, og fænomenet kaldes derfor selektionsfejning.

Bruges genomvariationen til at notere sig positioner, hvor den er fe-

jet væk (Figur 3), så giver ét hul et meget klart signal, nemlig et interval omkring genet, som koder for det enzym, der spalter mælkesukker—som også er genet der giver arvelig mælkesukkerintolerans. I de fleste befolkninger ophører evnen til at udnytte mælkesukker, når barnet bliver vænnet fra, men vi nordvesteuropæere beholder også som voksne evnen til at fordøje det. Voksne i det meste af verden kan ikke omsætte mælkesukker, og de reagerer kraftigt mod det. Vi nordvesteuropæere er tolerante overfor mælkesukker, men vi ikke de eneste. Vi og andre befolkninger, som historisk baserer sig påavl af malkekævæg, viser et signal for selektionsfejning omkring dette gen. Tutsier i Afrika viser et klart fejningssignal, og en nærmere analyse af toleransgenet viser, at nordvesteuropæeres tolerance-variant og tutsiernes variant er forskellige. Evolution hen-

imod mælkesukkertolerans er derfor foregået uafhængigt i nordvesteuropæere og tutsiere.

Genernes organisation i genomet ødelægger altså den klassiske evolutionstese om organismernes universelle tilpasning til deres miljø. Virkningen af genetisk overkrydsning forstyrre virkningen af selektion. Dette sampsil giver os imidlertid et nyttigt redskab i vores søgen efter spor af naturlig selektion på genetisk variation i vores genom. Darwin indså, at virkningen af naturlig selektion er tab af arvelig variation. Hans problem var, at med datidens beskrivelse af biologisk arv tabte populationen variation også uden selektion. Den neodarvinistiske revolution fejede den generelle tilpasningstese væk, men påviste observerbare langtidsvirkninger i form af lokale tab af arvelig variation i genomet.



af Viggo Andreasen
RUC, viggo@ruc.dk

Matematiske modeller af influenzas epidemiologi og evolution

Som årsag til 500-2000 årlige dødsfald alene i Danmark er influenza en af vores vigtigste og farligste infektionssygdomme. Dertil kommer, at influenza giver anledning til stor sygelighed: hver eller hver anden vinter rammes 5-10% af befolkningen af influenza, og endelig forårsager influenza med årtiers mellemrum pandemier, dvs store verdensomspændende epidemier med yderligere forhøjet sygelighed og dædelighed. Influenzas epidemiologi er imidlertid meget usædvanlig og spredningen af influenza kan ikke umiddelbart beskrives med de traditionelle epidemiomodeller som f.eks. Kermack og MacKendrick's differentialligningsmodel.

For at forstå, hvad der gør influenzas epidemiologi så usædvanlig er det nødvendigt at se nærmere på influenzas biologi. Der findes 3 typer af influenza

virus A, B og C, hvoraf type A er den farligste, og den vi vil koncentrere os om. To forskellige slags proteiner stikker ud af influenza-virusets overflade, hæmagglutinin og neuraminidase. Begge proteiner er antogener, dvs strukturer som vort immunsystem genkender som fremmedlegemer. Under infektionen produceres specifikke antistoffer der kan bruges til at bekæmpe de pågældende organismer, hvis de måtte trænge ind i kroppen igen. Uheldigvis sidder den funktionelle del af det vigtigste antigen, hæmagglutinin, gernt i en lommeholte, sådan at den ikke kommer i direkte kontakt med immunsystemet. I stedet opbygges der antistoffer over for fem andre dele af hæmagglutinin, såkaldte epitoper. Da epitoperne ikke spiller nogen rolle for virus' funktioner, kan de variere uden at påvirke virus' evne til at reproducer sig. Der er to forskellige mekanismer som forårsager variation i hæmagglutinin: virusdrift og virusskift. Drift og skift er forbundet med hhv punktmutationer og reassortering hvor hele gen-

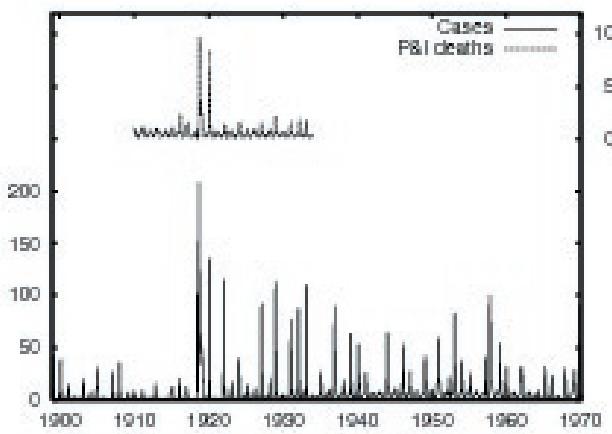


Figure 1: Influenza i Københavns Kommune 1899-1970. Antal influenza diagnoser per uge per 10.000 stillet af de praktiserende læger (fuldt optrukket linje og skala til venstre). Ugentlige dødsfald pr gr a. influenza, lungebetændelse og bronkitis per 10.000 (stiplet linje og skala til højre) [1]. Influenza-patienter dør ofte af følgesygdomme. Ikke influenza-relaterede dødsfald skal fjernes ved en tidsserie analyse, før man kan bestemme influenzens bidrag til dædeligheden.

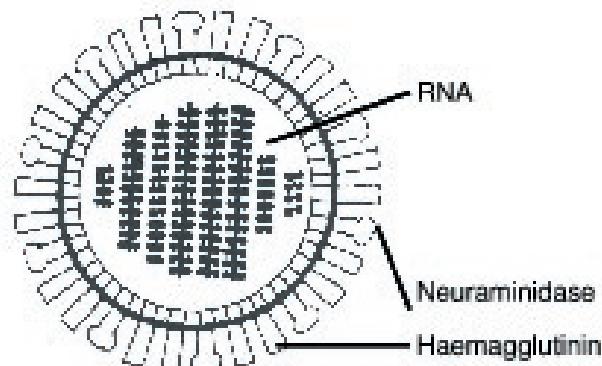


Figure 2: Influenza virus. Influenza virus består af 8 RNA stykker omgivet af en membran. To store proteiner, hæmagglutinin og neuraminidase stikker ud fra overfladen. Det menneskelige immunsystem genkender hæmagglutinin og i mindre omfang neuraminidase som fremmedlegemer, der skal bekæmpes.



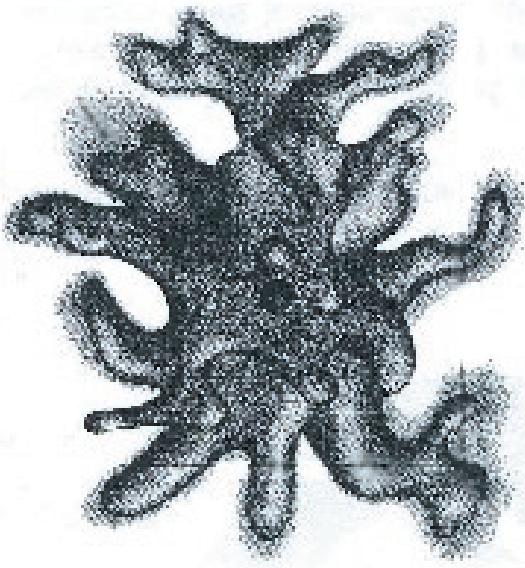


Figure 3: Quasi-art. En virus-“art” er en sky af beslægtede virus i rummet af gen-sekvenser. Virus i midten af skyen har højst fitness mens virus i udkanten klarer sig dårligere. Skyens udkant tilføres stadig nye virus fra centeret p gr a den høje mutationrate. For stabile vira, f.eks. mæslinge-virus, ligger skyens midtpunkt fast, men for influenza-virus bevæger midtpunktet sig i rummet af gen-sekvenser [5].

stykker udskiftes [2–4]. De giver anledning til meget forskellige modelleringssproblemer. Vi vil starte med modelleringen virusdrift, som kræver mest matematisk nytænkning.

Virus drift: influenzas evolutionære epidemiologi.

Den normale influenza, der optræder med jævnlige vinterepidemier, er et produkt af virusdrift og dermed af små genetiske ændringer i virus. En kvantitativ beskrivelse af fænomenet skal derfor kombinere matematiske modeller for infektionsepidemiologi med modeller for virus populationsgenetik.

Den type virus, som influenza tilhører, er de hurtigst muterende organismer, man kender. Langt de fleste mutationer er skadelige, men det store antal virus-partikler betyder, at der til stødighed også bliver dannet nok virus med høj fitness. Influenza er således en “quasi-art”, dvs en sky af beslægtede virus, der holdes sammen i en平衡 mellem mu-

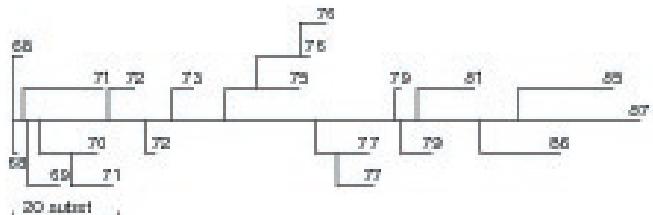


Figure 4: Phylogeni for influenza A. Endepunkterne i træet repræsenterer indsamlede influenza virus. Træets grene viser den simplest tænkelige forklaring på, hvordan disse virus kan være opstillet ved mutationer (substitutioner) ud fra en fælles stamfar, træets rod. Forgreninger svarer til hypotetiske virus, der er fælles ophav til virus på de to efterfølgende grene, mens grenenes vandrette længde angiver antallet af enkelt-mutationer, der adskiller de to virus ved grenens ender. Influenza's stamtræ vokser i længden med konstant hastighed mens sidegrenene forbliver korte. Influenza generngår jævnligt flaskehalse, hvor kun efterkommere af en enkelt variant overlever. Tallene yderst ved grenspidserne angiver indsamlingsåret [6].

tation og selektion mod mindre egnede mutanter med centrum i området med de mest fitte vira [5]. I M. Eigens oprindelige model for sådanne quasiarter er de enkelte virus-varianters fitness konstant, bestemt af variantens funktionelle egenskaber. Det specielle ved influenza er, at mutationer i det gen, som bestemmer hämagglutininets sammenstættning, kan føre til mutanter med højere fitness end deres oprindelse, og at størrelsen af denne fitness kan variere. Hvis mutationen ændrer en af epitopernes udseende, vil det nemlig ikke ændre virus even til at fungere, men den nye mutant vil kunne inficere individer, som var immune overfor det oprindelige virus. Derved får mutanten højere fitness, og skyens centrum flytter sig. Efterhånden som de modtagelige værter er blevet inficeret og immune, falder fitness for det nye “centrum,” og nye mutanter tager over.

Denne proces, virusdrift, kan aflæses direkte i de virus der rutinemæssigtindsamlies over alt i verden. Hvis man ser på ændringer i RNA-sekvensen for det gen-stykke, som koder for de vigtigste epitoper i hämagglutininet, kan man se, at der ofte kun er en enkelt mutant, som bliver stamfader alle senere stammer, figur 4, [6] mens alle vira som cirkulere på et bestemt tidspunkt er ganske nært beslægtede [7]. Virus driften kan også aflæses i det menneskelige immunrespons [3, 8]. Tabellen i figur 5 viser, hvor kraftigt antistoffer mod en bestemt virus-stamme

| Virus | Serum efter infektion med | | | | | | | |
|--------------------|---------------------------|--------|--------|--------|--------|--------|--------|--------|
| | H K/68 | Eng/72 | Vic/75 | Tex/77 | B k/79 | Phi/82 | Mis/85 | Sha/87 |
| A/Hong Kong/1/68 | 1280 | 320 | - | - | - | - | - | - |
| A/England/42/72 | 40 | 640 | 40 | - | - | - | - | - |
| A/Victoria/3/75 | - | - | 640 | - | - | - | - | - |
| A/Texas/1/77 | 40 | 40 | 80 | 1280 | 320 | 160 | - | 40 |
| A/Bangkok/1/79 | - | 40 | 40 | 320 | 1280 | 160 | - | 80 |
| A/Philippines/2/89 | - | - | 40 | 80 | 80 | 640 | 80 | 160 |
| A/Mississippi/1/85 | - | - | - | 40 | 80 | 80 | 1280 | 160 |
| A/Shanghai/11/87 | - | 40 | - | 40 | 80 | 80 | 40 | 640 |

Figure 5: Hæmagglutination af H3N2 influenza A virus. Influenza virus binder naturligt til de røde blodlegemer (hæmagglutinering), men processen hindres af antistoffer mod den pågældende virus. Tabellen viser hvor meget koncentrationen af antistoffer skal sænkes før der sker hæmagglutinering. Rækkerne giver de forskellige virusstammer - det sidste tal i betegnelsen angiver indsamlingsår. Tabellens søjler viser de vira, der har dannet antistofferne. Antistofferne reagere bedst mod de tilsvarende vira og reaktionen aftager med hvor lang tid der er gået mellem indsamlingen af de to virus [3].

reagerer ved mødet med en anden stamme. Som det fremgår falder immunreaktionen, hvis antistofferne svarer til virus, som er mere end 2-3 år ældre end den inficerende virus [9].

