FAMØS

Fagblad for Aktuar, Matematik, -Økonomi og Statistik
19. årgang, nr. 3, april 2006

Glædeligt forår!

Indhold

Den almindelige, den gode og den smarte 3
– Formidlingsaktivitet af Pia Mikkelsen

Trekantens medianer 9
– Side 9-sætning

Præmieopgaver 12
– Vind en Mystisk Flaske

Opgavebesvarelser 13
– Hvem vandt den mystiske flaske?

Konstanten Kenny & Eksponential-Erik 14
– Klassisk tegneserie af Damskur

Illustration 15
– Fordi vi elsker Maple

Fire farver er nok 16
– Formidlingsaktivitet af Karl Malassé

Efterårskryds 25
– Krydsogtværs af Damskur

Stabilitet i fysiske modeller 28
– Hvad forsker jeg i? $n$te del i serien

Glædeligt forår! 40
– Forårskåd forårsleder

Den almindelige, den gode og den smarte

Tre beviser for at ethvert primtal $p \equiv 1 \pmod 4$ *er en sum af to kvadrater*

Pia Mikkelsen

De fleste af os har i løbet af vores karriere som matematikstuderende stiftet bekendtskab med ovenstående kendte sætning. F.eks. i forbindelse med nærmere granskning af Anders Thorups 2 AL noter [1].
Sætningen går tilbage til Fermat. -Fermat opskriver sætningen, men giver ikke noget bevis for sin påstand.
Der er dog senere blevet givet en lang række beviser for denne sætning (bl.a. af Euler, Lagrange og Dedekind), og vi skal her se nærmere på tre forskellige indgangsvinkler.

Den almindelige

De fleste af os har nok før set et bevis baseret på teorien om kvadratiske talringe (se f.eks. RNG.6 [1]). -Mere præcist betragtes *Gauss' talring*

$$\mathbb{Z}[i] = \{x + yi \mid x,\ y \in \mathbb{Z}\}.$$

Elementerne i $\mathbb{Z}[i]$ kaldes *gaussiske heltal.*
Ved $\overline{\cdot}$ menes kompleks konjugering og vi minder kort om, at *normen* $N : \mathbb{Z}[i] \to \mathbb{Z}$,

$$x + yi \longmapsto (x + yi)\,\overline{(x + yi)} = x^2 + y^2,$$

er multiplikativ. Det er desuden nemt at overbevise sig om, at der for et element $\alpha \in \mathbb{Z}[i]$ gælder:

$$N(\alpha) = 1 \quad \Leftrightarrow \quad \alpha \text{ er en enhed i } \mathbb{Z}[i].$$

(Se evt. s.211 [1]).
Hvis p er reducibel i $\mathbb{Z}[i]$ -d.v.s. hvis der findes en opløsning

$$p = \alpha\beta, \quad \alpha,\ \beta \in \mathbb{Z}[i],$$

hvor $\alpha,\ \beta$ ikke er enheder i $\mathbb{Z}[i]$, må

$$p^2 = N(p) = N(\alpha)N(\beta).$$

Da α og β ikke er enheder, er $N(\alpha)$, $N(\beta)$ positive heltal forskellig fra 1, så $N(\alpha) = p$. I.e. hvis $\alpha = x + yi$, er

$$p = N(\alpha) = x^2 + y^2.$$

Hvis vi kan vise, at p er et reducibelt element i $\mathbb{Z}[i]$, har vi altså vores ønskede sætning. Da $\mathbb{Z}[i]$ er et PID (se s.217 [1]), er de irreducible elementer netop primelementerne, så det er nok at vise, at p ikke er et primelement i $\mathbb{Z}[i]$, i.e. vi søger α, $\beta \in \mathbb{Z}[i]$ for hvilket der gælder, at p går op i $\alpha\beta$, men p går hverken op i α eller β. Til dette benyttes følgende sætning også kendt som "Første supplement til den kvadratiske reciprocitetssætning":

Sætning 1. *Lad p være et primtal og antag $p \equiv 1 \pmod 4$. Da findes et heltal n, så*

$$p \,|n^2 + 1.$$

Bevis. Iflg. Fermat's lille Sætning (se GRP. 4, s. 82 [1]) gælder der for ethvert heltal z primisk med p, at

$$z^{p-1} - 1 \equiv 0 \pmod p.$$

Da $p \equiv 1 \pmod 4$, skrives $p-1 = 4k$. For ethvert heltal z primisk med p , gælder således, at

$$p|z^{4k} - 1 = \left(z^{2k} + 1\right)\left(z^{2k} - 1\right).$$

Da p er et primtal, må p gå op i $z^{2k} + 1$ eller $z^{2k} - 1$. I.e.

$$z^{2k} + 1 \equiv 0 \pmod p \quad \text{eller} \quad z^{2k} - 1 \equiv 0 \pmod p.$$

Tallene $z = 1, \ldots, p - 1 = 4k$ er alle primiske med p og altså løsninger til en af ovenstående kongruenser. Da $\mathbb{Z}_p$ er et legeme og altså specielt et integritetsområde, har polynomiet

$$z^{2k} - 1 \in \mathbb{Z}_p[z]$$

højst $2k$ rødder (se POL 3, s.236 [1]). Derfor findes et heltal z blandt tallene $1, \ldots, 4k$, så $z^{2k} + 1 \equiv 0 \pmod p$. Sættes $n = z^k$ fås det ønskede. □

Der findes altså et n, så

$$p \,|n^2 + 1 = (n + i)(n - i).$$

Da $\frac{n}{p} \pm \frac{1}{p}i \notin \mathbb{Z}[i]$, går p hverken op i $n + i$ eller $n - i$ og p er således ikke et primelement i $\mathbb{Z}[i]$.

Den gode[2]

Denne udgave af vores bevis bygger på følgende sætning:

Thues sætning. Lad $n > 1$ være et naturligt tal og lad k være det mindste heltal, således $k > \sqrt{n}$. -D.v.s. $k - 1 \leq \sqrt{n} < k$. For ethvert heltal a primisk med n findes $x,\ y \in \mathbb{N}$, så $x,\ y \leq k - 1$ og

$$ay \equiv \pm x \pmod{n}.$$

Bevis. Betragt tal på formen

$$ay + x, \quad x,\ y \in \{0, 1, \ldots, k-1\}.$$

Da der findes $k^2 > n$ sådanne tal og kun n restklasser modulo n, må mindst to af tallene ligge i samme restklasse (Skuffeprincippet). Der findes altså $x_1, y_1, x_2, y_2 \in \{0, 1, \ldots, k-1\}$, hvor

$$x_1 - x_2 \neq 0 \quad \text{eller} \quad y_1 - y_2 \neq 0$$

og

$$a\,(y_1 - y_2) \equiv x_2 - x_1 \pmod{n}\,.$$

Desuden må der gælde, at

$$0 < |x_2 - x_1|,\ |y_1 - y_2| \leq k - 1.$$

Antag nemlig $x_2 - x_1 = 0$. Så går n op i $a(y_1 - y_2)$. Da a og n er primiske vil n gå op i $y_1 - y_2$, hvorfor $y_1 - y_2 = 0$. Ligeledes giver $y_1 - y_2 = 0$, at $x_2 - x_1 = 0$. Vi kan antage, at $y_1 > y_2$. Sættes

$$y = y_1 - y_2 \quad \text{og} \quad x = \begin{cases} x_2 - x_1 & \text{hvis } x_2 > x_1 \\ x_1 - x_2 & \text{hvis } x_1 > x_2 \end{cases}$$

fås den ønskede kongruens. □

Vi så i foregående bevis, at hvis $p \equiv 1 \pmod 4$, har kongruensen

$$z^2 + 1 \equiv 0 \pmod{p} \tag{1}$$

en løsning. Da en sådan løsning z er primisk med p findes, iflg. Thues Sætning, $x,\ y \in \mathbb{N}$, så

$$yz \equiv \pm x \pmod{p}$$

og $x,\ y < \sqrt{p}$. Af kongruensen (1) haves så, at

$$y^2z^2 + y^2 \equiv 0 \pmod{p},$$

hvorfor

$$x^2 + y^2 \equiv 0 \pmod{p}.$$

Men da $x^2 + y^2 < 2p$, må $x^2 + y^2 = p$.

Den smarte[3]

En afbildning $\sigma : S \to S$ kaldes en involution (eller involutorisk), hvis den er selvinvers, i.e. hvis $\sigma^2 = \mathrm{Id}_S$. For en involution σ givet på en endelig mængde S haves, at

Lemma 2.

$$|S| \text{ er ulige } \Leftrightarrow \sigma \text{ har et ulige antal fixpunkter .}$$

Specielt må en involution, givet på en mængde med et ulige antal elementer, have mindst ét fixpunkt.

Bevis. Ved

$$(x,y,z) \sim (x',y',z') \overset{\text{def}}{\Leftrightarrow} (x,y,z) = (x',y',z') \vee \sigma(x,y,z) = (x',y',z')$$

defineres en ækvivalensrelation på S. Ækvivalensklasserne indeholder præcist et element, hvis dette er et fixpunkt, og ellers 2. Da ækvivalensklasserne udgør en klassedeling på S, kan elementantallet $|S|$ findes som summen af elementantallet i hver ækvivalensklasse. Involutionen σ har altså et ulige antal fixpunkter netop hvis $|S|$ er ulige. □

Betragt nu den endelige mængde

$$S := \{(x,y,z) \in \mathbb{N}^3 \mid x^2 + 4yz = p\}.$$

Det er klart, at $S \neq \emptyset$, da $(1,1,k) \in S$, $p = 4k+1$. Betragt desuden afbildningen $\sigma : S \to S$:

$$\sigma : (x,y,z) \mapsto \begin{cases} (x+2z, z, y-x-z) & \text{hvis } x < y-z \\ (2y-x, y, x-y+z) & \text{hvis } y-z < x < 2y \\ (x-2y, x-y+z, y) & \text{hvis } x > 2y \end{cases} .$$

Et punkt $(x,y,z) \in \mathbb{N}^3$, hvor $x = y-z$ eller $x = 2y$, ligger ikke i S, da p er et primtal. Derfor er σ defineret på hele S. Ved udregning ses desuden let, at σ faktisk afbildeder ind i S. Vores afbildning σ er således veldefineret.
Faktisk er σ involutorisk og har netop ét fixpunkt. Betragt nemlig de tre tilfælde

- Hvis $x < y - z$:

  Da $x + 2z > 2z$, er

$$\begin{aligned} \sigma^2(x,y,z) &= \sigma(x+2z, z, y-x-z) \\ &= ((x+2z)-2z, (x+2z)-z+(y-x-z), z) \\ &= (x,y,z). \end{aligned}$$

- Hvis $y - z < x < 2y$:

  Da $y - (x - y + z) < 2y - x < 2y$ er

$$\begin{aligned} \sigma^2(x,y,z) &= \sigma(2y - x, y, x - y + z) \\ &= (2y - (2y - x), y, (2y - x) - y + (x - y + z)) \\ &= (x, y, z). \end{aligned}$$

- Hvis $x > 2y$:

  Da $x - 2y < (x - y + z) - y$ er

$$\begin{aligned} \sigma^2(x,y,z) &= \sigma(x - 2y, x - y + z, y) \\ &= ((x - 2y) + 2y, y, (x - y + z) - (x - 2y) - y) \\ &= (x, y, z). \end{aligned}$$

I.e. $\sigma^2 = \mathrm{Id}_S$.
Det er umiddelbart, at der for et eventuelt fixpunkt $(x, y, z) \in S$, må gælde, at $y - z < x < 2y$ og i så fald er $x = y$. De mulige fixpunkter i S har altså formen (x, x, z) og opfylder

$$x^2 + 4xz = x(x + 4z) = p.$$

Da p er et primtal, må $x = 1$ og $z = k$, hvor $p = 4k + 1$. Afbildningen σ har altså netop ét fixpunkt - nemlig $(1, 1, k)$.
Iflg. vores Lemma er $|S|$ ulige og involutionen

$$(x, y, z) \mapsto (x, z, y)$$

har ligeledes et fixpunkt i S. For dette må gælde, at $y = z$ og altså er $x^2 + (2y)^2 = p$.

Det følger umiddelbart af vores sætning, at ethvert heltal, der er et produkt af primtal $\equiv 1 \pmod 4$, kan skrives som en sum af to kvadrater, idet

$$(x^2 + y^2)(a^2 + b^2) = (xa + yb)^2 + (xb - ya)^2.$$

Ingen af vores beviser ovenfor er konstruktive, men der findes faktisk konstruktive beviser. Et sådant bevis findes bl.a. i [4], som bygger på teorien om kvadratiske rester.

Man kan desuden spørge sig selv om opskrivning af primtal $\equiv 1 \pmod 4$ som en sum af to kvadrater er entydig? Svaret er ja til og med ombytning og fortegn. Dette kan vises på flere måder. F.eks. følger det af entydig primfaktorisering i $\mathbb{Z}[i]$. Ellers antag, at $p = x^2 + y^2 = a^2 + b^2$. Da der som før nævnt findes heltal z, så $z^2 \equiv -1 \pmod p$, sluttes

$$\begin{aligned} x^2 &\equiv -y^2 \equiv z^2 y^2 \pmod p \\ a^2 &\equiv -b^2 \equiv z^2 b^2 \pmod p. \end{aligned}$$

Da er

$$\begin{aligned} x &\equiv \pm zy \pmod{p} \\ a &\equiv \pm z \pmod{p} \end{aligned}$$

og ved eventuelt at erstatte y med $-y$ eller b med $-b$, kan vi antage, at

$$\begin{aligned} x &\equiv zy \pmod{p} \\ a &\equiv zb \pmod{p}. \end{aligned}$$

Altså må

$$xa \equiv z^2yb \equiv -yb \pmod{p}.$$

Da

$$p^2 = (x^2+y^2)(a^2+b^2) = (xa+yb)^2 + (xb-ya)^2$$

og p går op i $xa+yb$, er p^2 her skrevet som en sum af to hele ikke-negative tal begge delelig med p^2. Derfor må et af disse tal være nul.
Hvis $xa = -yb$, går x op i b, da x og y er primiske. Tilsvarende sluttes, at b går op i x. I.e. $x = \pm b$, hvorfor $x^2 = b^2$ og $y^2 = a^2$. Hvis $xb - ya = 0$ fås tilsvarende, at $x^2 = a^2$ og $y^2 = b^2$ [4].

[1] A. Thorup, Matematik 2AL, Algebra, 2.udg., 2.oplag, Matematisk Afdeling, Københavns Universitet, 1998.

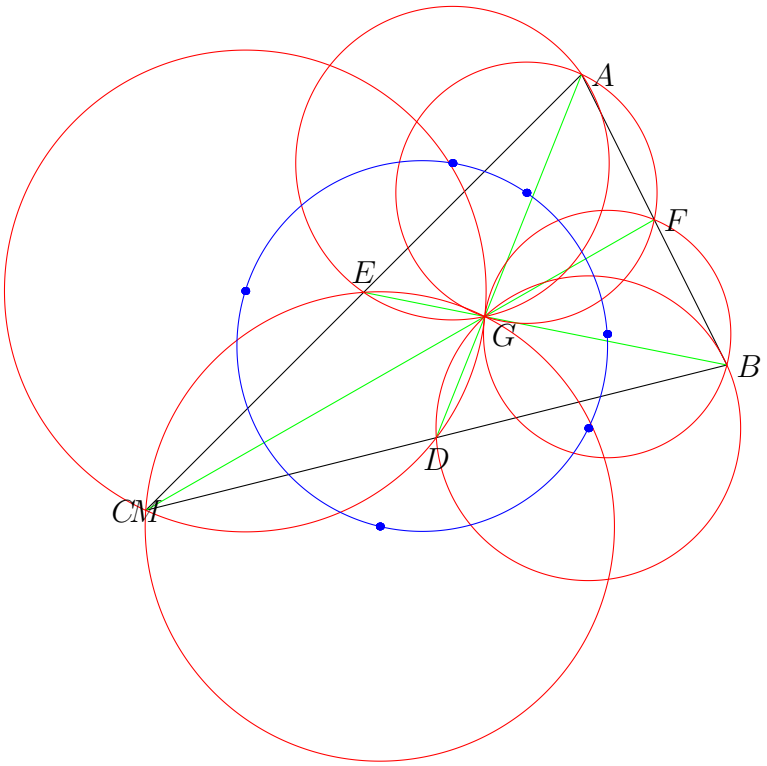
[2] T. Nagell, Introduction To Number Theory, Kapitel VI, s. 122-123, 188-189, John Wiley & Sons, Inc., New York, 1951.

[3] D. Zagier, A One-Sentence Proof That Every Prime $p \equiv 1 \pmod 4$ Is a Sum of Two Squares, American Mathematical Monthly, s.144, Bind 97, 1990. [4] C.U.Jensen, Matematik 3AL, Klassisk Algebra, HCØ Tryk, 2000.

Trekantens medianer

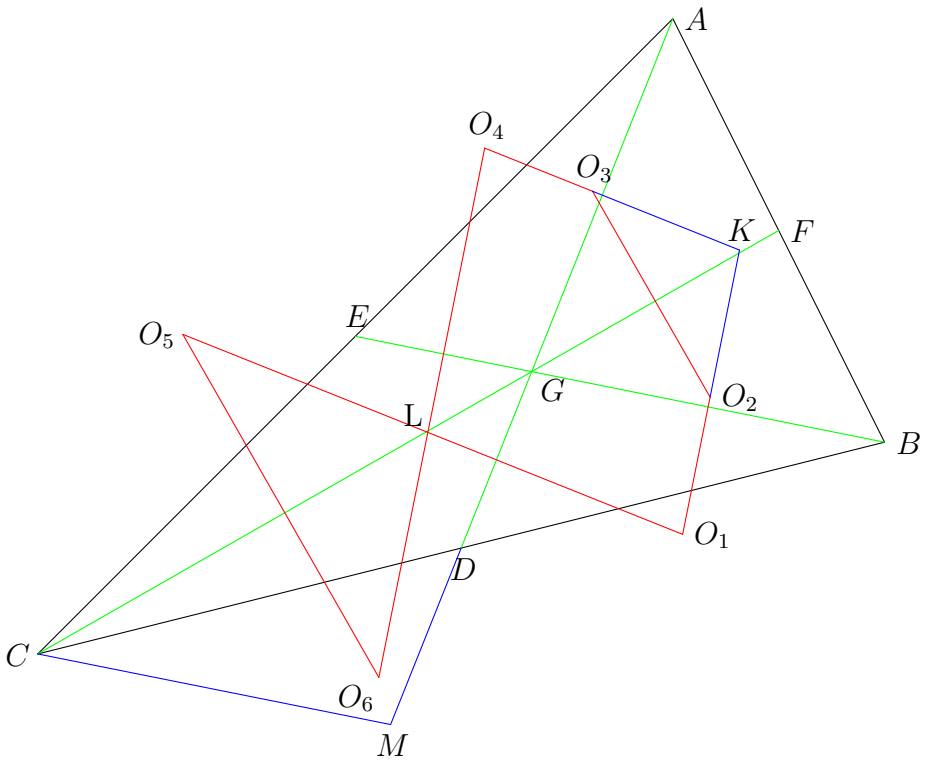
Nicolaj Strands

Side-9 sætningen er i dette nummer hentet fra plangeometrien. Der er tale om et overraskende resultat vedrørende medianerne i en trekant.



Side 9-sætningen. *En Trekants medianer skærer hinanden i et punkt og deler herved trekanten op i 6 mindre trekanter. Centrerne for disse trekanters omskrevne cirkler ligger på en cirkel.*

Bevis. Vi betegner midtpunkterne af trekantens sider D, E og F og medianernes skæringspunkt med G. Lad desuden O_1, O_2, O_3, O_4, O_5 og O_6 være centrer for de omskrevne cirkler for $\triangle DBG$, $\triangle BFG$, $\triangle FAG$, $\triangle AEG$, $\triangle ECG$ og $\triangle CDG$ (se figuren).