C Pease formulerede i 1987 den første model af influenzadrift, som giver en konsistent beskrivelse, der kobler disse genetiske og antigeniske ændringer sammen med influenzas epidemiologi [13]. Pease tog udgangspunkt i den velkendte Kermack-MacKendrik model (SIR-modellen på figur 6), der giver en udmærket beskrivelse af forløbet af en enkelt influenzaepidemi. I den klassiske Kermack-MacKendrik model standser epidemien, fordi der ikke er tilstrækkelig mange modtagelige at smitte. Hvis man vil bruge modellen over længere perioder, må man derfor gøre rede for, hvordan der tilføres nye modtagelige individer. For børnesygdomme som f.eks. mæslinger fremkommer nye modtagelige individer gennem fødsler. Nyfødte er også modtagelige over for influenza, men de spiller en mindre rolle, og Pease negligerer simpelthen fødsler. Modtagelige opstår først og fremmest ved, at individer taber deres immunitet, fordi den er dannet ved en infektion, som ligger for langt fra den aktuelle virus-stamme. Desuden antager Pease, at der til enhver tid kun er én virusstamme til stede i befolkningen, idet han ser bort fra de dimensioner i Eigens virus-sky, som ikke handler om variation i antigenerne. Ideen er nu at karakterisere populationens immunitet relativt til den aktuelle virus stamme, sådan at $S(a, t)$ er antallet af individer, hvis sidste infektion blev forårsaget af et virus som adskilder sig med a nukleotider (mutationer) fra

den aktuelle stamme - hvor a opfattes som en kontinuert variabel. Baseret på empiriske data antager Pease videre, at modtageligheden over for den aktuelle virus er proportional med a , og at centrum i virus skyen bevæger sig med en konstant hastighed k . Den samlede model bliver nu en ikke-lineær integro-differentialligning:

$$\begin{aligned}\frac{dI}{dt} &= \beta \int_0^\infty a S(a, t) da - \nu I(t) \\ \frac{\partial S}{\partial t} + k \frac{\partial S}{\partial a} &= -a \beta S(a, t) \\ kS(t, 0) &= \nu I(t),\end{aligned}$$

hvor vi har udeladt begyndelsesbetingelserne. Pease estimerede parametrene i modellen og viste at den kan forklare influenzas epidemiologi, dvs sygdomshyppighed, hyppigheden af reinfectioner og hyppigheden af influenza-epidemier. Pease model viser dæmpede svingninger om en stabil ligevægt. Senere arbejder har vist, at hvis proportionaliteten mellem modtagelighed og antal mutationer siden sidste infektion erstattes af en sigmoid kurve, opstår der regelmæssige svingninger gennem en Hopf-bifurkation [14, 15]. I influenza-fortolkningen af modellen betragtes systemet gennem en "moving frame" i rummet af virus typer. Modellen kan også bruges på visse andre infektionssygdomme, hvis man opfatter den voksende modtagelighed som et udtryk for tab af immunitet. Denne udgave af modellen, som kan spores tilbage til Kermack og MacKendrick [16], har fornødig været brugt af de amerikanske sundhedsmyndigheder til at vurdere

type i bliver nu

$$\Lambda^i = \beta_i \sum_{J \subseteq K \setminus i} f_J^i.$$

Idet vi går ud fra at even til at smitte ikke bliver påvirket af den smittedes immunitetsforhold. Modtageligheden over for virus type i afhænger af hvilke vira værten tidligere har været eksponeret for og kan beskrives ved en faktor $0 \leq \sigma_J^i \leq 1$, som skal fastlægges sådan, at den beskriver relationerne svarende til tabellen i figur 5. Idet vi for fuldstændighedens skyld inkluderer fødsel og død (med raten μ) bliver modellen nu

$$\begin{aligned} dS_J/dt &= \sum_{j \in J} \nu I_{J \setminus j}^j - \sum_{i \notin J} \sigma_J^i \Lambda^i S_J - \mu S_J \\ dI_J^i &= \sigma_J^i \Lambda^i S_J - (\mu + \nu) I_J^i. \end{aligned}$$

med oplagte modifikationer for $J = \emptyset$. I ovenstående ligninger mangler der desuden en beskrivelse af punktmutationerne som kan beskrives med små flow mellem klasserne af inficerede individer.

Modellen kan kun undersøges analytisk for små værdier af n og med stærke antagelser om symmetri af σ_J^i ; men allerede her kan modellen give anledning til ganske interessant dynamik [20–24]. Den blev oprindeligt formuleret med henblik på influenzadrift, men har i sær fundet anvendelse i beskrivelsen af malaria, dengue og kolera [25–28].

For influenza ved vi, at n er nødt til at være forholdsvis stor (typisk $n = 2^8$) for, at man kan indfange driftfænomenet, hvilket betyder at man er nødt til at gå over til computersimuleringer med den diskrete stokastiske analog, dels på grund af beregningernes størrelse, dels fordi der bliver så få individer i hver klasse, at man ikke kan betragte klasserne som kontinuerte variablene. Monte Carlo simuleringer med den diskretiserede model viser, at den ikke alene kan forklare influenzadrift. Krydssimmuniteten mellem beslægtede stammer givet ved σ_J^i er ikke i sig selv nok til at forhindre, at der opstår flere levedygtige afkom fra den samme oprindelse så skyen breder sig og bliver til en tåge. I arbejdet med computersimuleringer er der blevet foreslæbt to mekanismer som kan holde skyen samlet. Hvis inficerede personer ikke er modtagelige over en anden virus infektion de første 4–6 måneder efter endt sygdom, vil skyen holde sig samlet. Der findes ingen empirisk evidens for denne antagelse, men hvis undertrykkelsen af reinfektioner ikke er for voldsom vil den også være svær at igagtage [29, 30]. Skyen kan

også holdes samlet hvis de fleste mutationer producerer næsten identiske vira, mens meget sjældne mutanter er meget anderledes i deres antigene egenskaber [31].

Ideen med at indføre to forskellige typer mutationer er ny og er aldrig blevet undersøgt analytisk (men understøttes af molekylærbiologiske undersøgelser) [31]. Antagelsen om at reinfektioner bliver undertrykt, kan også fanges i en simple model hvor man undersøger om to forskellige vira begge kan forårsage epidemier i en population, der allerede er i ligevægt i forhold til Pease-modellen [32, 33]. Her viser det sig at influenzas smitsomhed være forholdsvis lav ($R_0 \approx 2 – 4$) eller også skal undertrykkelsen af reinfektioner være kraftig. Andre undersøgelser viser, at influenza's R_0 antagelig er så lav som 2–3 [34–36], så der er pt ingen afklaring på, hvordan skyen holdes sammen.

Selv driftprocessen kan beskrives analytisk som en diffusion i "rummet af mulige antigen-typer" ved at modificere Pease's model. Med passende parameterværdier opnås en drift hastighed, som passer med den observerede. Hovedproblemet er man nødt til at give afkald på en præcis beskrivelse af hele den immunitet der allerede er opnået [19, 37, 38] eller man må antage at immunitet kun påvirker evnen til at smitte andre [39–41].

Influenza skift og pandemier

Med lange mellemrum – typisk nogle årtier – gennerører influenza et såkaldt skift. Ved et skift erstattes 2 eller 3 af influenza-virusets 8 RNA-stykker med tilsvarende RNA stykker fra fugleinfluenza (andefugle). Da genet for hæmagglutinin altid er blandt de udskiftede gener, har den humane population ingen immunitet over for den nye sybotype. Influenza skift giver derfor anledning til en serie på 2–3 voldsomme, verdensomspændende epidemier (pandemi). Under influenza-pandemien forsvinder den tidligere influenza subtype af årsager, som er dårligt forstået, således at der normal kun cirkulerer en subtype ad gangen (figur 7).

Der er stor praktisk interesse for pandemier. Influenza pandemier er voldsomme begivenheder, hvor 50–80% af befolkningen bliver smittet på kort tid, så pandemier belaster sundhedsvæsenet og kan forstyrre almindelige samfundsfunktioner. Desuden koster en typisk pandemi 3–5000 menneskeliv alene i Danmark. Det lyder måske ikke af meget sammenlignet med at en normal vinterepidemi forårsager 1–2000 dødsfald, men pandemier rammer aldersmæs-

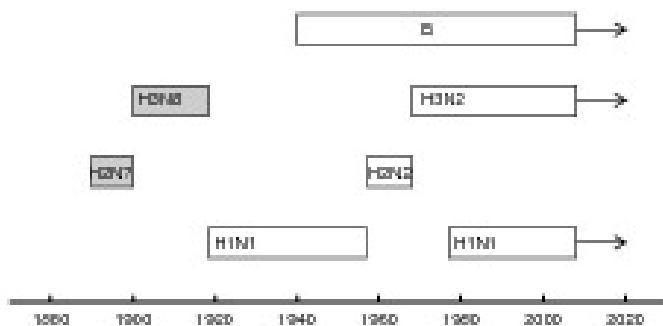


Figure 7: Influenza skift siden 1880. Figuren viser hvilke subtyper af human influenza, der har cirkuleret. Overgangen mellem subtyper er årsag til en verdenomspændende epidemi med den nye substype, en pandemi. Subtyperne betegnes med deres hæmagglutinin- og neuraminidase typer, men de har også "kælenavne". H2N2 kaldtes den "asiatiske" influenza, og H3N2 kaldes "Hong Kong" influenzaen. Der findes ikke virus-prøver fra perioden med de grå bokse, så subtypen er bestemt ud fra blodprøver. Skiftet i 1899/1900 kan kun ses i blodprøver og ikke i epidemiologiske optegnelser. H1N1 subtypen dukkede op igen i 1977 antagelig på grund af undslip fra et forskningslaboratorium [2, 42].

sigt bredere end vinterepidemierne som fortrinsvist rammer de ældre [43].

Da pandemier er sjældne, stammer vores erfaring fra historiske perioder, hvor den humane adfærd og demografi var markant annerledes end i dag. Matematiske modeller har været brugt til at forudse to aspekter af den næste pandemi, nemlig den globale spredning af det nye virus og den allertidligste lokale spredning i det område, hvor den nye subtype opstår [44, 45].

Den globale udbredelse af historiske pandemier er forholdsvis velbeskrevet. Man regner med, at de fleste pandemier starter i det sydlige Kina, hvor mennesker og andefugle er i tæt kontakt [46, 47]. Optegnelser i europæiske arkiver viser, at i det attende og nittende århundrede spredte pandemierne sig fra Rusland hurtigt langs Østersøen, lidt langsommere gennem Centraleuropa til Vesteuropa og derfra med skib til Nordamerika og i nogle tilfælde til Indien. Hvordan smitten er kommet fra Kina til Rusland vides ikke, men man antager at den er blevet spredt til Europa gennem Sibirien, mens Syd-øst-asien er blevet smittet direkte fra Kina. Den tyvende århundredes tre pandemier har fulgt et andet mønster. Den berømte "Spanske syge", op-

stod antagelig i foråret 1918 i nærheden af skyttegravene i Vesteuropa og spredtes til resten af verden med troppetransporter. De to seneste pandemier i 1956/57 og 1968/69 har formentlig spredt sig fra Kina med rutefly. De næede Europa og Nordamerika i løbet af 6-8 måneder. Spredningsmodeller (stokastiske kompartment modeller) der tager højde for stigningen i den internationale flytrafik viser, at den næste pandemi vil nå os meget hurtigere antagelig i løbet af 6-8 uger [48].

Det tager ca 6 måneder at producere en vaccine der beskytter mod en ny substype så masse-vaccinationer vil komme for sent til at få betydning for forløbet af den næste pandemi, selv hvis man forsøger at begrænse den internationale trafik. I stedet har to forskergrupper udviklet detaljerede modeller for at undersøge muligheden for at standse en kommende pandemi på et meget tidligt tidspunkt [44, 45].

Udgangspunktet for deres modeller er et scenarium hvor den nye virus bliver opdaget kort tid efter den er begyndt at spredes blandt mennesker. Ideen er nu at begrænse kontakten mellem personer i området ved at lukke skoler os samt at uddele anti-virus-medicin ("TamilFlu") som kan forhindre virus i at spredes. Modellerne viser at hvis virus ikke er særlig smitsom (R_0 skal være under 2) kan man faktisk standse virus spredningen og dermed købe tid til at få udviklet en vaccine. Hele historien virker imidlertid noget naiv. Erfaringerne fra den farlige 1918-epidemi er nemlig at virus var betydeligt mere smitsom (R_0 i København var f.eks. 4-5). Desuden forårsagede virus i de første 6 måneder ikke særligt mange dødsfald, så selv hvis man havde kunnet standse den tidlige epidemii, havde man nok ikke set nogen grund til at gribe ind, før den nye virus havde bredt sig over det meste af den industrialiserede verden.

De underliggende modeller er såkaldte "individ-baserede modeller." En slags stokastiske modeller, hvor hvert enkelt individ behandles særskilt, og hvor hvert individ kan komme ud for mange begivenheder med sandsynlighedsintensiteter, der kan afhænge af resten af populationens tilstand. Da modellerne bruges på store befolkninger som f.eks. Thailand, USA eller UK, er der tale om meget omfattende beregninger. Det er nemt at opskrive den slags modeller i moderne programmeringssprog som f.eks. C++ og programmerne egner sig glimrende til parallel-beregninger, men fra en matematikers synspunkt er modellerne så komplekse at det ikke er muligt at gøre dem til genstand for en



nærmere analyse.

Konklusion

Selv om modeller af influenzadrift har et grundvidenskabeligt sigte, er der en meget praktisk baggrund. Der foretages hvert år en omfattende vaccination mod influenza. I mange lande vaccineres op mod 70% af den voksne befolkning; i Danmark vaccineres færre og fortrinsvist ældre. Vaccinationerne skal getages hvert år, dels fordi vaccinens virkning forsvinder, dels fordi influenzadriften betyder at vaccinen muligvis ikke længere indeholder de antigener som svarer til den aktuelle influenzastamme. WHO anbefaler hvert år sammensætningen af den kommende vaccine til producenterne. Da vaccineproduktionen tager ca 6 måneder bliver den kommende sæsons vaccine-sammensætning for den nordlige halvkugle besluttet i januar-februar på baggrund af de virusvarianter som er blevet indsamlet på det tidspunkt. Denne procedure er bestemt ikke fejlfri; faktisk er WHOs anbefaling forkert i 1-2 ud af hver 10 sæsoner. De matematiske modeller kan naturligvis ikke i sig selv give bedre vaccineanbefalinger, men modellerne har skabt focus på samspillet mellem epidemiologi og naturlig selektion. Opdagelsen af at virusdrift antagelig foregår i små ryk med få års mellemrum blev f.eks. gjort af en forskergruppe, som længe har arbejdet med modellering af virusdrift.

Pandemimodellerne har et mere praktisk sigte og har karakter af begholderi-modeller hvor kendte fænomenes skaleres op til pandemi-niveau. Deres formål er altså at illustrere, hvordan en kommende pandemi kan udvikle sig. Variationer over disse modeller indgår da også i de "pandemiplaner," som alle industrialiserede lande har udarbejdet i de senere år ligesom modellerne har været forsidehistorier på såvel Nature som Science. Det er svært at peje på, hvor modellerne har frembragt ny erkendelse frem i en pandemi-sammenhæng, men tilsvarende modeller spillede en vigtig rolle under den store mund- og klovsygeepidemi i England i 2001. Modellerne viste nemlig at der gik for lang tid fra diagnose til nedslagtning [49]. Proceduren for nedslagtning blev ændret og epidemien kom kort efter under kontrol. For influenza kan en kombination af drift- og pandemimodeller måske bruges til kontrol af de efterfølgende belger i forbindelse med et influenzaskift.

Dette projekt er støttet af Forskningsrådet for Sundhed og Sygdom, projekt nr. 271-07-0555.

Referencer

- [1] Stadslægen i København, *Ugentlig oversigt over fødsler, sygdomme og dødsfald* (1898–1970).
- [2] N. J. Cox and K. Subbarao, *Annu. Rev. Med.* **51**, 407 (2000).
- [3] A. Hay, V. Gregory, A. Douglas, and Y. Lin, *Phil. Trans. Roy. Soc. London* **B356**, 1861 (2001).
- [4] D. J. D. Earn, J. Dushoff, and S. A. Levin, *Trends Ecol. Evol.* **17**, 334 (2002).
- [5] M. Eigen, *Scien. Amer.* **1993(7)**, 32 (1993).
- [6] W. M. Fitch, J. M. E. Leiter, X. Li, and P. Palese, *Proc. Natl. Acad. Sci. USA* **88**, 4270 (1991).
- [7] E. C. Holmes, E. Ghedin, N. Miller, et al., *PLoS Biol.* **3(9)**, e300 (2005).
- [8] D. J. Smith, A. S. Lapedes, J. C. de Jong, et al., *Science* **305**, 371 (2004).
- [9] C. B. Smith, N. J. Cox, K. Subbarao, L. H. Taber, and W. P. Glezen, *J. Infect. Dis.* **185**, 980 (2002).
- [10] W. O. Kermack and A. G. McKendrick, *Proc. Roy. Soc.* **115A**, 700 (1927).
- [11] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases* (Wiley, Chichester, 2000).
- [12] H. W. Hethcote, in *Mathematical Problems in Biology*, edited by P. van der Driessche (Springer-Verlag, Berlin, Heidelberg, New York, 1974), pp. 83–92.
- [13] C. M. Pease, *Theor. Pop. Biol.* **31**, 422 (1987).
- [14] H. Inaba, *Japan J. Indust. Appl. Math.* **18**, 273 (2001).
- [15] H. R. Thieme and J. L. Yang, *Math. Biosci.* **180**, 207 (2002).
- [16] W. O. Kermack and A. G. McKendrick, *Proc. Roy. Soc.* **141A**, 94 (1933).
- [17] H. W. Hethcote, *Math. Biosciences* **158**, 47 (1999).
- [18] B. T. Grenfell, O. G. Pybus, J. R. Gog, et al., *Science* **303**, 327 (2004).
- [19] V. Andreasen, J. Lin, and S. A. Levin, *J. Math. Biol.* **35**, 825 (1997).
- [20] J. H. P. Dawes and J. R. Gog, *J. Math. Biol.* **45**, 471 (2002).
- [21] J. Lin, V. Andreasen, and S. A. Levin, *Math. Biosci.* **162**, 33 (1999).
- [22] M. G. M. Gomes and G. F. Medley, in *Mathematical Approaches for emerging and reemerging diseases: Models, Methods, and Theory*, edited by C. Castillo-Chavez et al (IMA Volumes Mathematics and its Applications **126**, Springer-Verlag, New York, 2001), pp. 171–191.
- [23] L. J. Abu-Raddad and N. M. Ferguson, *Proc. Roy. Soc. London* **271**, 2431 (2004).
- [24] L. J. Abu-Raddad and N. M. Ferguson, *J. Math. Biol.* **50**, 531 (2005).
- [25] S. Gupta, J. Swinton, and R. M. Anderson, *Proc. Roy. Soc. London* **B256**, 231 (1994).
- [26] S. Gupta, N. Ferguson, and R. Anderson, *Science* **280**, 912 (1998).
- [27] N. Ferguson, R. Anderson, and S. Gupta, *Proc. Natl. Acad. Sci. USA* **96**, 790 (1999).
- [28] K. Koelle, M. Pascual, and M. Yunus, *Proc. Royal Soc. B-biological Sciences* **273**, 2879 (2006).
- [29] N. M. Ferguson, A. P. Galvani, and R. M. Bush, *Nature (London)* **422**, 428 (2003).

- [30] F. Tria, M. Lassig, L. Peliti, and S. Franz, *J. Stat. Mech. Theor. Experim.* **P**, 07008 (2005).
- [31] K. Koelle, S. Cobey, B. Grenfell, and M. Pascual, *Science* **314**, 1898 (2006).
- [32] V. Andreasen, *J. Math. Biol.* **46**, 504 (2003).
- [33] V. Andreasen and A. Sasaki, *Theor. Pop. Biol.* **70**, 164 (2006).
- [34] C. C. Spicer and C. J. Lawrence, *J. Hyg., Camb.* **93**, 105 (1984).
- [35] C. E. Mills, J. M. Robins, and M. Lipsitch, *Nature (London)* **432**, 904 (2004).
- [36] V. Andreasen, C. Viboud, and L. Simonsen, *J. Infect. Dis.* **197**, 270 (2008).
- [37] J. Lin, V. Andreasen, R. Casagrandi, and S. A. Levin, *J. theor. Biol.* **222**, 437 (2003).
- [38] J. Gog and B. Grenfell, *Proc. Natl. Acad. Sci. USA* **99**, 17209 (2002).
- [39] N. Ferguson and V. Andreasen, in *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods, and Theory*, edited by C. Castillo-Chavez et al (IMA Volumes Mathematics and its Applications **126**, Springer-Verlag, New York, 2002), pp. 157–169.
- [40] J. R. Gog and J. Swinton, *J. Math. Biol.* **44**, 169 (2002).
- [41] M. F. Boni, J. R. Gog, V. Andreasen, and F. B. Christiansen, *Theor. Pop. Biol.* **65**, 179 (2004).

REFERENCER

- [42] D. Buonagurio, S. Nakada, J. Parvin, M. Krystal, P. Palese, and W. Fitch, *Science* **232**, 980 (1986).
- [43] D. R. Olson, L. Simonsen, P. J. Edelson, and S. S. Morse, *Proc. Natl. Acad. Sci. USA* **102**, 11059 (2005).
- [44] N. M. Ferguson, D. A. T. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, and D. S. Burke, *Nature* **437**, 209 (2005).
- [45] I. M. Longini, A. Nizam, S. F. Xu, K. Ungchusak, W. Hanshaoworakul, D. A. T. Cummings, and M. E. Halloran, *Science* **309**, 1083 (2005).
- [46] C. W. Potter, *J. Appl. Microbiology* **91**, 572 (2001).
- [47] P. K. D., *Pandemic Influenza 1700–1900* (Rowman and Littlefield, 1986).
- [48] T. D. Hollingsworth, N. M. Ferguson, and R. M. Anderson, *Nature Med.* **12**, 497 (2006).
- [49] N. M. Ferguson, C. A. Donnelly, and R. M. Anderson, *Nature* **413**, 542 (2001).



DTU initiativet



af Michael Pedersen,
Institut for Matematik, DTU.
M.Pedersen@mat.dtu.dk

Anvendelse af matematiske modeller i biologi har en lang historie, og biologiske problemstillinger har, som fysiske – dog i en helt anden størrelsesorden - gennem tiderne inspireret matematikere i mange felter. Etablering som egentlig matematisk subdisciplin kan man måske tidsfæste til stiftelsen af *Society for Mathematical Biology* (SMB) i 1975, men man kan næppe sige at feltet fødtes med vind i sejlene. Fra at i starten være betragtet af mange som en lidt obskur, tenderende til primitiv, legeplads for en skøn blanding af matematikere, fysikere, kemikere og biologer, har feltet i det seneste tiår gennemgået en i alle henseender positiv udvikling. Med udnævnelsen af en højt anset matematiker som Avner Friedman som den seneste præsident for SMB, har man her klart signaleret at matematisk biologi er et seriøst felt i det matematiske spektrum, og på bevillingsområdet boomer feltet internationalt, fra NSF i USA over EU og helt til det tilsvarende kinesiske NSF-China satses umådelige summer på at få etableret forskningscentre og netværk med det formål at øge den matematiske forskningsindsats i alt fra genterapi til havøkologi. En ikke uvæsentlig motivation kommer oplagt fra den nyeste lovgivning i USA, der dikterer at ethvert lægemiddels (også indopererbare hjælpemidlers) virkning, for at kunne godkendes skal være dokumenteret ved en matematisk model med tilhørende (numerisk) simulation. Men den intensiverede klimadebat og øget fokus på økologiske katastrofer har bestemt også øget fokus på feltet. Dette har affødt et stort (også kommersIELT) behov for "biomathematicians" og et af formålene med de igangværende initiativer har været at få uddannet sådanne. En af de store forhindringer har her været sprogsbarrierer mellem biologer og matematikere, biologerne næsten uden kvantitative og analytiske færdigheder og matematikerne med en generisk forkærlighed for ultimativ abstraktion. De udenlandske erfaringer viser tydeligt at det er muligt at lære matematikere den nødvendige biologi, men næsten umuligt for biologer at lære nok matematik.



i Matematisk Bioscience

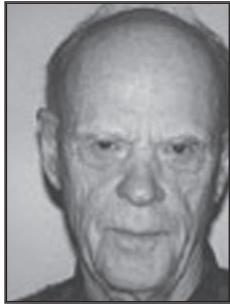
I Danmark har der på universitetsniveau kun været få tiltag inden for feltet, og i lyset af den internationale udvikling startede i 2006 på tværs af universiteterne det FTP-støttede forskningsnetværk MECOBS: "Modelling, Estimation and Control of Biotechnological Systems", som har til formål at undersøge mulighederne for at etablere forskning og uddannelse i feltet herhjemme. Jeg citerer fra hjemmesiden (www.mecobs.org):

"Mathematical modelling and subsequent mathematical analysis has for decades been the common and firm ground for fruitful cooperation between the various technical sciences, the major breakthroughs in one field often being the impetus of research in the other. Being a relatively new field, biotechnology and systems biology has not yet had the need for very advanced mathematical modelling and analysis, but the rapid development of production techniques as well as more refined experimental methods calls for more elaborate mathematics than typically comprised in the biotechnological environments themselves. From a strict mathematical point of view, the areas of biotechnology and systems biology are, to a large extent, using mathematical modelling and estimation based on both "imprecise definitions and inadequate methods". This is only a natural consequence of the fact that the fields have been driven so far by scientists with a background in chemistry and biology, but now the need for more advanced mathematical modelling and theory has become evident. This is in particular the case when the description of the systems in question is changed from a purely biological level to a description of the overall system dynamics and quantitative considerations. This also calls for the development of entirely new mathematical tools, leaving the hope that biotechnology and mathematics can develop the same synergy in the future that physics and mathematics has done for decades. By barrier-breaking workshops and knowledge-exchanging activities with the biotechnological and biological environments both nationally and internationally, the hope is to attract master- and Ph.D.-students to this interdisciplinary field and in this way establish a common ground for future research. The interdisciplinary nature of the subject is reflected in the theoretical background of the network members, and the initiative to co-operation with the scientific biotechnological environments at the universities and private companies is a major priority."

De matematiske modeller er typisk baseret på flere koblede, ikke-lineære ordinære eller partielle differentialligninger og helt konkret er der brug for specialister i PDE's, ODE's, dynamiske systemer, modellering, fysik, numerik, fluid dynamik, simulering, klassisk kontrolteori og uendeligtdimensional kontrolteori. Jeg har så ikke medtaget den statiske indfaldsvinkel, der forsøger at modellere systemerne stokastisk (som Markov kæder), idet man betragter dem som værende i forskellige tilstande med forskellige overgangssandsynligheder. Det giver anledning til stokastiske differentialligninger på en vis skala, men kan også modelleres ved Fisher-Kolmogorov eller Swift-Hohenberg ligninger (højereordens, ikke-lineære PDE's).