Linjen O_3O_4 er midtnormal til AG da $|O_3A| = |O_3G|$ og $|O_4A| = |O_4G|$. På samme måde er O_6O_1 midtnormal til GD. De er altså begge vinkelret på AD og deres indbyrdes afstand er $\frac{1}{2}|AD|$. Tilsvarende er O_1O_2 og O_4O_5 vinkelrette på BE med indbyrdes afstand $\frac{1}{2}|BE|$.

Lad K være skæringspunktet mellem linjerne O_1O_2 og O_3O_4 og L skæringspunktet mellem O_4O_5 og O_1O_6. Betragt parallelogrammet KO_4LO_1. En højde i parallelogrammet er afstanden mellem de parallelle linjer O_3O_4 og O_1O_6, dvs. den er $\frac{1}{2}|AD|$, så parallelogrammets areal er

$$\mathrm{Ar}(\square KO_4LO_2) = \frac{1}{2}|AD||KO_4|.$$

Den anden højde i parallelogrammet er afstanden mellem de parallelle linjer KO_1 og LO_4 som er $\frac{1}{2}|BE|$. Parallelogrammets areal er altså også

$$\mathrm{Ar}(\square KO_4LO_2) = \frac{1}{2}|BE||KO_1|.$$

Altså er

$$|AD||KO_4| = |BE||KO_1|$$

eller

$$\frac{|KO_1|}{|KO_4|} = \frac{|AD|}{|BE|}. \tag{1}$$

Vi ser nu på $\triangle KO_2O_3$. Vi har

$$KO_2 \perp BG, \quad O_2O_3 \perp FG, \quad KO_3 \perp AG,$$

hvoraf

$$\angle KO_2O_3 = \angle BGF \quad \text{og} \quad \angle KO_3O_2 = \angle FGA. \tag{2}$$

Så får vi i $\triangle KO_2O_3$ at

$$\begin{aligned} O_2KO_3 &= 180^\circ - \angle KO_2O_3 - KO_3O_2 \\ &= 180^\circ - \angle BGF - \angle FGA \\ &= 180^\circ - \angle BGA = \angle DGB. \end{aligned}$$

En linje gennem C parallel med BG skærer AG i M. Så er

$$\angle MCG = \angle BGF = \angle KO_2O_3 \quad \text{og} \angle MGC = \angle FGA = \angle KO_3O_2,$$

hvor vi har brugt (2). Dermed er $\triangle KO_2O_3$ og $\triangle MCG$ ensvinklede så

$$\frac{|KO_3|}{|KO_2|} = \frac{|MG|}{|MC|}. \tag{3}$$

Nu er $\triangle MCD$ og $\triangle GBD$ kongruente fordi de er ensvinklede og $|CD| = |BD|$. Så er

$$|MG| = 2|GD| = \frac{2}{3}|AD| \quad \text{og} \quad |MC| = |GB| = \frac{2}{3}|BE|.$$

Dette giver sammen med (3) og (1) at

$$\frac{|KO_3|}{|KO_2|} = \frac{|MG|}{|MC|} = \frac{|AD|}{|BE|} = \frac{|KO_1|}{|KO_4|},$$

eller

$$|KO_1||KO_4| = |KO_3||KO_4|.$$

Dette medfører ifølge *korde-sekantsætningen* at punkterne O_1, O_2, O_3 og O_4 ligger på en cirkel. På samme måde ligger O_2, O_3, O_4 og O_5 på en cirkel og O_3, O_4, O_5 og O_6 ligger på en cirkel. Da cirklerne to og to har tre punkter til fælles må de være sammenfaldende. □

Kilde: Matematiske miniaturer, Jens Carstensen og Alija Muminagić.

Præmieopgaver

Mikkel Abrahamsen

Denne uge stiller vi to opgaver, hvor kun den første er bare en anelse puzzle-agtig.

Husk, at der er mulighed for at vinde en my(s)tisk flaske!

Den onde

Sune ('05) havde engang en mængde. Han var rigtig glad for sin mængde, og brugte meget af sin tid på at betragte og beundre den. Mængden var skrevet op kun vha. krøllede parenteser og kommaer. Desværre skete der det uheldige, at Sune (ved en fejl selvfølgelig) kom til at opbevare sin mængde et lidt usikkert sted. Uheldigvis tog en fæl kanalje chancen og lemlæstede Sunes mængde ganske udtalt. Den fæle kanalje tog nemlig en gevaldig mukkert og en ambolt frem, og hamrede alle de krøllede parenteser helt rette. Da Sune fandt sin mængde igen var der til hans altoverskyggende ærgrelse kun følgende tilbage:

$$|||, ||||, |||, |||||, |||, ||||, |||, |||||||$$

Opgaven lyder nu: hjælp Sune med at rekonstruere mængden, og gør rede for entydigheden af rekonstruktionen!

Den virkeligt sadistiske

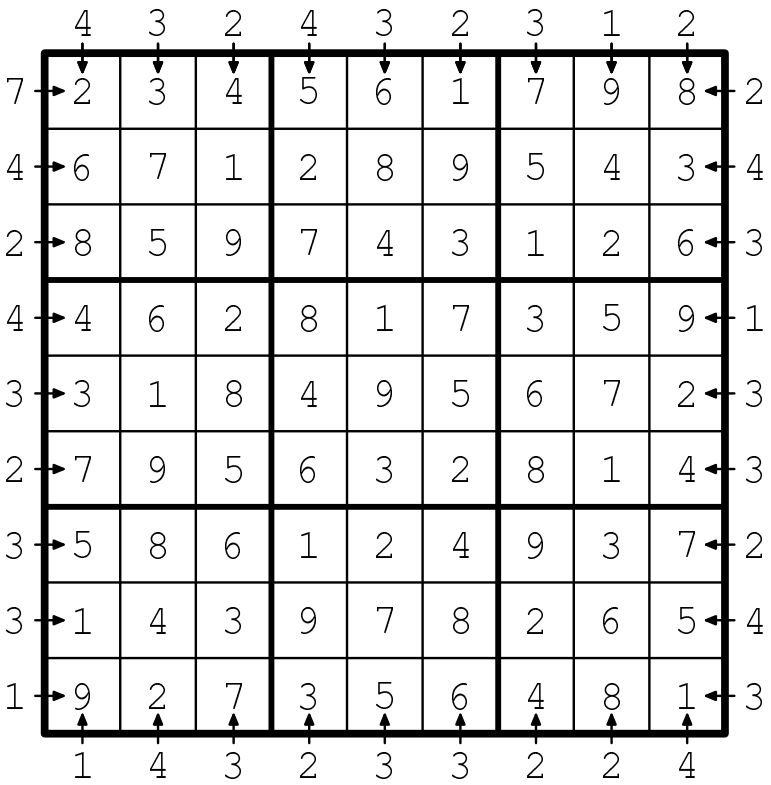
Lad for ethvert naturligt tal $n \in \mathbb{N}$ $\sigma(n)$ være summen af divisorerne i n. Altså er fx $\sigma(3) = 4$ og $\sigma(12) = 28$. n siges at være *overvægtigt* hvis $\sigma(n) > 2n$. Et naturligt tal m siges at være *forårskådt* hvis det ikke kan skrives som en sum af to overvægtige tal.

Afgør om der findes forårskåde tal. Afgør i bekræftende fald om der er en øvre grænse for hvor store de kan være, og find i bekræftende fald det største forårskåde tal.

Opgavebesvarelser

Taus Brock-Nannestad

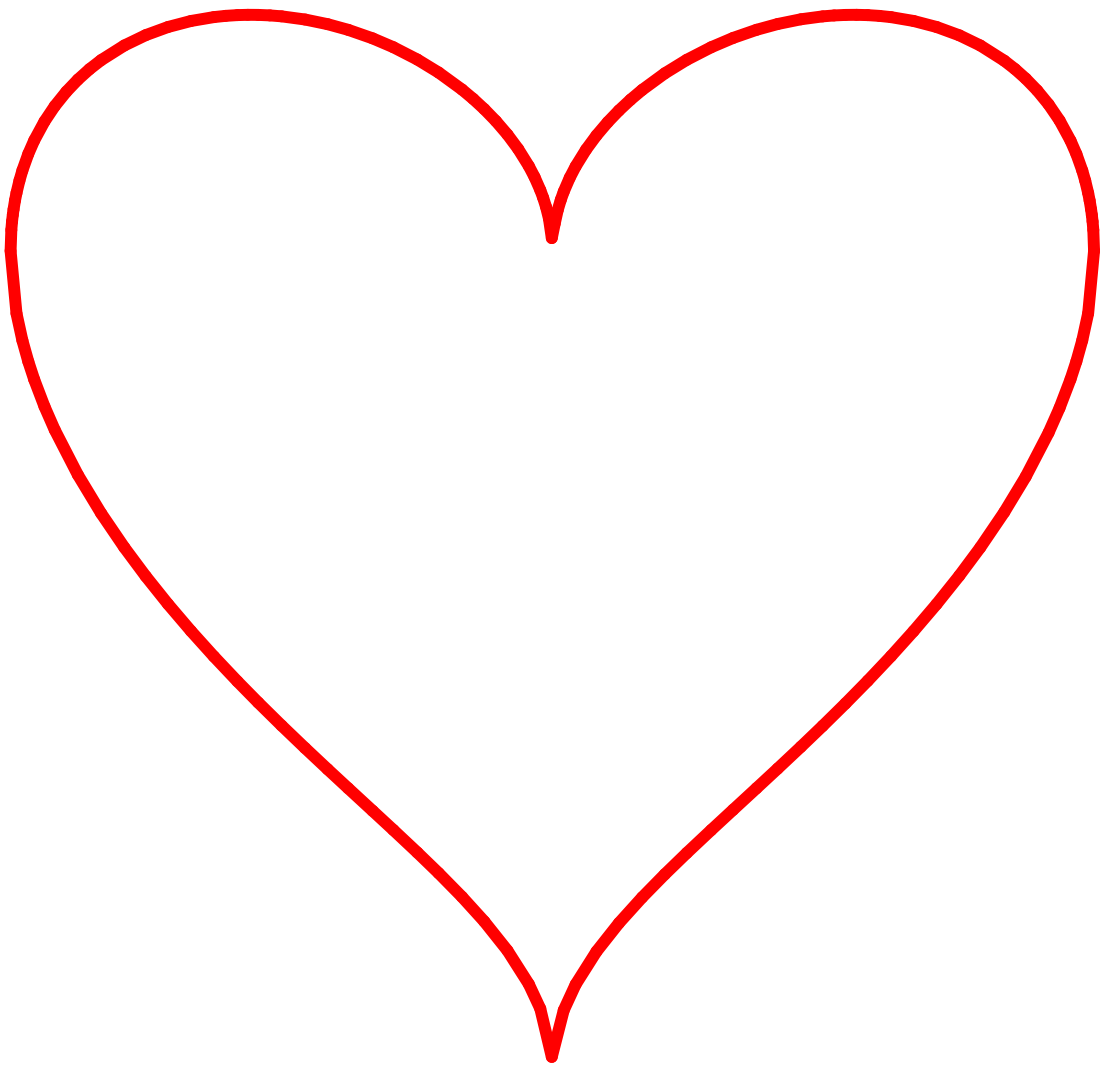
Løsningen til sidste nummers præmiepuzzle kan ses her:



Der blev modtaget korrekte besvarelser af: Kirsten Andersen, Louise Licht, Rune Kaasen, Pablo V. Holm-Nielsen, Helle Petersen og Mads Birkholm.