Matematisk bioscience er et særligt indsatsområde på DTU, i krydsfeltet mellem forskningsområderne anvendt funktionalanalyse og dynamiske systemer. Her startes eksamensprojekter, phd- og postdoc-forløb og området er søgt styrket ved stillingsopslag. I forbindelse med DTU's fusion med forskellige tidligere sektorforskningsinstitutter er der endvidere kommet særdeles relevante, nye samarbejdspartnere på banen. Man kan sige at dette område pr. definition ligger lige til højrebenet for DTU; den tværvidenskabelige naturvidenskab kondenseret i ingeniørvidenskab, poly-teknik par excellence, men som antydet er der problemyper på ethvert niveau og næsten i enhver matematisk disciplin - det svære er måske afgrænsningens kunst. Pointen er blot at alle de danske matematiske universitetsmiljøer burde kunne finde sin egen profil i dette umådelige univers af udfordringer, (der i øvrigt tilsyneladende tiltrækker en del studerende, som ikke traditionelt ville være startet på en matematisk-naturvidenskabelig uddannelse). På DTU er det "engineering"-aspektet, der er centralt, den traditionelle matematiske arbejdsmetode, hvor der abstraheres fra den konkrete model til den generelle eksistens- og entydighedsteori i passende funktionsrum skal nødvendigvis kulminere i kvantitative overvejelser, parametrene i modellerne skal estimeres og numeriske simulationer spiller en væsentlig rolle. Dette er DTU-vinklen, andre institutioner kan givetvis finde tilgange der passer bedre til deres lokale forskningstradition, hvis de vil.

Det er naturligvis håbet at DTU-initiativet vokser sig stort og at vi i samarbejde med andre universiteter kan få etableret et stærkt dansk forskningsmiljø i dette spændende felt, og vi står til rådighed med vores erfaringer for interesserede.



af Lars Gårding
Lunds Universitet
Lars.Garding@math.lu.se

Lämmelgåtan



Låt mig först tala om smågnagare, dvs sork och lämmel. Av båda finns olika arter på norra halvklotet. Dessa djur är små och bor i gångar under jord eller snö och äter växter, t.e.x mossor och gräs. De får också stora barnkullar och försvarar sig med sin stora fruktsamhet mot rovdjur som har dem som favoritföda, t.ex. räv, vessla, uggla och hök.

Någon gång under tjugotalet besökte den engelske zoologen Charles Elton¹ de norra delarna av Fennoskandia och upptäckte då att det med regelbundna tidsintervall mellan 3 och 5 år fanns många smågnagare men få däremellan. Detta var naturligtvis bekant för folk på platsen men Elton kallade fenomenet en kosmisk oscillation och upphöjde det till ett vetenskapligt problem: varför dessa svängningar, varför så regelbundna? När han kommit hem och funnit att smågnagarna där inte oscillerade regelbundet kunde problemet utvidgas med Och varför bara i Arktis? Alla dess frågor är det som jag kallar Lämmelgåtan.

Smågnagarforskningen i fält går till så att man någon fix tid på året sätter upp fällor och räknar hur många man får, eventuellt också kön och ålder. Om man samlar sina data från samma plats från år till år blir det statistik över en smågnagarpopulations storlek. Sådana data finns från Norge, Sverige och Finland och från den japanska ön Hokkaido. Detta utgör kontrollmaterialet för alla uppställda teorier.

Litteraturen på området är enorm och har gått i vågor. Just nu är det ganska lugnt. Det finns ett antal förklaringar till de cykliska fenomenen men ingen allmänt omfattad. I Norge tror man på en förklaring

som har hittats på av Seldal i Bergen. Han fann att då sork äter sin föda mycket intensivt så försvarar sig födan genom att bli litet giftig genom någon kemisk process. Detta tar en viss tid och detta förklarar visserligen svängningarna, men det är å andra sidan svårt att få fram cykellängden. Andra har hittat mera svävande förklaringar eller har pekat på än den ena än den andra faktorn i smågnagarnas liv och leverne. Man har bland annat framhållit (och jag måste tala engelska här) en *delayed density dependence* med en vag betydelse om att man kan förutsäga en populations storlek från ett år till ett annat genom att mäta förekomsten.

Naturens kretslopp, i detta fall en växlande förekomst av fisk i Tibern, väckte intresse i Italien omkring 1920. Matematikern Vito Volterra (1860-1940) var intresserad av matematikens tillämpningar och skrev 1927 en lång artikel om differentialekvationer som bland annat handlar om samspelet mellan två arter där en av dem äter och lever på den andra. Om den ena arten har N medlemmar och den andre M medlemmar så beskriver differentialekvationerna

$$\begin{aligned}N' &= aN + bNM \\M' &= cM + dMN\end{aligned}$$

arternas gemensamma öde. Här är M, N är funktioner av tiden och a, b, c, d är omsorgsfullt valda konstanter. Talet a betyder tillväxten hos N och c tillväxten hos M och b, d hur starkt och i vilken riktning de påverkar varandra. Genom val av dess konstanter fick han lösningskurvorna i M, N -planet att bli slutna och ringformade. Precis så som det är hos smågnagarna i Arktis. Samma ekvationer hittades även av en amerikan, Lotka², och det har lett till en åttio-

*Föredrag i Matematiska Sällskapet, Lund, 28.2.07

¹Charles Sutherland Elton (1900-1991) anses vara populationsekologins grundare. Periodic Fluctuations in the Number of Animals: Causes and Effects *British Journal of Experimental Biology* (1924) red. anm

²Alfred J. Lotka (1880-1949) ursprungligen från Lemberg





årig användning av dem i ekologi. Ekvationerna har inneburit en teoretisk framgång men de har medfört svårigheter med det numeriska och förklaringskraften är begränsad. Dessutom fordrar de fyra konstanterna många uppgifter om både ätare och föda utanför interaktionen.

Min ekologiske vän Sam Erlinge höll kort före sekelskiftet ett föredrag i Fysiografiska Sällskapet om en nordisk-rysk expedition till Sibirien med uppgift bland annat att studera den sibiriska lämmeln som är litet raskare än den skandinaviska med en kortare cykellängd på omkring 3.5 år. De såg en gång hur en lämmel som vågat sig upp i ljuset darrade av rädsla för rovdjur. Efter föredraget frågade jag Sam om det fanns en bra matematisk modell för lämmelcykeln och han svarade nej. Detta lockade mig och jag försökte hitta en modell.

Jag började först med differentialekvationer men slutade snart eftersom det var svårt att få cykellängden att passa. Så jag övergav de små stegen och beslöt att göra en beräkning från år till år och låta denna gälla ätarnas interaktion med födan. Det skulle ske så att ätarna antas få barn varje år som blev vuxna på ett år dessutom skulle vara så många att en viss del normalt skulle dö före årets slut. Jag bestämde att räkna ätare och föda i enheter sådana att varje enhet ätare skulle i medeltal förbruka en enhet föda per år. Jag bestämde att antalet enheter ätare E och antalet enheter föda F skulle räknas en gång varje år. I en första approximation skulle de vara oförändrade. Det blev början till följande formler

$$\begin{aligned} E &\rightarrow E + a(F - E) \\ F &\rightarrow F - E + C \end{aligned}$$

som förutsäger hur E och F ändras från år till år. Till höger äter ätarna E enheter föda och ett årligt tillskott på C enheter kommer av växtligheten. Till vänster händer ingenting då $F = E$ och ätarna har precis så mycket de behöver för året för att leva, dö och växa till med nya. Talet a är en positiv konstant. Den betyder att ätarna ökar vid ett överskott av föda och minskar vid ett underskott. Termen skall alltså betyda överskott eller underskott av ätare per

(Lviv). Känd för Lotkas lag eller *maximum power principle*
red. anm

enhets föda eftersom $F - E$ kan betyda enheter föda. Om a är födelsetalet per enhet ätare så kan man säga att $a(F - E)$ är antalet överlevande ungar från födelseöverskottet och motsvarande vid underskott. Medan födelsetalet är svårt att mäta i fält, betyder den cykiska rörelsen att man ta medeltal, t.ex. ätarnas genomsnittsalder L år. Då förnyas en genomsnittspopulation ätare efter L år och alltså tillkommer en $1/L$ -tedel per år, dvs genomsnittsfödelsetalet är $1/L$ och det tar vi som en definition av a . Jag fick höra av min expert Sam Erlinge att en lämmel lever mellan 4 och 8 månader så att jag bestämde att $a = 2$. Återstår matematiken.

Man ser att jämvikt i formlerna ovan betyder att $E = F = C$ och att iterationens banor kretsar kring denna punkt. Inför man variablerna

$$x = E - C, y = F - C$$

som passar bra i matematiska sällskapet ser man att iterationssteget blir homogent

$$\begin{aligned} x &\rightarrow (1 - a)x + ay \\ y &\rightarrow -x + y \end{aligned}$$

Nu frågar man sig *hur* ser iterationsbanorna ut? Det kan man se genom att avbildningen ovan har en invariant kvadratisk form

$$Q(x, y) = x^2 - axy + ay^2.$$

Med papper och penna kan man verifiera detta själv och dessutom visa att denna form är definit pris då $0 < a < 4$ vilket vi antar i fortsättningen. Då ser man att iterationsbanorna ligger på likformiga ellipser med centrum i C, C . Om $a > 4$ blir banorna hyperbler och vi kan kanske tänka oss gräshoppor i stället för smågnagare. Nu går vi vidare med litet linjär algebra. Med $X = \begin{pmatrix} x \\ y \end{pmatrix}$ och $A = \begin{pmatrix} 1-a & a \\ -1 & 1 \end{pmatrix}$, kan iterationssteget skrivas i matrisform

$$X \rightarrow AX$$

som kan upprepas genom att man tar potenser av A , dvs $X \rightarrow A^2 X, X \rightarrow A^3 X$ osv. Vi ska nu beräkna A 's egenvärden genom att ta determinanten av $A - \lambda I$



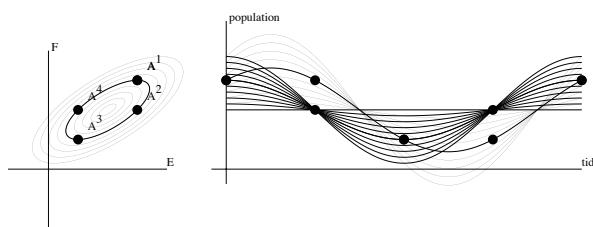
Sork

och räkna ut nollställena. Den blir $\lambda^2 - \lambda(2 - a) + 1$. Rötterna är $1 - a/2 \pm \dots$ Mer behöver vi inte räkna eftersom rötterna är konjugerade med produkten 1 och alltså ligger på enhetscirkeln och har formen $e^{\pm iv}$ där $v = \arccos(1 - a/2)$. (här ser man varför $a < 4$) och cykellängden är $360/v$. Här kan vi sätta in lämmelns $a = 2$ och får $v = 90$ grader och cykellängden 4 år. Sorken lever litet längre och får cykellängden 5 år som den ska vara. Vi kan nu vila oss efter att ha löst halva lämmelgåtan: Arktisk interaktion med födan ger rätt cykellängd.

Men nu ska vi göra en paus med den kanadensiska snöskoharen som har tioårig cykellängd i sin interaktion med en stor katt, lo. Loks genomsnittliga livslängd, uppmätt av zoologer i Kanada, passar precis denna cykellängd. Sambandet kan utläsas av räkenskaperna från artonhundratalet hos The Hudson Bay Company som anställde trappers som fångade lo och hare. Men det finns en hake, nämligen interaktionen hare-föda. Men den kan inte vara viktig eftersom haren har lätt att förflytta sig och föder sina ungar i tillfälliga bon på snön. Slut på pausen.

Men inverkar inte de som äter lämmel på cykellängden? Jo, litet grand. De är vesslor, rävar och den stora lappugglan. Men de har mycket längre genomsittliga livslängder än lämmeln som ger små värden av a under 1 och inte inverkar nämnvärt. Dessutom är deras inverkan inräknad i lämmelns genomsnittliga livslängd. Så är den saken ut världen. Men Arktis då? Utanför Arktis i tempererade zoner finns inte lämmel men sork. Där har de det gott jämförelsevis och har inte den stränga interaktion med mat som i gångarna i Arktis under långa vintermånader. Och är interaktionen svag så blir cykeln ojämnn. Ingen cykel har heller observerats. Alltså har vi hittat svaret på frågan Varför bara i Arktis? Jo, bara där domineras smågnagarnas tillvaro av tillgången på mat, dvs av interaktionen med föda. Vi har alltså löst lemmelgåtan.

Men: det finns andra lösningar. I Norge tror man på Seldal och i Finland på en annan professor. Vad man tror och inte tror har blivit en politisk fråga. Så om vår (min och er) lösning inte blir trodd så är det kanske inte en lösning, bara en komplikation till av lämmelgåtan som bara ökar dess mystik.



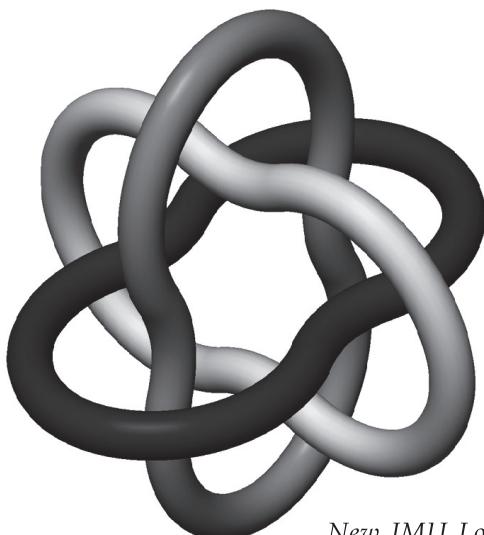
Referens: Lars Gårding *Interactions driving the population cycle of Arctic small rodents* Polish Journal of Ecology Vol. LIII No 4 2005. 579-584.





Christian Berg, formand

Beretning for 2007 om arbejdet i Nationalkomiteen for matematik



New IMU Logo based on the tight Borromean rings

Nationalkomiteen er bundeledet mellem dansk matematik og Den Internationale Matematiske Union (IMU) og har i 2007 haft følgende sammensætning: Henning Haahr Andersen, Christian Berg (formand), Morten Brøns, Hans Christian Hansen, Mikael Rørdam (sekretær), Michael Sørensen og Bent Ørsted.

Medlemskabet af de internationale faglige unioner som IMU ligger formelt i Videnskabernes Selskab, som nedsætter nationalkomiteer efter indstilling fra diverse foreninger, bl.a. DMF. Det er planlagt at oprette hjemmesider for nationalkomiteerne og vi regner med at vores komite får sin hjemmeside i løbet af foråret.

IMU's medlemslande er inddelt i 5 grupper I-V. Store lande som USA og Frankrig er medlem i gruppe V, Danmark og Finland er medlem i gruppe II og Sverige i gruppe IV. Tallet angiver hvor mange delegerede landet kan sende til generalforsamlingen og dermed hvor mange stemmer landet har. Kontingentets størrelse udregnes efter formlen $F(G)U$, hvor G er gruppen, $F(G)$ er en faktor og U er en enhed i Schweizer Franc. For 2008 er $U=1455$ CHF. Faktorerne $F(G)$ er fastsat til $F(I)=1$, $F(II)=2$, men derefter er der en progression $F(III)=4$, $F(IV)=8$, $F(V)=12$.

Iran og Korea er i 2007 rykket op fra gruppe II til henholdsvis gruppe III og IV. Nationalkomiteen har stemt for disse forslag. Efter afstemning i begyndelsen af 2008 er Norge rykket op fra gruppe II til III, og det må

give anledning til overvejelse om Danmark skal forblive i gruppe II eller søge om flytning til gruppe III. Det vil i givet fald betyde en fordobling af vores kontingent fra cirka DKK 13.500 i 2008 til 27.000. I øjeblikket dækkes vores kontingent og en række andre naturvidenskabelige fags kontingenter via en bevilling fra FNU til Videnskabernes Selskab, som sørger for det praktiske. Det sundhedsvidenskabelige forskningsråd har for et par år siden besluttet ikke at betale kontingenter for medlemskab af de internationale unioner, og overladt det til de faglige miljøer. Hvis noget sådant også sker i FNU, må finansieringen findes andetsteds. Det vil være en ubetydelig udgift for DMF, hvis den skal påtage sig opgaven. Derfor skal vi tænke os godt om før vi beslutter os til at søge om opgradering. Der henvises til IMU's hjemmeside, <http://www.mathunion.org> for yderligere information om medlemslandenes fordeling på grupper.

I 2007 har IMU optaget Ecuador og Kyrgyzstan som associerede medlemmer. Det betyder at de får information om arbejdet i IMU, men betaler ikke kontingent og har ikke stemmeret.

Den næste ICMI konference (International Commission on Mathematical Instruction) finder sted i Monterrey, Mexico i juli 2008. Som forberedelse af ICMI's generalforsamling i forbindelse med konferencen blev nationalkomiteerne af IMU bedt om at indstille kandidater til diverse poster. Vi valgte at indstille Tine Wedege til ICMI's nomineringskomite, der som formand har Jeremy Kilpatrick, USA. Der blev trukket lod blandt de indstillede fra nationalkomiteerne. Valget faldt på Attia Ashour, Egypten og Evgenia Sendova, Bulgarien, så Tine kom desværre ikke på banen i denne omgang. Nomineringskomiteen har til opgave at forberede valget af ICMI's kommende Executive Committee.

En væsentlig del af IMU's arbejde er planlægningen af de internationale matematiker kongresser ICM (International Congress of Mathematicians), der finder sted hvert fjerde år, og ved samme lejlighed afholdes IMU's generalforsamling. Som meddelt sidste år finder næste ICM sted i 2010 i Hyderabad i Indien. Som led i planlægningen af det videnskabelige program er Hendrik Lenstra blevet formand for Program Committee. Han skal sammen med en række chairmen for de forskellige hovedområder i gang med at overveje, hvem der skal inviteres som foredragsholdere.

En bro mellem forskning og skole for voksnes matematiklæring

- et fælles nordisk projekt

af Projektgruppen ved Johan Forsell og Knud Søgaard

Med Åsö Vuxengymnasium, Stockholm som initiativtager og projektkoordinator påbegyndtes i efteråret 2007 et Nordplus Voksen-projekt. Projektet, som bærer navnet CORMEA (COnnecting Researchers and Mathematics Education for Adults), skal bygge bro over kløften mellem lærere i matematik og forskere i matematikkens didaktik med fokus på voksnes læring. Målet er at skabe mødesteder for lærere og forskere. Projektet afsluttes juli 2008, men gruppens klare hensigt er at generere et nyt projekt baseret på de indhøstede erfaringer.

Nordisk samarbejde

Projektet er fælles nordisk med deltagende organisationer fra Danmark, Norge og Sverige. Et bredt felt af erfaringer er repræsenteret. De fleste arbejder med voksenuddannelse mens andre er forskningsorienterede eller arbejder på nationalt plan med udvikling og koordinering af uddannelser. Fælles for alle deltagere er en speciel interesse for matematikundervisning for voksne, hvilket ikke på nogen måde udelukker erfaringer fra andre uddannelsesområder i matematik. Al erfaring med forsker- og lærerkontakter indenfor matematikuddannelsesområdet har projektets interesse.

Behovet for dialog mellem forskere og lærere

At der er behov for flere og mere aktive mødesteder/dialoger mellem matematiklærere og didaktikforskere er projektets deltagere overbevist om. En artikel i Skolvärlden 29/8 2007 under rubriken "Kan lärare och forskare gå varandra till mötes?" fokuserer på samme tema.

"Om lärare hade bättre koll på forskning, skulle det inte vara så många frälsningsläror i skolan. Broarna mellan forskare och lärare behöver bli fler. Lärare som inte har tid att läsa forskningsrapporter. Forskare som inte forskar om sådant som intresserar lärare. Ibland framställs forskarna och lärarna som två läger. Kanske tjänar den bilden ingenting till. Men

sant är att det bedrivs en massa forskning som inte når lärarna..."

Alle som arbejder med matematikundervisning for voksne ved, hvor kompleks og dynamisk læreprocessen er og at ingen metode alene løser alle knuder. En mere omfattende dialog mellem lærere og forskere ville kunne bidrage med at belyse disse komplikerede processer på en frugtbar måde. Deltagerne i projektet er overbevist om, at flere broer mellem lærere og forskere vil være en berigelse for begge parter.

Gode matematikkundskaber er vigtige for at kunne fungere godt i arbejdslivet. Mange voksne fornemmer, at deres matematikkundskaber er mangelfulde og møder undertiden store vanskeligheder, når de begynder med matematikstudier i voksenuddannelsen. Det er en udfordring for alle arrangører af voksenuddannelse at møde den studerende på rette niveau samt på alle måder at bidrage til et succesfuldt resultat. Dette forudsætter, at lærerne gives muligheder for at udvikle og uddybe deres kompetencer. Et grundlag for en sådan uddybning kan blandt andet være en mere aktiv dialog mellem forskere og de som praktiserer voksenuddannelsen. Den matematikdidaktiske forskning i Norden er under kraftig udvikling, men dialogen mellem forskere og lærere opleves at være begrænset.

Kortlægnings- og kontaktprojekt

Projektet har til formål at kortlægge allerede etablerede kontakter mellem matematiklærere og forskere og forskningsresultater samt at belyse behov og udviklingsmuligheder i forbindelse med disse kontakter. Hvordan formidles de didaktiske forskningsresultater, og hvordan bliver de til gavn for den studerende? Hvordan kan lærere, som møder matematikunderisningens komplekse (og stimulerende) daglige problemer, gives mulighed for at udpege interessante og nødvendige forskningsområder? Forhåbentligt kan projektdeltagerne skitsere modeller for, hvordan fremtidige mødesteder kan se ud samt generere et udviklingsprojekt med fokus på disse spørgsmål.

I alle tre deltagende lande Danmark, Norge og Sverige gælder, at matematiklærere som regel har et stort behov



Nogle af projektdeltagerne ved det første træf på DPU i København

for kompetenceudviklende aktiviteter. Sådanne eksisterer og handler ofte om at lærere deler ud af deres erfaringer, eller at et forlag eller en virksomhed informerer om sine bøger og produkter. Lærernes efteruddannelse handler i for ringe omfang om kommunikationen mellem lærere og forskere. Dette burde være et nøgleområde for udvikling, hvis matematikundervisningen for voksne skal forbedres.

Første træf mellem projektdeltagerne var i København på Danmarks Pædagogiske Universitetsskole (DPU) i slutningen af august 2007. Her blev retningslinierne for det etårige projekt tegnet.

Websted og spørgeskema

For at kunne gennemføre kortlægningen og for at så mange som muligt kan følge med og deltage aktivt har projektet sit websted <http://www.cormea.org>. Her findes bl.a. artikler og referencer vedrørende voksnes matematiklæring, yderligere information om projektet og aktivitetskalender.

Et andet nøgleværktøj i kortlægningen er et spørgeskema. Via dette forsøger projektet at nå ud til så mange interesserter som muligt. Spørgeskemaet kortlægger kontaktfladen mellem forskere og lærere i voksenuddannelsen. Hvad er erfaringerne, og hvad er behovene? Spørgeskemaet, som henvender sig både til forskere og lærere, kan besvares online på hjemmesiden. Samme sted vil det samlede resultat af undersøgelsen kunne læses.

Gennemførte og kommende aktiviteter

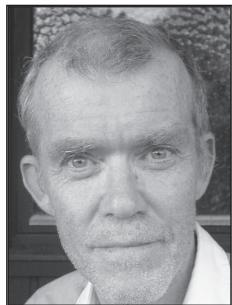
Projektet blev præsenteret på MADIF 6 (6. matematikdidaktiske forskningssseminar) i Stockholm 30. januar 2008 samt i forbindelse med to seminarer på Matematikbiennalen i Stockholm 1. februar 2008. I et af disse seminarer inviterede projektgruppen til diskussion om dialogen mellem forskere og lærere under temaet "Ut med didaktikken!". Det gav udbytterige diskussioner og nye kontakter. Udvilket projektet i forbindelse med konferencer samt i relevante medier og tidsskrifter, vil der blive etableret direkte kontakt med interessererde lærere og forskere.

Indbydelse til seminar i Oslo

Som et vigtigt led i at nå projektets målsætning afholdes et seminar i Oslo den 5. juni 2008. Indbudte nordiske forskere / udviklingspersoner afholder workshops for matematiklærere. Vi håber, at disse workshops vil føre til, at en række oplagte forskningsområder kan specificeres og dermed danne grundlag for en videreførelse af CORMEA som et udviklingsprojekt. Vi satser endvidere på at få nye kontakter og personer, som vil hjælpe med til at sprede erfaringerne eller som vil være aktører i det kommende udviklingsprojekt. Vi håber, at netop du har mulighed for at komme til Oslo. Brug projektets hjemmeside og hold dig opdateret om programmet.

Projektdeltagerne ser frem til en spændende afslutning for projektet med mange nye kontakter. Projektgruppen håber på, at mange interessererde (udbydere af voksenundervisning, lærere, skoleledere og forskere) vil deltage og aktivt bidrage til kortlægningen og til at udvikle den så nødvendige dialog.

Selberg i Normat



af Ulf Persson

Hösten 2005 intervjuade två norska matematiker - Nils A. Baas och Christian F. Skau från Trondheim- Atle Selberg i dennes kontor i Fuld Hall, Princeton¹. Intervjuerna skedde vid två olika tillfällen, vid det första togs intervjun upp på band, vid det senare videofilmades det hela. Totalt rör det sig om cirka tolv timmars upptagning. Delar av videoupptagningen har bildat underlaget för en dokumentär om Selberg som under hösten 2007 visades på norsk television.

I denna intervju ger Selberg sin egen version på den konflikt han hade med Erdős i samband med det elementära beviset för primtalssatsen sommaren 1948² ven om intresset för denna prioritetskonflikt må ha mattats något med åren, (det elementära beviset visade sig inöte vara det genombrott det vid den tiden förmodades vara), är det fortfarande historiskt intressant. Rykten om denna kontrovers har sedan länge varit i omlopp och kan numera även återfinnas i de populärmatematiska böcker som på senare år publicerats³.

Men Selberg talar om mycket annat, inte bara om andra matematiska resultat utan även om

sitt liv och ger många intressanta och klart formulerade synpunkter om både det ena och det andra. Intervjun har karaktären av en lång monolog ty de bågge intervjuarna har varit ytterst obenägna att bryta det naturliga flödet. Intervjun är att se som Selbergs testamente till eftervärlden. Det är givet att mycket av detta är högintressant, speciellt för skandinaviska matematiker och bör komma till allän kändedom, vilket också var Selbergs ambition. Publicering av stora delar av materialet är därmed angeläget. På grund av att det dels rör sig om en transkribering av ett talat språk, dels om den vändlande gången av en monolog kommer en relativt sträng redigering att vara av nöden, dock utan att i möjligaste mån inkränkta på Selbergs egna ord.

Man skall aldrig sälja skinnet förrän björnen är skjuten, men jag är personligen ganska övertygad om att Normats läsare skall snart kunna ta del av innehållet i intervjuerna, förhoppningsvis även kompletterade med artiklar med anknytning till Selbergs matematik. Denna notis är således avsedd att i tillägg till dess informationsvärde även utgöra reklam för att locka nya läsare (och prenumeranter) till Normat.

*Sedan 1 januari 2007 huvudredaktör för Normat

¹Detta kontor i Institutets huvudbyggnad tillhörde ursprungligen von Neumann, men övertogs av Selberg vid den förres död 1957.

²Elementärt i den meningens att inga tekniker från Komplex Analys utnyttjades. Det första beviset för primtalssatsen presenterades som bekant av Hadamard och de Vallé Poussin ett halvsekel tidigare - 1896. Atle Selberg erhöll Fields-medaljen i Cambridge (MA) ett par år senare, delvis motiverat av hans elementära bevis, och Erdős fick ungefär samtidigt Cole-priset för sin insats.