Efter konsultation med en passende tilfældighedsgenerator har redaktionen fundet frem til, at vinderen er Mads. Tillykke! Den mystiske præmie kunne dårligt være mere mystisk. Det drejer sig nemlig om en flaske Mystery. Den kan hentes hos redaktionen ved lejlighed.

```
> plot([sqrt(abs(x))+sqrt(1-x^2),sqrt(abs(x))-sqrt(1-x^2)],
   x=-1..1,color=red,axes=none,thickness=3);
```



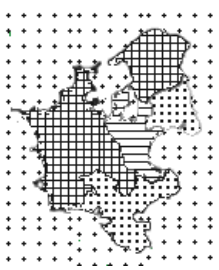
Fire farver er nok

Karl Malassé

100 år i tvivl

Et kort kan farves med fire farver, så områder, som har en grænse til fælles, får forskellige farver. Problemstillingen blev først formuleret af Francis Guthrie i 1852, og fik en afklaring, da det første bevis for 4-farve-sætningen kom i 1977. Det to-delte bevis kom først fra K. Appel, og W. Haken[1], og senere sammen med J. Koch[2] for den anden del. Et lignende, men nemmere bevis kom næsten 20 år efter fra Neil Robertson, Daniel P. Sanders, Paul Seymour og Thomas Robin[5]. Disse er de to eneste tilgængelige beviser for 4-farve-sætningen. Jeg gengiver her de begreber fra grafteori, som er nødvendige for forståelsen og giver et overordnet indblik i bevisernes fælles ramme.

Kort VS. graf



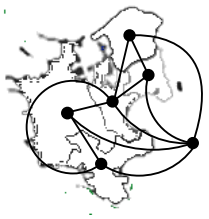
En farvning af Sjællands amtskort med fire farver (her er det mønstre) ses på figuren til venstre. Københavns Amt og Storstrøms Amt er prikket, Vestsjællands Amt og Frederiksborg Amt er gitret, Roskilde Amt er stribet, og det eksterne område (havet) er iført små krydser. Naboområder, altså områder, som har en grænse til fælles, har forskellige farver, og fire farver er også her det mindste antal farver, så dette er tilfældet. Roskilde amt, Københavns amt, og Frederiksborg amt er nemlig indbyrdes naboer, og alene dér skal der bruges tre farver. Disse områder er allesammen naboer til det eksterne område, og farven i dette område skal derfor være forskellig fra de tre øvrige. Resten af farvningen følger nemt.

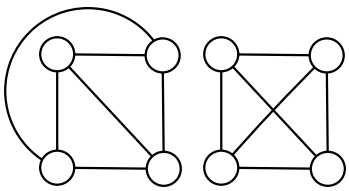
Skulle man bevise 4-farve-sætningen udfra ovenstående kort-formulering, skulle man definere formelt, hvad der forstås ved et kort. En definition vil medtage topologiske informationer, såsom afstand og områders indbyrdes placering. Disse oplysninger er irrelevante i vores sammenhæng. En farvning af et område forudsætter blot kendskab til naboernes farve. Vi kan derfor med fordel definere en symmetrisk relation $\mathcal{R}$ på mængden V af områderne, ved at to områder er i relation, såfremt de er naboområder. Udfra denne relation kan vi konstruere mængden $E = \{(v, w) \mid v\mathcal{R}w,\ v \in V,\ w \in V\}$. Parret $G = (V, E)$ kaldes **den duale graf til kortet.**

Duale grafer er blot en lille mængde af, hvad der normalt forstås ved grafer i grafteori. Givet en ikke-tom mængde V og en (muligvis tom) mængde

$E = \{(v, w) \mid v, w \in V\}$, er parret $G = (V, E)$ en **graf**. Elementerne i V kaldes normalt **knuder,** og V siges at være knudemængden. Elementerne i E kaldes normalt **kanter,** og E siges at være kantmængden. For os er elementerne i kantmængden uordnede par, hvorfor de grafer, vi har med at gøre, er uorienterede. Hvis kanten (v, w) findes, siges v og w at være **naboknuder**. En ækvivalent problemstilling af 4-farve-problemet er nu at farve knuderne i den duale graf med fire farver, så naboknuder ikke får samme farve.

En repræsentation af en graf kan have forskellige egenskaber. Særligt er vi interesseret i de grafer, som **kan** repræsenteres således, at grafens kanter ikke mødes andre steder end i fælles endeknuder. Disse typer grafer kaldes **planare** grafer, og en sådan repræsentation kaldes en planar repræsentation. Man begyndte formentligt at interessere sig for denne type grafer netop i forbindelse med 4-farve-problemet. Specielt er den duale graf til et kort en planar graf, og en planar repræsentation af den fås ved, at en hovedstad udpeges i hvert område, og to hovedstæder forbindes med en kant, såfremt deres respektive områder har en grænse til fælles. En planar repræsentation af den duale graf til Sjællands amtskort ses på figuren herover til højre.





Figur 1:

Om en graf er planar eller ej kan være svært at afgøre. At en graf er planar er ikke ensbetydende med, at samtlige repræsentationer af den er planare. Omvendt er det ikke fordi, der ikke findes en umiddelbar planar repræsentation af en graf, at den ikke er planar. På figur 1(a) er en planar graf illustreret med en planar repræsentation. På figur 1(b) er illustreret en ikke-planar repræsentation af den samme graf.

De planare grafer, som er relevant for os, er dem, som er mættet i kanter, d.v.s planare grafer, som mister planariteten ved at tilføje blot én kant. Grafer med denne egenskab kaldes for **maksimale planare grafer**. Den planare graf repræsenteret på figur 1(a) er en maksimal planar graf med fire knuder.

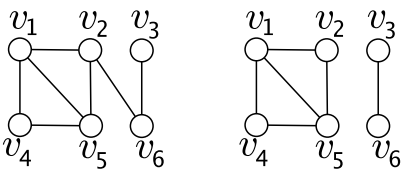
En planar tegning af en planar graf inducerer en naturlig inddeling af tegneplanen i **flader**. Fladerne har kanter fra kantmængden som grænse og udstyrer grafen med et topologisk aspekt. En planar tegning af en planar graf har således flader, men det har en graf ikke.

En 4-farvning af en graf er en farvning af grafens knuder med fire farver, så naboknuder får forskellige farver. Formuleringen af 4-farve-problemet i graf-versionen omfatter lidt flere grafer end de grafer, som kan være duale. Dog forbydes kanter af formen $(v, v), v \in V$, for ellers vil en knude være nabo til sig selv. En sådan kant kaldes en **løkke**. Problemformuleringen af 4-farve-problemet i graf-versionen betegnes fra nu af *4CT,* og er *som følger.*

Hovedsætning (4-farve-sætningen)**.** Enhver løkkefri planar graf kan 4-farves.

Minimale modeksempler

En **vej** er en følge af forskellige knuder $v_1, \ldots, v_n$, hvor kanterne $(v_i, v_{i+1})_{1 \leq i < n}$ er i kantmængden. En graf siges at være **sammenhængende**, såfremt der findes en vej mellem to vilkårlige knuder. Grafer, som er usammenhængende, består af flere komponenter. Den **komponent**, som en knude v er en del af, er den største sammenhængende delgraf, som indeholder v. Følgelig har en sammenhængende graf én komponent. En repræsentation af en sammenhængende graf ses på figur (a), medens figur (b) repræsenterer en usammenhængende graf. I den sidstnævnte findes der ingen vej mellem f.eks. v_1 og v_3. Delgrafen, bestående af knuderne v_3 og v_6, og kanten (v_3, v_6), er den største sammenhængende delgraf, som indeholder både v_3 og v_6 og er derfor en komponent. En anden komponent i den samme graf af delgrafen med knudemængden $V = \{v_1, v_2, v_4, v_5\}$ og kantmængden $E = \{(v_1, v_2), (v_1, v_5), (v_1, v_4), (v_5, v_2), (v_4, v_5)\}$.



(a) (b)

En **triangulering** er en repræsentation af en planar graf, hvor hver flade (også den uendelige flade) er en trekant, d.v.s. at de hver er begrænset med tre kanter. Figur1(a) kan man se et eksempel på en triangulering. Bemærk, at denne graf også er maksimal planar. Dette er ingen tilfældighed. Enhver planar repræsentation af en maksimal planar graf er en triangulering[4, 6].

Antag, at 4-farve-sætningen ikke gælder. Der findes altså planare grafer, som ikke kan 4-farves. Disse grafer er modeksempler til *4CT* og er derfor løkkefrie planare grafer. Betragt mængden af alle disse modeksempler. Man kan dele mængden i minimale klasser ved eksempelvis blot at betragte modeksemplerne med færrest knuder eller færrest kanter m.v., bare de er minimale i en eller anden forstand.