³Sautoy presenterar en relativt objektiv skildring i sin 'The Music of the Primes', medan Schechter i sin bok 'My brain is open' tar mera oförbehållsamt parti för Erdős.



Selberg



BASE – et kursus i matematisk modellering på den naturvidenskabelige basisuddannelse ved RUC



af Morten Blomhøj og Tinne Hoff Kjeldsen,
NSM, IMFUFA, RUC

1. Baggrund, organisering og evalueringskrav

BASE-kurset er et grundkursus i matematisk modellering ved RUC's naturvidenskabelige basisuddannelse (Nat-Bas). Formålet med kurset er at kvalificere de studerende til overbygningsstudier, hvor matematik anvendes integreret i modellering og databehandling bl.a. i forbindelse med eksperimentelt arbejde. Hovedparten af de studerende på BASE-kurset planlægger at læse andre fag end matematik og fysik i deres bachelor, og kurset er klart målrettet studerende med interesser inden for det biologiske område.

Kurset er en integreret del af "plottet" med RUC's naturvidenskabelige basisuddannelse. I hvert af basisuddannelsens fire semestre gennemfører de studerende i grupper af 4-6 personer et problemorienteret projektarbejde, som udgør halvdelen af arbejdsbelastningen i et semester. Den anden halvdel bruges på to kurser. I løbet af Nat-Bas skal man gennemføre mindst to matematikkurser og BASE, der er et dobbeltkursus, er her det naturlige valg for mange studerende. I mindst et af en studerendes fire projekter på Nat-Bas skal der indgå eksperimental arbejde. Projekterne i de første tre semestre er underlagt semesterbindinger, der sikrer belysning af naturvidenskab fra forskellige vinkler. I daglig tale hedder semesterbindingerne *med, i og om-naturvidenskab*, hvor *med* drejer sig om anvendelser af naturvidenskab i samfund og teknik, *i* handler om indre naturvidenskab, hvor det er sammenspillet mellem model, teori og eksperiment, der skal blyses, mens det i tredje semester er *om- eller meta-aspekter af naturvidenskab*, der skal undersøges.

Matematiske modeller spiller en væsentlig rolle i projektarbejdet, men erfaringerne viser, at de studerende, på trods af god forståelse af modellerne og deres relation til den problemstilling de arbejder med, og på trods af at de har de nødvendige tekniske matematiske forudsætninger, som oftest ikke er i stand til selv at bruge matematik til at opstille og analysere matematiske modeller. Disse og andre lignende erfaringer fra overbygningsstudier, hvor matematiske modeller indgår i anvendelser og teoriopbygning, tyder på at modelleringskompetence, altså kompetence til matematisk modellering og brug af matematiske modeller, ikke udvikles automatisk ved at lære matematik. For at imødegå dette problem og for derved at øge kvaliteten i projektarbejdet inden for modeller og modellering begyndte vi for snart ti år siden udviklingen af BASE.

BASE-kurset, som det foreligger i dag, er designet til at varetage følgende tre formål:

1. At vise, at matematisk modellering er relevant, ofte uundværlig, i studie og anvendelse af teori i naturvidenskaberne, og derved øge de studerendes værdættelse af matematik og motivere dem til at lære matematik og modellering.
2. At støtte de studerendes konstruktion af vigtige og fundationale matematiske begreber og metoder knyttet til differential- og integralregningen, differentialligninger og systemer af differentialligninger.

3. At bidrage til udviklingen af de studerendes modelleringsskompetence ved at gøre dem i stand til selvstændigt at opstille, analysere og kritisere matematiske modeller inden for det naturvidenskabelige område.

I denne artikel ser vi specielt på (2) og viser et eksempel på, hvordan de studerende kan få støtte til forståelse af centrale faglige begreber ved at arbejde med analyse af differentialaligningsmodeller. (1) og (3) tilgodeses i høj grad gennem samspil med semester-projekterne, hvor de studerende kan komme til at arbejde med hele modelleringssprocessen under et mere holistisk perspektiv.

Kursets matematiske emner omfatter lineær regression, datatransformation, differential- og integralregning, kompartmentsmodellering, ordinære differentialaligninger og systemer af differentialaligninger. Alle emner behandles og fortolkes gennem arbejdet med matematisk modellering, og de indledende emner tjener som forudsætning for opstilling og analyse af modeller for dynamiske systemer vha. differentialaligningssystemer i den sidste halvdel af kurset. Der anvendes både analytiske og numeriske metoder, og programmet MatLab er en integreret del af kurset, se kursusnoterne (Blomhøj, Kjeldsen og Ottesen, 2005).

Selv undervisningen foregår i hold af 25-40 studerende (et hold per overgang) og omfatter 48 kursusgange af 2.5 time fordelt med to kursusgange per uge over to semestre. Kurset er opbygget omkring 6 mini-projekter, der udføres i grupper af 2-4 studerende. Hvert mini-projekt løber over 2-3 uger, og der afsættes typisk svarende til 1½ kursusgang til vejledning af og arbejde med mini-projekterne. Hver gruppe kan vælge mellem 2-3 problemstillinger til hvert mini-projekt. Sammensætningen af grupperne varierer fra mini-projekt til mini-projekt, da de studerende ofte vælger mini-projekt efter interesse. Grupperne udarbejder rapporter af typisk 15 sider for hvert mini-projekt og fremlægger nogle af dem til diskussion i plenum eller over for en anden gruppe.

Kravene for at bestå kurset er (1) rettidigt afleverede grupperapporter over alle miniprojekterne. Er grupperapporterne ikke tilfredsstillende, får de studerende mulighed for at aflevere en forbedret udgave efter nærmere aftale med kursuslæreren. (2) Udarbejdelse af en portefølje med skriftlige opgaver. I hvert semester svarer porteføljemappen til 4-5 individuelle opgavesæt med mindre matematiserings- og modelleringsopgaver. (3) Gennemførelse med et tilfredsstillende resultat af en individuel skriftlig test i midten af hvert semester. (4) Gennemførelse af en individuel mundtlig prøve med et tilfredsstillende resultat ved afslutningen af hvert semester. Prøven tager enten udgangspunkt i den studerendes porteføljemappe eller i et af de tre mini-projekter, den studerende har lavet i semesteret. Valget foregår ved lodtrækning. Der er 20 minutters forberedelse og selve prøven tager ca. 20 minutter. De seneste fem år er gennemførelsесprocenten på 75-80 %.

2. Mini-projekterne og deres relation til modelleringsprocessen

De problemstillinger, de studerende arbejder med i mini-projekterne, vælges for langt de flestes vedkommende fra et katalog – hæfte – der indgår i kursusmaterialet, men de

studerende kan også selv foreslå problemstillinger – evt. med relation til igangværende eller tidligere semesterprojekter. Problemstillingerne er alle formuleret som et overordnet problem, som f.eks.:

1. Hvor gammel er jorden?
2. 10=44 kampagnen – passer det?
3. Hvor mange arter er der plads til på en ø?
4. Hvordan vil skarvbestanden i Dk udvikle sig?
5. Hvordan udvikler verdensbefolkningen sig?
6. Hvor stor bliver en regndråbe og hvor hurtigt falder den?
7. Hvad er systemet i planetssystemet?
8. Hvad styrer en enzymreaktion?
9. Hvordan skal astmamedicin doseres?
10. Hvordan skal strålebehandling af cancersvulster doseres?
11. Hvordan skal pankuron doseres under en operation?
12. Hvad styrer dynamikken mellem rovdyr og bytte-dyr?
13. Hvorfor er det ikke alle, der får influenza?
14. Hvad styrer forekomsten af gonorré?
15. Hvad styrer udbrud af knopormepest?

Udover den overordnede problemstilling er der en introducerende tekst, der forklarer baggrunden og konteksten for problemstillingen, samt i langt de fleste mini-projekter nogle autentiske data vedrørende problemstillingen. Kravene til mini-projekterne er angivet i meget generel form, der inviterer de studerende til at være med til at definere ambitionsniveauet. Derudover er der i hvert mini-projektoplæg formuleret en samling vink og opgaver/spørgsmål, der forventes belyst i rapporten, og som dermed indikerer den indholdsmæssige forventning til omfanget af besvarelsen.

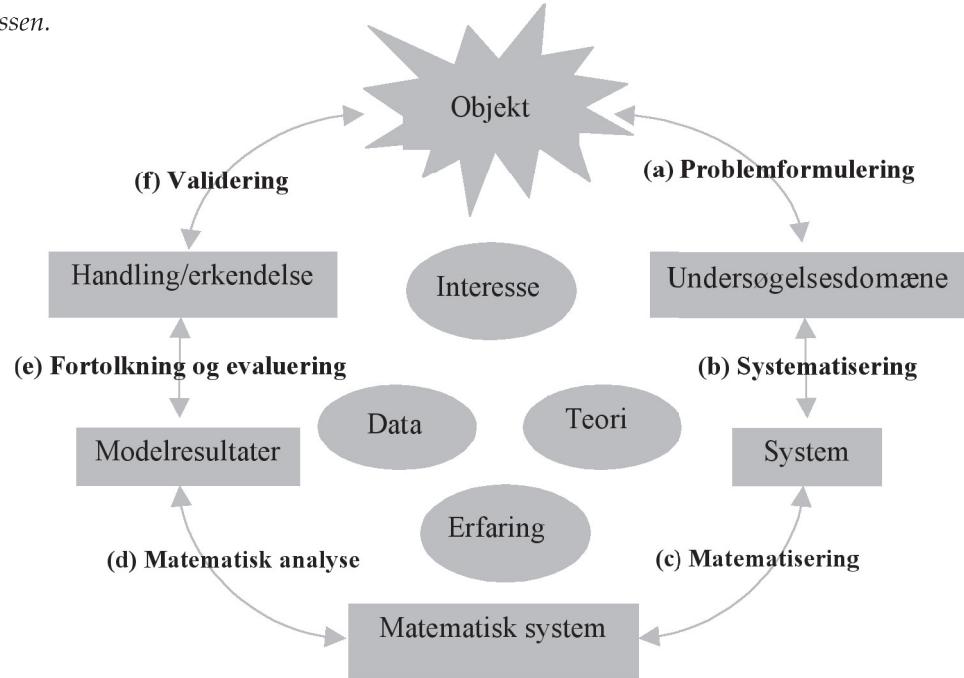
Problemstillerne til mini-projekterne er udarbejdet løende, og jagten på gode problemstillinger fortsætter. Kriterierne er, at de skal være genkendelige eller umiddelbart forståelige for de studerende, at de skal give plads til, at de studerende selv kan modellere, at de skal være autentiske (læs: af interesse uden for matematikundervisning) og helst indeholde autentiske data, at de skal give mulighed for interessante (kontra intuitive) resultater eller nye indsigtter, være åbne over for kritik, og at de til sammen skal bidrage til at tegne et repræsentativt billede af anvendelsen af matematiske modeller i de naturvidenskabelige fag inden for de matematiske emner kursets behandler.

Som et nødvendigt krav skal problemstillerne udfordre de studerende til at arbejde med de indre dele af modelleringsprocessen, det vil sige processerne (c), (d) og (e) i modelleringsprocessen (se figur 1) (Blomhøj, 2006, s. 88).

3. Eksempler på de studerendes arbejde

Vores udviklingsarbejde har været tæt knyttet til den konkrete tilrettelæggelse af undervisningen på kurset, pædagogiske observationer af de studerendes arbejde, vurdering af de studerendes udbytte af kurset, samt de studerendes evaluering af kurset. De pædagogiske ob-

Figur 1: Model af modelleringsprocessen.



servationer har i særlig grad fundet sted under mini-projektførsløbne. Gennem dialoger med grupperne og enkelte studerende har vi på den ene side fået en dybere indsigt i de studerendes forståelse af de matematiske begreber, der indgår i arbejdet med problemstillingerne, og på den anden side bevidst udfordret deres begrebsforståelse gennem indgribende spørgsmål til deres modelleringsarbejde. Denne del af vores arbejde har ført til, at vi bl.a. har designet og analyseret mini-projektproblemstillinger ud fra deres potentiale til at udfordre og øge studerendes forståelse af centrale matematiske begreber.

Igenmenn modelleringsaktiviteter kan de studerende, i læringsprocessen, skabe forbindelser mellem deres erfaringer og viden fra andre områder og den matematik, der indgår i den konkrete modelleringsaktivitet. Det er qua modelleringskonteksten, at de studerendes begrebsforståelse udfordres, idet de studerende i modelleringsprocessen udfordres til at bruge nye sider af deres matematiske begreber og til at bringe dem i spil ud fra nye perspektiver og i nye situationer. Det ser ud til, at studerendes kontekstualiserede refleksioner i sådanne modellerings-situationer spiller en væsentlig faktor for realiseringen af læringspotentialer ved matematisk modellering i forhold til forståelse af matematiske begreber.

I det følgende illustrerer vi, hvordan modelleringsaktiviteter kan støtte de studerendes konstruktion af vigtige matematiske begreber og metoder knyttet til analyse af systemer af to koblede ordinære differentialligninger. Det drejer sig om faseplansanalyse som metode til kvalitativ undersøgelse af differentialligningssystemer og om begreberne banekurven og begyndelsesværdiproblem, nulhældningskurver og ligevegtsløsning. Eksemplet er modellering af dynamikken mellem en rovdyr- og en byttedyrspopulation med udgangspunkt i den klassiske Lotka-Volterra model.