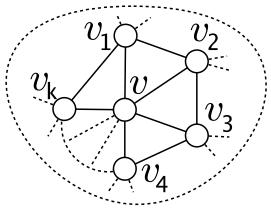
Til vores brug betragter vi de modeksempler, som har færrest knuder. De er **minimale** i antallet af knuder og kaldes af den grund **minimale modeksempler**. Graferne i denne klasse har følgelig samme antal knuder og har den egenskab, at planare grafer med færre knuder kan 4-farves, og dette er en direkte konsekvens af minimaliteten. Er en graf et minimalt modeksempel til *4CT*, medfører det dog ikke, at planare grafer med samme antal knuder også er det. En formel definition følger:

Definition 1. *Et* ***minimalt modeksempel*** *til* 4CT *er en planar graf G, som ikke kan 4-farves, og for enhver planar graf $G^\star$, som opfylder, at $|G^\star(V)| < |G(V)|$, kan $G^\star$ 4-farves.*

Når man nu vil vise, at *4CT* gælder, er vores mål at samle så mange oplysninger som muligt om minimale modeksempler og håbe på, at de er tilstrækkelige til at vise, at minimale modeksempler ikke findes.

En undersøgelse af minimale modeksempler viser, at det er maksimale planare grafer[6, 4], og dermed, at enhver planar repræsentation af dem er en triangulering.

Vi definerer **graden** af en knude v til at være antallet af kanter, som v er endeknude for. Følgelig er graden af en knude v også antallet af dens naboer, og en knude af grad k kaldes en k-knude. En **kreds** er en vej, som starter og ender i samme knude. Længden af en kreds er antallet af forskellige knuder i kredsen. En kreds af længde k kaldes en k-kreds.



At et minimalt modeksempel er maksimal planar medfører, at vi ved præcist, hvordan nabolaget til en k-knude ser ud. Naboerne danner nemlig en kreds af længde k, som illustreret på figuren til venstre. Fjerner man kredsens knuder fra grafen, og alle de kanter, som de er endeknuder for, får man en ny (usammenhængende) graf bestående af to komponenter, hvorfor kredsen er **adskillende**. I det tilfælde, som er illustreret på figuren, vil den ene komponent bestå af blot knuden v, da den adskillende kreds består af naboerne. Dette er imidlertid kun et specielt tilfælde, for kredse kan også adskille grafen i større komponenter.

En videreundersøgelse af minimale modeksempler viser, at de ikke har nogen adskillende kreds af længde tre eller fire. Det første tilfælde vises senere, mens beviset for den anden kan findes i [6, 4]. Endvidere ligger de eneste adskillende 5-kredse i minimale modeksempler omkring 5-knuder. En graf med denne egenskab kaldes **internt 6-sammenhængende**. Dette har til følge, at minimale modeksempler ingen knuder har af grad tre eller fire, for ellers vil deres naboer udgøre en adskillende kreds af længde hhv. tre eller fire.

Borte borte ...

At samle oplysninger om, hvordan et minimalt modeksempel ikke ser ud, spiller en central rolle i 4-farve-problemet. Vi betragter en delgraf G af et minimalt modeksempel T, og groft sagt 4-farver induktivt $T \backslash G$ og udvider denne 4-farvning til en 4-farvning af T. Induktionen starter ved grafer med højst fire knuder, som jo trivielt kan 4-farves. Hvis dette lykkes, kan denne fremgangsmåde påpege nogle delgrafer, som ikke kan være indeholdt i et minimalt modeksempel.

For at kunne bruge denne metode, betragter vi delgrafer med visse egenskaber. Antag, vi har en 4-farvning af G og betragt to knuder v_1 og v_2 i G. Hvis disse knuder er naboknuder i T, men ikke i G, har de måske fået samme farve i 4-farvningen af G, og vi får besvær, når vi skal udvide 4-farvningen af G til en 4-farvning af T. Vi ønsker derfor at undgå dette problem og definerer de delgrafer, vi ønsker:

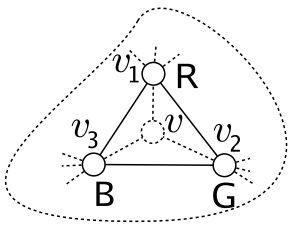
Definition 2. *G er en **induceret delgraf** af en maksimal planar graf T, såfremt*

G er en delgraf af T og hver kant i T med begge endeknuder i $G(V)$ er i G.

Når vi betragter en induceret delgraf, har vi brug for at vide, hvordan delgrafen er indeholdt i den oprindelige graf. En konfiguration er et værktøj, som udstyrer en induceret delgraf med oplysninger om knudernes grad i den oprindelige graf.

Definition 3. *En **konfiguration** er en induceret delgraf G af T udstyret med en afbildning $\gamma : G(V) \to \mathbb{N}$, så $\forall v \in G(V), \gamma(v) = grad_T(v)$, hvor $grad_T(v)$ er graden af knuden v i T.*

Definition 4. *En konfiguration (G, γ) **indgår** i T, såfremt G er en induceret delgraf af T og hver knude v i G har grad $\gamma(v)$ i T.*



Figur 2:

Et eksempel på en meget simpel konfiguration er en knude af grad tre. Vi viser her, at minimale modeksempler ingen 3-knude har. Lad derfor G_m være et minimalt modeksempel, og lad v være en 3-knude. Som illustreret på figuren , danner naboerne omkring v en 3-kreds, som er en adskillende 3-kreds, idet kredsens indre ikke er tom. Betragt nu grafen $G_m \setminus \{v\}$, hvor man fjerner v fra den oprindelige graf, samt alle de kanter som v er endeknude for. Denne graf har én knude færre end G_m og kan induktivt 4-farves, hvor naboerne til v får tre forskellige farver, da de er indbyrdes naboer. På figuren har naboerne fået farverne rød, grøn, og blå og repræsenteres med deres første bogstav. Denne 4-farvning kan udvides til en 4-farvning af G_m, hvor v får den fjerde farve. G_m kan derfor ikke være et minimalt modeksempel, og dermed har et minimalt modeksempel ingen 3-knude.

Formålet er at finde konfigurationer, som ikke kan indgå i et minimalt modeksempel. En konfiguration med denne egenskab kaldes **reducibel**. En knude af grad tre er en reducibel konfiguration. Om lidt viser jeg, at en anden konfiguration ved navn Birkhoffs diamant er en reducibel konfiguration. En (meget) determineret og (meget) opfindsom person kan dermed sætte sig ned og begynde at lede efter reducible konfigurationer. Men det er ikke nok. For selvom man kan danne et bibliotek af reducible konfigurationer, siger det kun noget om, hvordan et minimalt modeksempel **ikke** ser ud. Hvis man derimod kan vise, at mængden af de fundne reducible konfigurationer er **uundgåelig**, d.v.s at mindst en af konfigurationerne i mængden indgår i ethvert minimalt modeksempel, har man modbevist eksistensen af minimale modeksempler og har bevist 4-farve-sætningen.

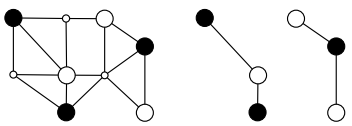
For dette er hele essensen i de to tilgængelige beviser for *4CT*. Både Appel, Haken og Koch i første omgang og *RSSR* senere fremviste en uundgåelig mængde af reducible konfigurationer. Den førstnævnte indeholdt 1476 konfigurationer,

mens den seneste "kun" indeholdt 633. Dér spillede EDB en stor rolle, for indlejrede Kempe-kæde argumenter sammen med manipulationer af grafer bliver hurtigt uoverskuelige. Reducibiliten af konfigurationerne sikres af et computerprogram. Ligeledes vises konfigurationerne uundgåeligt ved hjælp af en computer.

Diamant og kæder

Jeg viser om lidt reducibiliteten af en mere kompliceret konfiguration. Beviset kræver imidlertid et værktøj, kaldet Kempe-kæder, som jeg vil starte med at præsentere.

Betragt en 4-farvet planar graf G og to farver `a` og `b`. Delgrafen $G_{\texttt{a,b}}$, bestående af knuder farvet med `a` eller `b` og alle kanterne mellem dem, er selv en planar graf. Det er dog ikke sikkert, at $G_{\texttt{a,b}}$ er sammenhængende, og $G_{\texttt{a,b}}$ består derfor af én eller flere komponenter. Disse komponenter kaldes for `a-b` **Kempe-kæder**. Begrebet Kempe-kæder stammer fra Kempes forsøg på at bevise *4CT* i 1879 og bliver i høj grad brugt i behandlingen af 4-farve-problemet. Kempes "bevis" blev bredt anerkendt som værende et bevis til 4-farve-sætningen, men indeholdt imidlertid en fejl, det blev påpeget af P. J. Heawood i 1890.

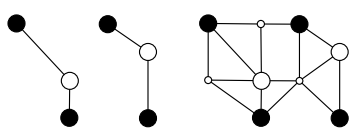


(a) (b)

Betragt en 4-farvning af en planar sammenhængende graf som illustreret på figuren (a) til venstre. Blandt de fire farver er farverne sort og hvid. På figuren er knuderne farvet med en af disse to farver repræsenteret med store hvide eller sorte cirkler. De øvrige knuder farvet med en af de to andre farver er repræsenteret ens med små cirkler. En planar repræsentation af delgrafen $G_{\texttt{sort,hvid}}$ ses på figur (b) og består af to komponenter. Hver af disse er en sort-hvid Kempe-kæde.

Kempe-kæder bruges til at ændre en 4-farvning af en graf ved et bytte om på farverne i en hel Kempe-kæde. For sådan, som vi har konstrueret dem, har knuderne i `a-b` Kempe-kæde ingen nabo i den oprindelige graf farvet `a` og `b`, og som ikke allerede er en del af den samme `a-b` Kempe-kæde. Dette betyder formelt, at man kan permutere farverne `a` og `b` i en `a-b` Kempe-kæde uden at bryde 4-farvningen af den oprindelig graf. Man får naturligvis en anden 4-farvning af grafen, men ikke desto mindre er denne 4-farvning gyldig.

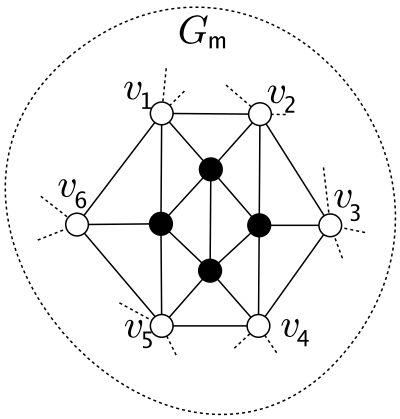
Vi fortsætter med vores eksempel og bytter om på farverne sort og hvid i den ene sort-hvide Kempe-kæde. Den anden lader vi stå uændret. En repræsentation af den nye farvning ses på figur (a). Den oprindelige 4-farvning af grafen ændres tilsvarende og illustreres på figur (b).