De studerende får opgivet Lotka-Volterra modellen:

$$x' = ax - bxy$$

$$y' = -cy + dxy$$

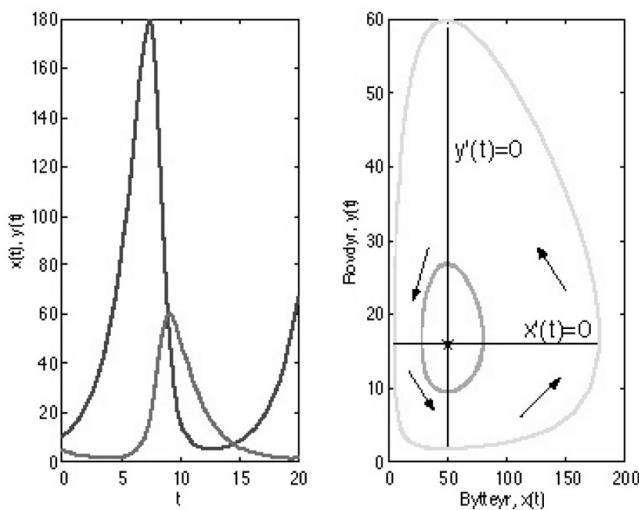
hvor $x(t)$ og $y(t)$ angiver henholdsvis antallet af bytte- og rovdyr til tiden t , og hvor a, b, c og d er positive konstanter.

Første udfordring er at redegøre for forudsætningerne for opstillingen af modellen. Som hjælp hertil opfordres de studerende til at tegne et kompartmentdiagram for systemet af de to koblede differentialligninger. De har flere gange tidligere benyttet kompartmentdiagrammer til opbygning af differentialligningsmodeller – fortrinsvis lineær modeller. I denne sammenhæng er der altså tale om, at de studerende skal gå "baglæns" i processen og læse mening ind i ligningerne ved at identificere "ind- og udflow" fra to "kompartments", der beskriver henholdsvis størrelsen af bytte- og rovdyrpopulationen. Det er i øvrigt ikke helt uproblematisk for alle studerende at erkende, at der ikke er tale om et flow af individer fra byttedyrspopulation ind i rovdyrpopulationen, men om at størrelsen af den ene population påvirker henholdsvis ud- og indflow af den anden population. Samtidig giver dette arbejde anledning til diskussion i grupperne om fortolkningen af de indgående konstanter. For eksempel kræver det lidt overvejelse at erkende, at a og $-c$ skal fortolkes som konstante "netto-fødselsrater", der så kan antages at være positiv for byttedyrene og negativ for rovdylene.

Det er som regel først i denne proces, at de studerende bliver klar over, at der er tale om en ikke-lineær model. De kan se af deres diagram - evt. med hjælp fra et spørgsmål fra læreren - at f.eks. udflow fra byttedyrene afhænger både af x og y . Efter at have fået hul på diskussionen kan de fleste grupper nå frem til en stort set udtømmende liste over biologiske forudsætninger, der skal være opfyldt for at nå frem til Lotka-Volterra modellen. Det kommer ofte frem som en overraskelse for grupperne under dette arbejde, at det forhold, at parametrene er konstante, betyder, at systemets tilstand (i modellen) er entydigt fastlagt ved størrelsen af x og y (en begyndelsestilstand), og at differentialligningerne determinerer systemets udvikling entydigt til evig tid. Sådanne overvejelser begrunder også,

at man kan interessere sig for at afbilde løsninger til differentialligningssystemet i faseplanen.

Efter at have arbejdet med forudsætningerne for modellen bliver de studerende udfordret til at løse ligningerne numerisk for givne parameterværdier og forskellige begyndelsesbetingelser. Det foregår ved hjælp af MatLab, og det er relativt uproblematisk for de fleste grupper at producere løsningskurver som gengivet i figur 2a. Grupperne kan som regel også godt få tegnet banekurven for en løsning i faseplanen, men det betyder imidlertid ikke nødvendigvis, at de enkelte studerende forstår sammenhængen mellem de to grafiske repræsentationer. Banekurven bliver tegnet i MatLab, i et glimt efter at løsningen er beregnet, så de studerende ser således ikke det tidslige forløb i tegneprocessen. Dette forhold bruger vi ofte til at udfordre de studerende til at beskrive det tidslige forløb i faseplanen og forklare sammenhængen mellem de to diagrammer, de selv har frembragt. Også i denne sammenhæng refererer grupperne typisk til den biologiske kontekst. "Vi startede i (10,4). Først vokser byttedyrene fordi, der ikke er ret mange rovdyr. Dvs. vi bevæger os ud ad x -retningen, indtil der kommer så mange byttedyr at rovdyrene begynder at vokse".



Figur 2: (a) Viser udsnit af en løsning $(x(t), y(t))$ til Lotka-Volterra modellen med begyndelsespunktet $(10, 5)$. (b) Faseportræt af samme model med banekurver for løsningen i (a) samt af en løsning svarende til begyndelsespunktet $(40, 10)$. Nulhældningskurverne og det ikke-trivielle ligevægtspunkt er indtegnet.

I det videre arbejde bliver de studerende bedt om at finde ligevægtspunkter for differentialligningssystemet. Det kan de gøre relativt nemt analytisk baseret på tidligere erfaringer, og $x'(t) = 0$ og $y'(t) = 0$ kan ydermere løses separat i dette tilfælde. Ofte overser grupperne imidlertid $(0,0)$ som et ligevægtspunkt. Også her er den biologiske kontekst en støtte til fortolkning af ligningerne: "Hvis der hverken er byttedyr eller rovdyr, hvad sker der så?" og "Hvis der kun er byttedyr eller kun rovdyr?"

Selv når grupperne har fundet ligevægtspunkter er det som oftes ikke helt simpelt for dem at få repræsenteret den information i faseplanen. Nulhældningskurverne bliver indtegnet, men ofte byttes om på $x' = 0$ og $y' = 0$, og her er konteksten igen et godt grundlag, hvorpå de studerende kan kontrollere deres forståelse: "Hvad er det,

der er konstant, når banekurven har henholdsvis lodret og vandret tangent?"

Når grupperne har tegnet banekurver for forskellige begyndelsespunkter som f.eks. vist i figur 2b, er de ofte meget modtagelige for og interesseret i at diskutere, hvad et sådan modelresultat egentlig kan bruges til, når nu modellen bygger på en lang række biologiske forudsætninger, som typisk ikke er opfyldt i virkelige biologiske systemer. Modellen er imidlertid en fin illustration af, at kendskab til et systems grundlæggende dynamik kan give en helt anden indgang til regulering af systemet. Hvis man ønsker at undgå store udsving i systemet, skal man prøve at bringe systemet tættere på en stabil ligevægtstilstand.

Problemstillingen fortsætter med udfordringer til at udbygge modellen med "en befiskningsparameter" og til at modificere differentialligningen for byttedyrene således, at populationen vokser logistisk i stedet for eksponentielt ved fravær af rovdyr. Det giver anledning til en ny faseplansanalyse og et nyt ligevægtspunkt, der er tilstrækende stabilt. Herved bliver der mulighed for at styrke grundlaget for de studerendes forståelse af de centrale matematiske begreber.

4. Erfaringer fra afholdelse af BASE-kurset

Gennem løbende evalueringer og diskussioner kursuslærerne imellem har vi efterhånden fået udviklet rammerne om kurset og vores undervisningspraksis således, at de studerende, der består kurset, faktisk udvikler en modelleringskompetence, der bringer dem i stand til at opstille, analysere og kritisere matematiske modeller inden for kursets område. Dette viser sig allerede for nogle studerende i anden halvdel af kurset i forbindelse med deres 2. semesterprojekter, hvor matematisk modeller og modellering ofte indgår centralt.

Vi har erfaret, at mini-projekterne er det element i kurset, der har størst betydning for de studerendes udbytte. Mini-projekterne udfordrer de studerende til selvstændig modellering, hvilket sjældent kommer på banen i mere traditionel anvendelsesorienterede opgaveregning. Det er i særlig grad gennem modelleringsarbejdet, at de studerendes forståelse af centrale matematiske begreber bliver udfordret og udviklet, og det er også i dette arbejde, at de studerendes forståelsesvanskeligheder kommer op til overfladen. Dette gælder især når de studerende skal anvende matematik i (c) – (e) i modelleringsprocessen (se figur 1). Netop derfor er det en væsentlig pointe i BASE-kurset, at de studerende arbejder med disse processer i forbindelse med mange forskellige problemstillinger.

Det er imidlertid en afgørende forudsætning for, at mini-projekter fungerer didaktisk, at de studerende har forstået og accepteret kursets formål. Formålet er at udvikle matematisk modelleringskompetence, og man kan tænke på arbejdet med mini-projekterne som en første approksimation heraf. Men det er ikke trivielt at bibringe de studerende en sådan forståelse. Særligt i starten af kurset, bruger vi således megen tid på at tale med de studerende om, hvad matematisk modellering er, og hvordan de enkelte bestanddele af kurset kan bidrage til kursets formål. Samtidig fremhæver vi ved hver en given lejlighed, hvordan modelleringskonteksten kan give mening til de matematiske begreber og støtte begrebsforståelsen.



Figur 3: Studerende fra BASE-kurset fordybet i arbejdet med mini-projektet om dynamikken i rovdyr-byttedyrssystemer.

I denne sammenhæng er det en væsentlig pointe ved mini-projekterne, at de giver mulighed for differentiering. Vi ser ofte rapporter over samme problemstilling, som varierer meget i tilgang og matematisk sværhedsgrad. Mini-projekterne giver mulighed for at støtte de studerendes begrebsforståelse på mange forskellige niveauer, og det ser vi som en didaktisk pointe ved denne organisering.

I forhold til kursets formål er det vigtigt, at de studerende oplever mini-projekterne som autentiske i den forstand, at de forekommer relevante og interessante, og at de rækker ud over det matematikfaglige indhold. For mange studerendes motivation og arbejdsindsats spiller mini-projekterne en væsentlig rolle. Det, at der i hvert mini-projekt er en overordnet problemstilling, der sætter rammerne for arbejdet, og ikke blot en samling af mere eller mindre løsrevne opgaver, er afgørende i denne sammenhæng. De studerende opererer ikke i forhold til en "er det godt nok"-taktik, men tager selv vurderingsansvaret på sig. Det er "godt nok", når de føler, at de vil være i stand til at give en udtømmende besvarelse af problemstillingen til den individuelle mundtlige evaluering. Det er naturligvis et subjektivt kriterium, hvor vi som kursuslærere definerer den nedre grænse gennem vores godkendelse/ikke-godkendelse af rapporterne. I den sammenhæng fungerer de individuelle mundtlige prøver efter hvert semester godt, idet de giver en hensigtsmæssig tilbagevirkning på engagementet i arbejdet med mini-projekterne.

Læringspotentialerne ved matematisk modellering i forhold til forståelse af matematiske begreber realiseres ikke automatisk. Fra vores egen undervisningspraksis har vi erfaret, at det er essentielt at udfordre de studerende til at reflektere over deres matematiske forståelse af de relevante begreber i selve modelleringsaktiviteten. Disse refleksioner kan ikke styres eller "tvinges" ind i de studerendes arbejde gennem eksplizit formulerede krav i oplæggene til mini-projekterne. Her er der en hårfin balance mellem udfordringerne i projektoplæggene og interaktionen med de studerende under gruppearbejdet. Forsøges disse refleksioner skrevet ind i oplægget, sker der ofte det, at problemstillingen tømmes for udfordrin-

ger og bliver til "gøre"-opgaver. På den anden side er det vigtigt, at man som lærer er opmærksom på, hvornår i gruppens arbejde sådanne læringspotentialer er til stede og at være forberedt på at udfordre og støtte de studerendes refleksioner i situationen. En stor del af udviklingen af vores egen undervisningspraksis på kurset har været, og er fortsat, at identificere sådanne læringspotentialer i de forskellige mini-projekter og forberede os på, hvordan vi vil udfordre de studerende i situationen.

Noter:

- 1) BASE er udviklet og kursusnoterne skrevet af Morten Blomhøj, Tinne Hoff Kjeldsen og Johnny Ottesen (2005), IMFUFA. For uddybning af historien bag udviklingsarbejdet se (Blomhøj, Jensen, Kjeldsen og Ottesen, 2001).
- 2) Denne problemstilling er analyseret i detalje i forhold til integralbegrebet i (Blomhøj og Kjeldsen, 2007) og i forhold til systemer af to koblede lineære differentialligninger i (Blomhøj og Højgaard Jensen, 2003).