(a) (b)

Som et eksempel på en mere avanceret reducibel konfiguration betragter vi Birkhoffs diamant. Birkhoffs diamant består af en 5-knude med tre på hinanden følgende naboknuder af grad fem. Disse knuder er repræsenteret ved sorte knuder

i figuren nedenfor til højre. Denne konfiguration er opkaldt efter G. D. Birkhoff[3], som har vist i 1913, at denne konfiguration ikke kan indgå i et minimalt modeksempel. Jeg gengiver her hans argumenter.

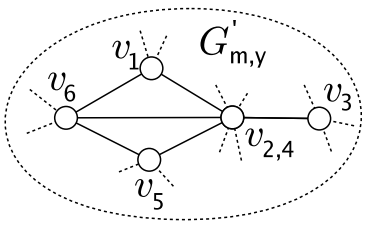


Antag, at en sådan konfiguration er indeholdt i et minimalt modeksempel G_m, altså, at grafen G_m indeholder en 5-knude, som har tre naboer af grad fem. En planar repræsentation af G_m illustreres på figuren til højre og konstrueres udfra, at G_m er maksimal planar. På denne repræsenation har de sorte knuder grad fem. Lad C betegne 6-kredsen $v_1, v_2, v_3, v_4, v_5, v_6, v_1$. Vi ved ikke noget om graderne af knuderne i C, men ved derimod, at kredsen C er en adskillende 6-kreds.

Betragt grafen $G_{m,y}$, hvor man fjerner det indre af 6-kredsen i G_m, altså, hvor man fjerner 5-knuderne fra Birkhoffs diamant (de sorte knuder) og alle de kanter, som de er endeknuder for. Vi konstruerer en ny graf $G'_{m,y}$ udfra $G_{m,y}$ på følgende måde:

- Knudemængden af $G'_{m,y}$ er $G_m(V) \setminus \{v_2, v_4\} \cup \{v_{2,4}\}$.
- Kanterne af $G_{m,y}$, som hverken har v_2 eller v_4 som endeknude, er med i kantmængden af $G_{m,y}$, medens naboerne til v_2 og v_4 i $G_{m,y}$ bliver naboer til $v_{2,4}$ i $G'_{m,y}$.
- Kanten $(v_{2,4}, v_6)$ tilføjes kantmængden.

Den nye graf $G'_{m,y}$ ser indviklet ud, men er faktisk resultatet af en ganske almindelig manipulation i grafteori. Før kanten $(v_{2,4}, v_6)$ er tilføjet, har vi blot **identificeret** knuderne v_2 og v_4 i knuden $v_{2,4}$. Dette betyder, at vi har smeltet knuderne sammen i én og samme knude, og at de identificerede knuders naboer bliver naboer til den nye knude.



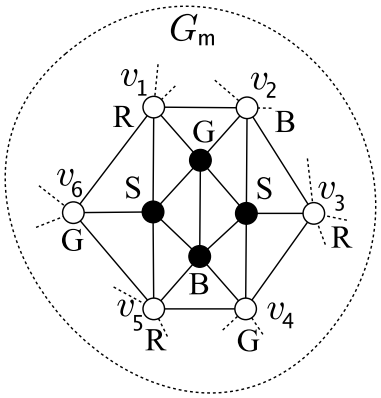
Den planare repræsentation af $G'_{m,y}$ til venstre kan hjælpe læseren med at indse resultatet af manipulationerne. Jeg påstår nu, at $G'_{m,y}$ er en løkkefri og planar graf. Hvis v_2 og v_4 var naboknuder i G_m, vil planariteten af G_m tvinge v_3 til at være en 3-knude, hvilket vi har vist, ikke er indeholdt i et minimalt modeksempel. Derfor findes kanten $(v_{2,4}, v_{2,4})$ ikke, hvorfor $G'_{m,y}$ er løkkefri. Planariteten sikres af figuren til venstre. Endvidere har $G'_{m,y}$ færre knuder end G_m. Grafen $G'_{m,y}$ kan derfor 4-farves induktivt, og vi betragter en 4-farvning af den. Da $v_1, v_6, v_{2,4}$ er indbyrdes naboer, skal tre af farvene bruges på dem. Har vi farvene rød, sort, blå og grøn til at farve med, kan vi antage, at v_1 eksempelvis er farvet rød, $v_{2,4}$ er farvet blå, og at v_6 er farvet grøn. Der

er ikke så mange muligheder for de andre knuder. Knuden v_3 har fået en farve blandt rød, sort og grøn, og knuden v_5 har enten fået rød eller sort.

En 4-farvning af $G'_{m,y}$ inducerer en foreløbig delvis 4-farvning af G_m, hvor knuderne v_2 og v_4 får samme farve. Det er ikke et problem, for et ovenstående argument sikrer, v_2 og v_4 ikke kan være naboknuder i G_m. Vi mangler nu blot at farve det indre af kredsen, og det er lige netop det, vi gør nu.

Vi kan ikke forudsige mere og er derfor nødt til at kigge på alle de seks muligheder, som de mulige farvninger af v_3 og v_5 medfører. Disse muligheder er illustreret på figur 3. Et hurtigt tjek viser, at en udvidelse af den delvis 4-farvning til den indre af kredsen er umiddelbar i alle tilfældene på nær i tilfældet, hvor v_3 og v_5 er farvet rød (tilfældet (c)). Her bliver vi nødt til at anvende et Kempe-kæde argument.

Hvis v_3 og v_5 er en del af den samme rød-sorte Kempe-kæde, består denne kæde af knuder i det ydre af C. Den grøn-blå Kempe-kæde, som v_4 er en del af, vil derfor ikke have fælles knuder med den rød-sorte Kempe-kæde, som v_3 og v_5 er en del af. Derfor sikrer planariteten af G_m, at v_2 og v_4 ikke er en del af den samme grøn-blå Kempe-kæde. Vi kan derfor bytte om på farvene i den grøn-blå Kempe-kæde, som v_4 er en del af, og vi får en ny delvis 4-farvning af G_m, hvor v_4 er grøn, og vi kan udvide denne farvning til en 4-farvning af G_m, som illustreret på figuren til højre.



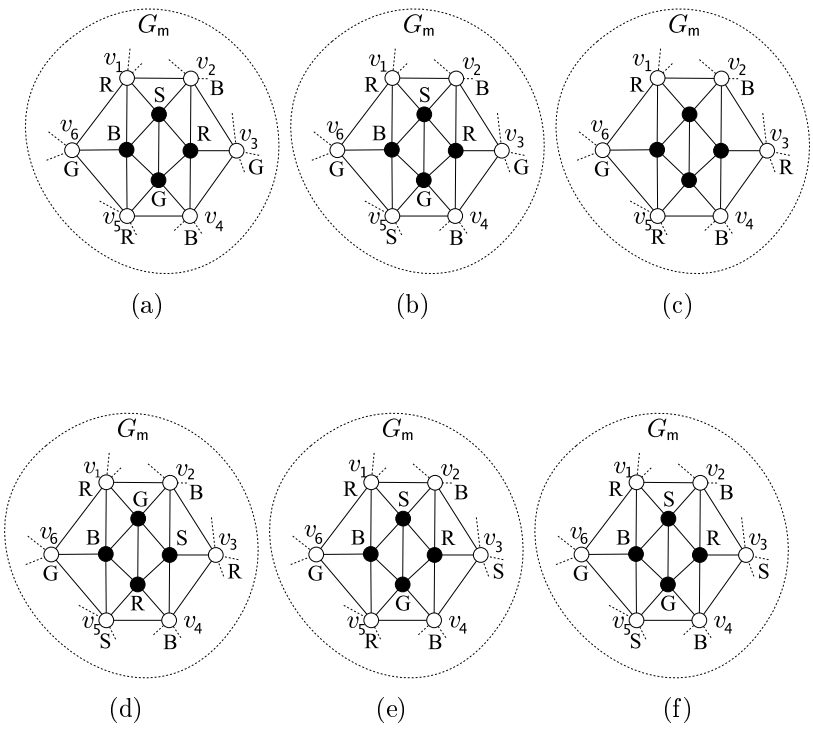
Hvis v_3 og v_5 derimod ikke er i den samme rød-sorte Kempe-kæde, betragter vi den rød-sorte Kempe-kæde, som v_3 er en del af. Hvis ikke v_1 er en del af den, bytter vi om på farvene og får en ny delvis farvning af G_m, hvor v_3 er i sort. Vi kan nu udvide til en 4-farvning af G_m som i tilfældet (e) illustreret på figur 3.

Vi mangler nu det tilfælde, hvor v_1 og v_3 er en del af den samme rød-sorte Kempe-kæde, og v_5 er ikke en del af den. I dette tilfælde kan vi bytte om på farvene i den rød-sorte Kempe-kæde, som v_5 er en del af. I denne nye farvning er v_5 i sort, hvorfor vi ender med tilfældet (d) af figur 3.

Vi har været alle mulighederne igennem og kan dermed konkludere, at Birkhoffs diamant ikke kan indgå i et minimalt modeksempel, og at det derfor er en reducibel konfiguration.

Litteraturliste

[1] K. Appel and W. Haken. Every planar map is four colorable part I: Discharging. *Illinois J. Math*, 21:429–490, 1977.



Figur 3:

[2] K. Appel, W. Haken, and J. Koch. Every planar map is four colorable part II: Reducibility. *Illinois J. Math*, 21:491–567, 1977.

[3] G. Birkhoff. The reducibility of maps. *Amer. J. Math.*, 35:114–128, 1913.