Refencer:

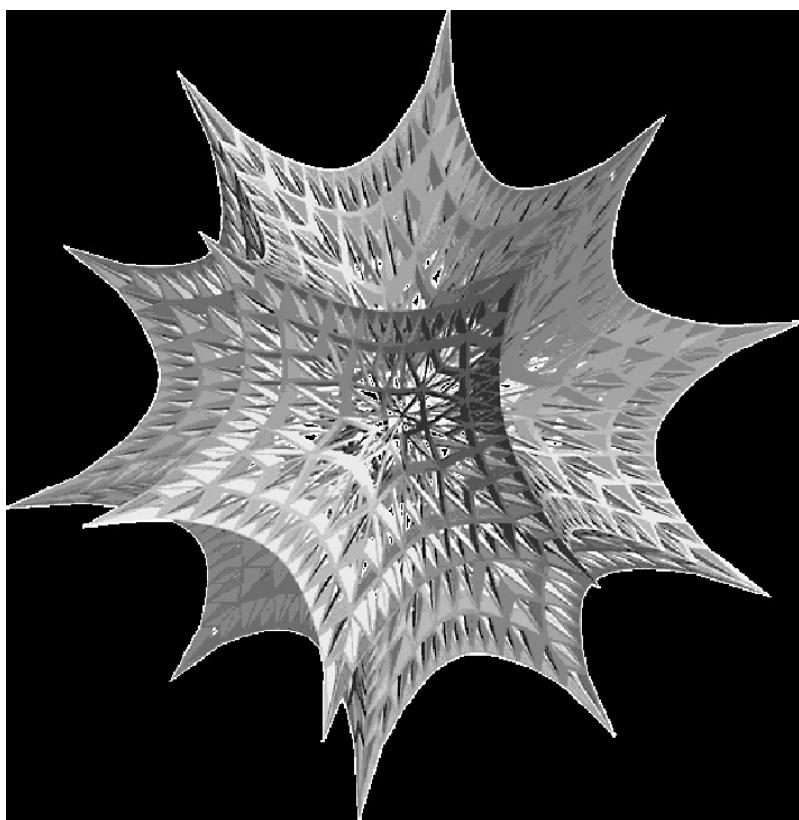
- Blomhøj, M. (2006). Mod en didaktisk teori for matematisk modellering. I Skovsmose, O. og Blomhøj, M. (red.) (2006): *Kunne det tænkes? – om matematiklæring* (s.80-109). København: Malling Beck.
- Blomhøj, M., Kjeldsen, T.H. og Ottesen. J. (2005). *BASE – et grundkursus i matematisk modellering*, 5. udgave. Noter i tre bind til kursus ved den Naturvidenskabelig Basisuddannelse. Roskilde Universitet: IMFUFA. Kan bastilles via vedel@ruc.dk.
- Blomhøj, M. og Højgaard Jensen, T. (2003). Developing mathematical modelling competence: Conceptual clarification and educational planning, *Teaching Mathematics and its applications* 22 (3), pp. 123-139.
- Blomhøj, M. og Højgaard Jensen, T. (2007). What's all the fuss about competences? Experiences with using a competence perspective on mathematics education to develop the teaching of mathematical modelling. In: W. Blum (red.): *Modelling and applications in mathematics education*, pp. 45-56. The 14th ICMI-study 14. New York: Springer-Verlag.
- Blomhøj, M., Kjeldsen, T.H., Højgaard Jensen, T. og Ottesen, J. (2001). *Matematisk modellering ved den naturvidenskabelige basisuddannelse – udvikling af et kursus*. Roskilde Universitetscenter, IMFUFA-tekst 402.
- Blomhøj, M. og Kjeldsen, T.H. (2007). Learning the integral concept through mathematical modelling. I Pitta-Pantazi, D., og Philippou, G. (eds.): *European Research in Mathematics: Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (s. 2070-2079). Cypræs: ERME, Department of Education, University of Cyprus.



Kim Knudsen er pr. d. 1/8 2008 ansat som lektor ved DTU Matematik. Kim er uddannet cand. scient. i Matematik og Datalogi fra Aalborg Universitet (1999) og ph.d. (2002) i Matematik også fra Aalborg Universitet. Hans forskning omhandler teoretiske og numeriske metoder til løsning af inverse problemer for partielle differentialligninger. Sådanne problemer opstår typisk, når man ønsker at undersøge et fænomen, som kun kan observeres via indirekte målinger, eksempelvis i forbindelse med medicinske scanninger og geofysik.

På DTU Matematik tilknyttes han gruppen Anvendt funktionalanalyse.

Kim er gift med Janne, som er lærer. Sammen har de twilingerne Johanne og Karl Emil på fire år. Fritiden bruger Kim sammen med familien og på motionsløb.



Aftermath

ved Mogens Esrom Larsen



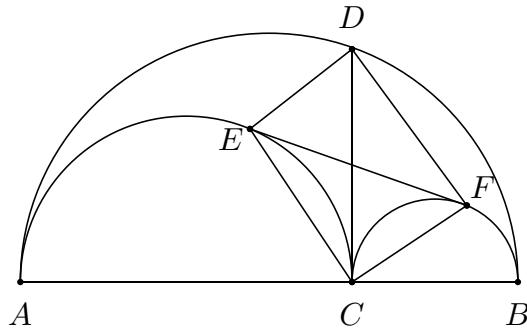
LØSNINGER

Opgaverne i sidste nr. var ren nostalgi; et sæt studenteksamensopgaver fra tiden før Kristensen og Rindung, in casu sættet fra 1961. Problemklubben "Con Amore" har været så venlig at give en smuk og udførlig løsning, som er ret omfattende. Jeg vælger derfor at bringe et eksempel her, mens hele historien findes på hjemmesiden:

<http://www.matilde.mathematics.dk/arkiv/M34/conamore.pdf>

NYE OPGAVER

Firkantet



Lad C være et vilkårligt punkt på liniestykke AB mellem A og B , og tegn halvcirkler til samme side over diametrerne AB , AC og CB . Lad D være det punkt på halvcirklen AB , der har CD vinkelret på AB , og lad EF være fællestangenten til de to små halvcirkler.

Vis, at $ECFD$ er et rektangel.

Trekantet

En trekant er tegnet på ternet papir, så alle tre hjørner er i skæringspunkter (punkter med heltallige koordinater). Lad nu r være antallet af skæringspunkter på randen og i antallet af skæringspunkter i det indre af trekanten. Vis, at arealet af trekanten er

$$i + \frac{1}{2}r - 1$$

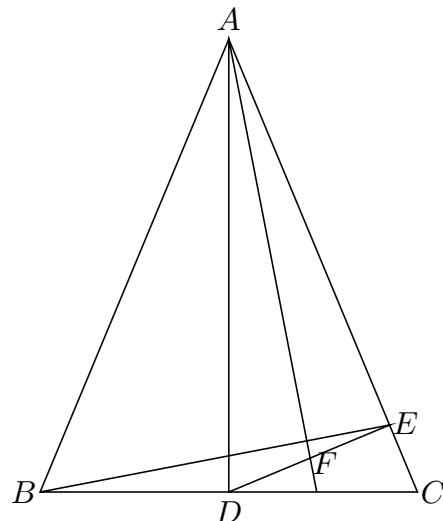
Eksponentielt

Når man får at vide, at tallet 2^{29} er 9-cifret, og at de 9 cifre alle er forskellige, kan man så uden at udregne tallet bestemme, hvilket ciffer der mangler?

Kvadratisk

n er et helt tal, så $2n + 1$ er et kvadrattal. Vis, $n + 1$ er sum af to sukcessive kvadrattal.

Trekantet



Trekanten $\triangle ABC$ er ligebenet med $AB = AC$, D er midtpunktet på BC , E på AC er det punkt, hvor ED er vinkelret på AC og F er midtpunktet af DE .

Vis, at AF står vinkelret på BE .

Heltalligt

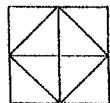
Givet et naturligt tal, n , bestem antallet af firsæt, (a, b, c, d) , så at $0 \leq a \leq b \leq c \leq d \leq n$.

Polynomiet

Lad a , b og c være ulige, hele tal. Vis, at ligningen

$$ax + bx + c = 0$$

ikke kan have en rational rod.



CON AMORE PROBLEMGRUPPE

9. januar 2008.

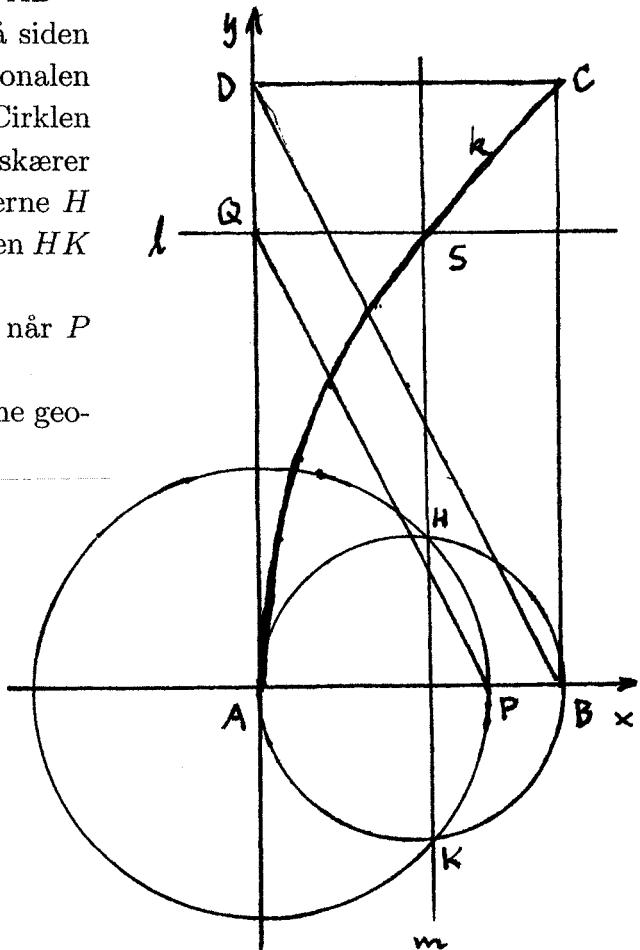
1. $ABCD$ er et rektangel, hvori siden $AB = a$ og siden $BC = 2a$. P er et punkt på siden AB . Linien gennem P parallel med diagonalen BD skærer siden AD i punktet Q . Cirklen med A som centrum og AP som radius skærer cirklen over AB som diameter i punkterne H og K . S er skaeringspunktet mellem linien HK og linien gennem Q parallel med AB .

Bestem det geometriske sted for S , når P gennemløber siden AB .

Angiv art og beliggenhed af det fundne geometriske sted.

Vi indfører et retvinklet koordinatsystem med begyndelsespunkt i A , B på den positive del af x -aksen og D på den positive del af y -aksen. Da er $A = (0,0)$, $B = (a,0)$, $C = (a, 2a)$ og $D = (0, 2a)$. Lad P være $(t, 0)$ (hvor $0 \leq t \leq a$); da PQ er parallel med AB , er $Q = (0, 2t)$, og linien l gennem Q parallel med AB har ligningen $y = 2t$.

Den cirkel, der har centrum i A og går gennem P , har ligningen $x^2 + y^2 = t^2$, og cirklen med AB som diameter har ligningen $(x - \frac{a}{2})^2 + y^2 = (\frac{a}{2})^2$ eller $x^2 + y^2 = ax$. Linien m



gennem cirkelernes skæringspunkter H og K har da ligningen $t^2 = ax$ eller $x = \frac{t^2}{a}$.

Da S er skæringspunktet mellem l og m, har den kurve k, som er det geometriske sted for S, parameterfremstillingen

$$x = \frac{t^2}{a}, \quad y = 2t, \quad 0 \leq t \leq a,$$

og vi ser, at både x og y vokser med t, og at det gælder $0 \leq x \leq a$ og $0 \leq y \leq 2a$. Ved elimination af t finder vi, at k har ligningen $ax = (\frac{y}{2})^2$ eller

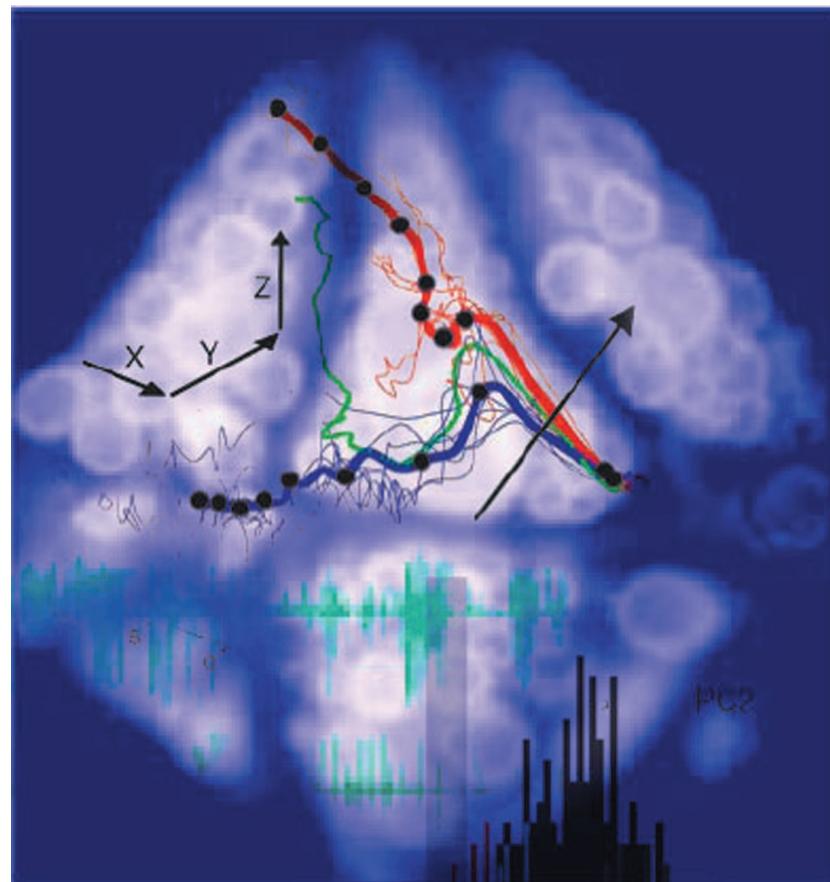
$$y^2 = 4ax, \quad \text{hvor } 0 \leq x \leq a, \quad 0 \leq y \leq 2a.$$

Herav ses vi, at k er den del af parablen ved x-aksen som symmetriakse og y-aksen som toppunktstangent, med toppunktet i $(0,0)$ og brandpunktet i $(a,0)$, som forbinder punkterne $(0,0)$ og $(a,2a)$, eller udtrykt direkte ved det givne: Det geometriske sted for S er den del af parablen ved linien gennem A og B som symmetriakse og linien gennem A og D som toppunktstangent, med toppunktet i A og brandpunktet i B, som forbinder A og C.

*

Ved uanbringelighed returneres bladet til afsender:

Matilde
Institut for Matematiske Fag
Aarhus Universitet
Ny Munkegade Bygning 1530
8000 Århus C



Brigman Montage