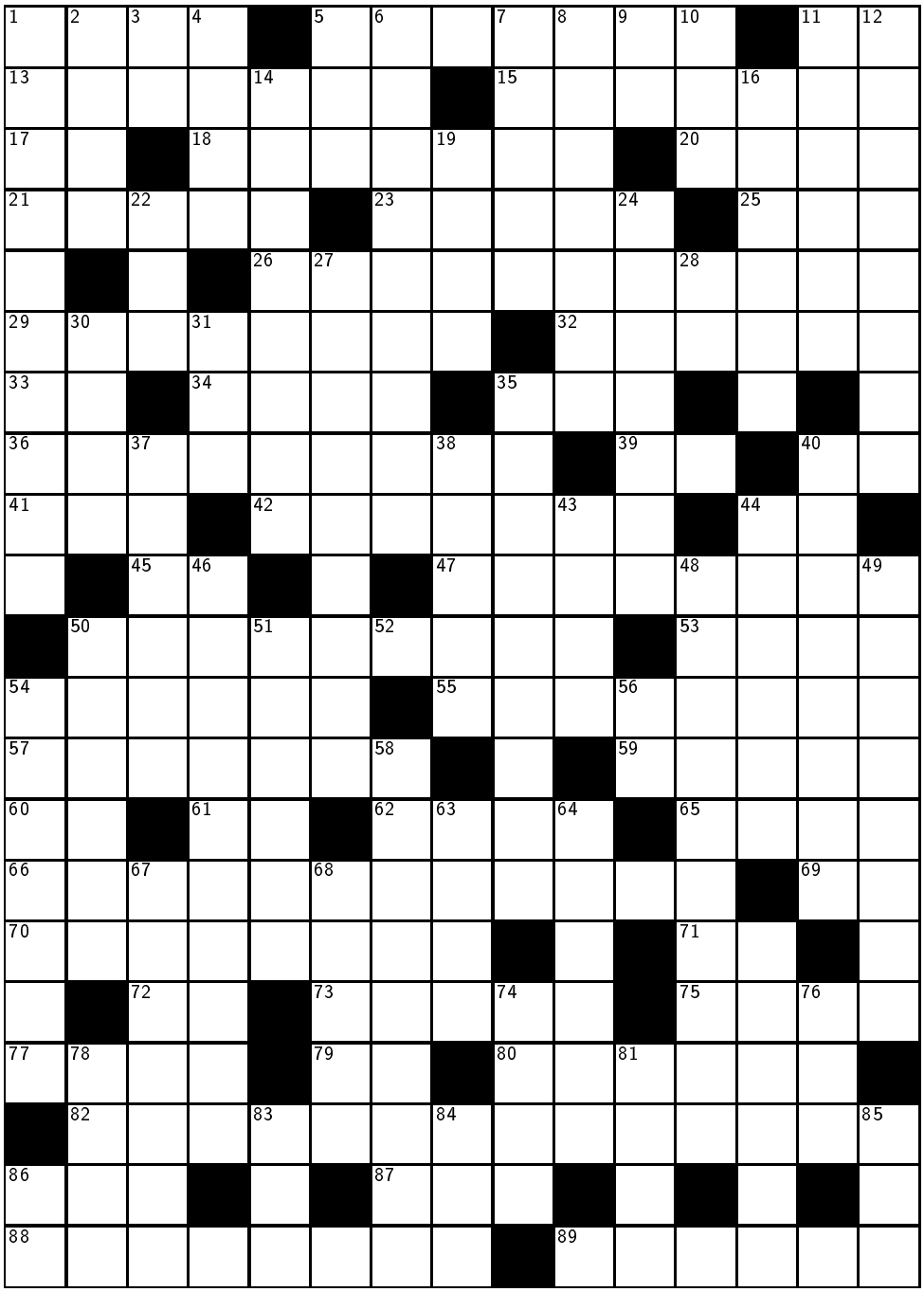
[4] R.A. Wilson. *Graphs, colourings and the four-colour theorem.* Oxford University Press, 2002.

[5] Neil Robertson, Daniel Sanders, Paul Seymour, and Robin Thomas. The four-colour theorem. *J. Comb. Theory Series B*, 70(1):2–44, 1997.

[6] Thomas Saaty and Paul Kainen. *The four-color problem.* McGraw-Hill, New York, 1977.

Efterårskryds

Martin "Damskur" Damhus



Vandret:

1: Muslimsk pilgrimsfærd
5: Led med stor bevægelighed
11: Biord, optræder gerne i danske infinitivkonstruktioner
13: 2001 – 2006
15: Skønfarvning
17: Bilkode for Holland
18: Udstøde artskarakteristisk lyd
20: Bibelsk eks-beboer i Edens Have (der var noget vrøvl med æblerov)
21: ***** Minogue: Søster til Kylie
23: 4 øjne fordelt ligeligt på to terninger
25: Virksomhed, der administrerer arbejdsmarkedets tillægspension
26: New York
29: Frafald, typisk i betydningen religionsfornægtelse
32: Komedie af Voltaire
33: Spil
34: Fransk-baseret NGO med base i Kabul, der arbejder (bl.a. gennem uddannelse af journalister) for at støtte udviklingen af en fri og uafhænging afghansk presse
35: Mobilstandarden „Globalt system for mobil kommunikation“
36: $\forall$
39: Kvinde
40: Fonetisk er ** en affrikativ konsonantisk lyd, der kunne erstatte „zz“ i „pizza“
41: Indisk by; techno-genre
42: Instant-drik af høj nyttighed for studerende
44: Findes
45: Tidsrum på 10^{-9} s
47: Som træffer beslutninger kun på basis af egne vurderinger
50: Komet, der blev opdaget 20/1–1996 og opkaldt efter sin opdager, en japansk amatørastronom
53: Asiatisk land
54: Afrikansk land
55: Som fastsætter hvordan noget bør være, modsat „deskriptiv“
57: Ressource (en sådan lovprises fx. indenfor management)
59: Gal streg
60: Landekode for stillehavsøstaten Republikken Kiribati, der fik sin uafhængighed i 1979
61: ** masse: I massevis
62: Lille kystby på nordsiden af Guadalcanal-øen, Salomonøerne, Oceanien
65: Ugerning, hvorunder en eller flere personer ihjelslås
66: Lydisolerende
69: Forkortelse for „ditto“
70: Essentiel aminosyre, som eksempelvis sesam og mange slags kød er rig på
71: Avis
72: Erbium
73: Bær-bærende busk
75: Titulatur for religiøs entitet, der hyldes i gospelmusik; **** Voldemort: Magtgal mørkets troldmand i „Harry Potter“-bøgerne
77: Hønsefugl med fjerklædte fødder og tæer (latin: lagopus)
79: Strøm; ** Salvador: Land
80: Anordning, der giver lys på cyklen uden brug af batterier
82: Uformuende; i overført betydning: om en person, der kun har meget få brikker at rykke rundt med
86: Landområde, der mod visse forpligtelser tildeles en adelsmand af en højerestående, fx. kongen
87: Megagram
88: Territorium af omskiftelig størrelse, der gennem tiderne har inkluderet dele af N-Tyskland, N-Polen, den nuværende Kaliningrad-region, samt Litauen
89: Centraleuropæisk land

Lodret:

1: Lille rygsæk
2: Svensk-dansk andelsmejeriselskab, der for nylig ramtes af større handelsboykot i mellemøsten
3: Junior
4: Gennemsnitlig
5: August Krogh Instituttet; japansk fornavn
6: Parameter (fx. diskontoen), der fastlægger långivers udbytte
7: Våben
8: Stemme *******
9: Dansk Industri
10: Underverdenen i polynesisk mytologi; beskatningstype i fx. Spanien, Latinamerika og Mexico („Impuesto al Valor Agregado“)
11: Halvædelstenen
12: Hellige bygningers
14: Drengenavn
16: Alvorsfuldt musikstykke i langsomt tempo
19: Partier (af fx. musikstykke), der udføres af blot en enkelt person
22: Ny-
24: Hollandsk ost
27: Dukke
28: ** excelsis: I det høje
30: Sport
31: Redskab
35: Nåletræsfrugt, opbygget af fortykkede blade, der beskytter frøene
37: Hul nål
38: Vandmasse
40: Hovedø i ø-republik i det sydlige Caribiske Hav, 11 km nord for Venezuelas kyst
43: Eventyrlige hunkønsvæsner
44: Tekstbehandlings-program
46: Bilindehaveren
48: Impressionist-motiv
49: Fødselsdagsmøbel
50: Kort beretning fra profeten Muhammeds liv
51: Afrikansk flod med udspring i Angolas højland, som flyder mod syd og ind i Namibia
52: TRUTTERUT! Foræringsordforklaringen! Men kun for nørder. $\texttt{ASCII}(*) = (84, 116)$.[1]
54: Blanding
56: LaTeX-kode for μ; kemisk symbol for muoniumpartikel; landekode for Mauritius
58: Varm (helst) drik
63: Få
64: Unavngiven
67: Lidelse
68: Bemærker
74: Paradis
76: Næringsrig rodfrugt
78: Surmælksprodukt
81: Pyntegenstande
83: Slummer
84: *** appetit
85: Bygningsdel; metadata (information om data, adskilt fra selve data, fx. på hjemmesider); graffiti-underskrift
86: Plade

[1] `ASCII` er her afbildingen af det engelske alfabet $\{\mathrm{a}, \mathrm{b}, \mathrm{c}, \ldots, \mathrm{z}\}$ ind i $\mathbb{N} \times \mathbb{N}$, så

$$\texttt{ASCII}(x) = \big(\text{decimal ASCII}^{2}\text{-kode for } x \text{ (upper case)}, \text{decimal ASCII-kode for } x \text{ (lower case)}\big).$$

[2] ASCII er her *American Standard Code for Information Interchange.*

Stabilitet i fysiske modeller

Jan Philip Solovej

Indledning

Man skulle måske tro, at det er enhver forskers drøm, når man, som det ret ofte sker, bliver stillet spørgsmålet "kan du forklare, hvad du egentlig forsker i?" Det er da også rigtigt, at der ikke er noget, en forsker hellere vil end fortælle, hvad det er man pusler med og hvad det er, der kører rundt i hovedet på en i mange af døgnets timer. Desværre er det for de fleste matematikere en temmelig utaknemmelig opgave at skulle fortælle om abstrakt matematisk forskning til ganske almindelige mennesker. Jeg har måske en lidt lettere opgave end mange andre matematikere, fordi jeg beskæftiger mig med matematisk fysik og kan beskrive det jeg forsker i i fysiske termer uden nødvendigvis at skulle komme ind på abstrakt matematik. Jeg forsøger mig normalt ved først at forklare, at det jeg forsker i er at forstå, hvorfor alt det, vi ser omkring os, har den størrelse, det har og hvorfor de fleste objekter omkring os er stabile. Hvorfor får man f.eks. to liter vand, når man hælder to gange en liter vand sammen?

Desværre når jeg sjældent meget længere i min forklaring, inden "åh nej! Hvorfor spurgte jeg?" står malet med skræk i den stakkels spørgers ansigt.

Min opgave her er betydelig lettere. Jeg regner med, at du kære læser ikke er helt almindelig (en kompliment!) og vil kunne fordøje en god portion avanceret matematik. Jeg vil forsøge at starte, hvor en førsteårs studerende i matematik vil kunne følge med. Om det er lykkedes mig at holde hele beskrivelsen på det niveau, ved jeg ikke.

Kræver det stor indsigt i fysik at læse denne artikel? Det skader ikke, men jeg vil forsøge ikke at antage noget større kendskab til fysik. Jeg vil give matematiske definitioner på de begreber der optræder. På den anden side vil jeg ikke gøre noget forsøg på alt for lange forklaringer af de fysiske motivationer for definitionerne.

Et af de spørgsmål, der har interesseret mig igennem det meste af min forskningskarriere, er, om de fysiske modeller for mikrokosmos virkelig kan forklare den makroskopiske verden. Min angrebsvinkel er relativt beskeden. Jeg har ingen ambitioner om at kunne beskrive detaljerede materielle egenskaber eller biologiske organismer. Spørgsmålet er på en måde mere overordnet: Kan de fysiske modeller for mikrokosmos overhovedet forklare eksistensen af en makroskopisk verden? Jeg håber i denne artikel at illustrere, at dette langt fra er oplagt.

Når jeg taler om fysiske modeller, så opfatter jeg det som matematiskt formulerede teorier, hvis objekter kan fortolkes i fysiske termer. Det er selvfølgeligt et

interessant spørgsmål om disse teorier beskriver virkeligheden korrekt, men det er uden betydning for den matematiske analyse. Alle kendte fysiske teorier har deres begrænsninger i beskrivelsen af virkeligheden. Som matematiske modeller kan man studere dem udover disse begrænsninger og i mange tilfælde få en forståelse for, hvorfor de er begrænsede. Lad mig beskrive den model jeg først vil diskutere og specielt hvilke begrænsninger den har.

Jeg vil tage udgangspunkt i beskrivelser af verden som bestående af positivt ladede atomkerner og negativt ladede elektroner. Atomkernerne består i sig selv af protoner og neutroner, men det har kun betydning ved energier langt ud over, hvad man normalt finder i vores omgivelser. For ikke at gøre det hele for kompliceret ser vi bort fra kernernes indre struktur (og fra eksistensen af andre partikler, hvis eksistens kun har betydning ved meget høje energier).

Ladede partikler vekselvirker med hinanden gennem elektriske og magnetiske kræfter. Da partiklerne desuden har masse, tiltrækker de også hinanden ved tyngdekraften.

Både den magnetiske kraft og tyngdekraften er langt svagere end de elektriske kræfter, derfor vil jeg også til at starte med begrænse mig til at se bort fra disse. Tyngdekraften og magnetiske kræfter har betydning for, hvordan verden omkring os ser ud og at se bort fra dem er en stor begrænsning, derfor vil jeg senere vende tilbage til de problemer, der er involveret i at inkludere dem.

Kvantemekanik og brintatomets stabilitet

To elektriske partikler siddende i punkterne $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^3$ (det tre-dimensionale rum) med ladninger $q_1, q_2 \in \mathbb{R}$ har en elektrisk energi givet ved Coulombs lov

$$\frac{q_1 q_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|}.$$

Hvis vi forestiller os et atom (brintatomet) bestående af en kerne med ladning $+1$ og masse M og en elektron med masse 1 og ladning[1] -1 vil energien i klassisk mekanik være givet ved følgende funktion af kernens position $\boldsymbol{X} \in \mathbb{R}^3$ og impuls $\boldsymbol{P} \in \mathbb{R}^3$ og elektronens position $\boldsymbol{x} \in \mathbb{R}^3$ og impuls $\boldsymbol{p} \in \mathbb{R}^3$

$$E(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{P}) = \frac{1}{2M}\boldsymbol{P}^2 + \frac{1}{2}\boldsymbol{p}^2 - \frac{1}{|\boldsymbol{x} - \boldsymbol{X}|}.$$

Problemet er, at denne funktion ikke er defineret for $\boldsymbol{x} = \boldsymbol{X}$, men mere alvorligt er det, at funktionen ikke er nedadtil begrænset. Det betyder, at den klassiske beskrivelse af et atom muliggør, at det kan have vilkårlig negativ energi. Et sådant atom er ikke stabilt, veksevirkning med omgivelserne vil tillade atomet at afgive vilkårlige store energimængder og elektronen vil falde tættere mod kernen.

[1] Vi vælger at måle alle masser i enheder af elektronmassen og alle ladninger i enheder af minus elektronladningen

Det er et af de problemer, der blev løst af den kvantemekaniske beskrivelse af atomare partikler. Lad mig kort skitsere, hvordan energien af brintatomet beregnes i kvantemekanik. Her beskrives et brintatom ikke ved at angive sted og impuls koordinaterne for elektronen og atomkernen. I stedet beskrives atomet ved at angive en funktion (bølgefunktionen) $\psi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{C}$ afhængig af elektronens og kernens stedkoordinater $(\boldsymbol{x}, \boldsymbol{X})$. I virkeligheden afhænger bølgefunktionen af to yderligere parametre: elektronens spin og kernens spin. Spinnet har stor betydning for fysikken, men ikke umiddelbart for de matematiske aspekter jeg vil diskutere her (sålænge der ses bort fra magnetiske felter). Jeg vil derfor, for ikke at vanskeliggøre notationen for meget, tillade mig at ignorere spinnet[2]. Energien i kvantemekanik afhænger af den funktion ψ, der beskriver atomets tilstand.

Det er ikke alle funktioner ψ, der er relevante. Det er et ikke helt nemt matematisk spørgsmål at præcisere den rigtige klasse af funktioner og den kan endda afhænge af, det fysiske system man betragter. For brintatomet er den "rigtige" klasse af funktioner et såkaldt Sobolev rum, men det vil jeg ikke benytte mig af her. I stedet vil jeg fokusere på den mindre mængde af funktioner, der består af *de en gang kontinuert differentiable funktioner, som er nul uden for en begrænset mængde*. Jeg bruger notationen C_0^1 for denne klasse (superskriptet 1 står for antal gange funktionerne er differentiable og subskriptet 0 står for, at de er nul uden for en begrænset mængde). For brintatomet er funktionsklassen altså $C_0^1(\mathbb{R}^3 \times \mathbb{R}^3)$. Desuden skal funktionen ψ opfylde *normaliseringen*

$$\iint |\psi(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} = 1. \tag{1}$$

Denne normalisering os tillader at fortolke $|\psi(\boldsymbol{x}, \boldsymbol{X})|^2$ som sandsynlighedsfordelingen for elektronens og kernens positioner.

For brintatomet er energien udtrykt ved ψ givet ved[3]

$$\begin{aligned} \mathcal{E}_{\mathrm{H}}(\psi) &= \frac{1}{2M} \iint |\nabla_{\boldsymbol{X}} \psi(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} + \frac{1}{2} \iint |\nabla_{\boldsymbol{x}} \psi(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} \\ &\quad - \iint |\psi(\boldsymbol{x}, \boldsymbol{X})|^2 \frac{1}{|\boldsymbol{x} - \boldsymbol{X}|} \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X}. \end{aligned}$$

Her er $\nabla_{\boldsymbol{x}}\psi$ og $\nabla_{\boldsymbol{X}}\psi$ gradienterne mht. de variable $\boldsymbol{x}$ og $\boldsymbol{X}$. Man ser relativt nemt, at $\mathcal{E}_{\mathrm{H}}(\psi)$ er veldefineret for $\psi \in C_0^1(\mathbb{R}^3 \times \mathbb{R}^3)$.

Definition 1 (Brintatomets grundtilstandsenergi)**.** *Brints grundtilstandsenergi er defineret ved*

$$E_{\mathrm{H}} = \inf \left\{ \mathcal{E}_{\mathrm{H}}(\psi) \;\middle|\; \psi \in C_0^1(\mathbb{R}^3 \times \mathbb{R}^3), \iint |\psi(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} = 1 \right\}.$$

[2]Elektronens spin inkluderes ved blot at lade $\boldsymbol{x} \in \mathbb{R}^3 \times \{1, -1\}$

[3]I energiudtrykket er der, udover de allerede valgte enheder, valgt enheder, hvor Planck's konstant $\hbar = 1$. D.v.s. at længde måles i Bohr radius, som er ca. $0,5 \times 10^{-10}$meter, energi måles i 2Rydberg som er ca. $4,4 \times 10^{-18}$Joule og tid måles i en enhed som er ca. $2,4 \times 10^{-17}$sekunder.

Stabilitet af brintatomet er påstanden $E_{\mathrm{H}} > -\infty$, altså at energien ikke kan være vilkårlig negativ. Denne ulighed er på ingen måde oplagt, men det er så heldigt, at man faktisk kan udregne *energien af brint* eksakt.

Sætning 2 (Brinatomets energi)**.**

$$E_{\mathrm{H}} = -\frac{1}{2}\frac{M}{M+1}. \tag{2}$$

Schrödinger var den første til at indse dette tilbage i 1927 (selvom han måske ikke havde et fuldstændigt bevis). For bl.a. dette resultat modtag Schrödinger i 1933 nobelprisen[4] i fysik. Niels Bohr havde allerede tidligere forklaret stabiliteten af brint (belønnet med nobelprisen i 1922), men ikke så matematisk veldefineret som Schrödinger.

Stabilitet og usikkerhedsrelationen

Man er ikke altid så heldig i matematik at kunne finde eksakte udtryk for relevante størrelser, som f.eks. E_{H}. Man er derfor interesseret i at kunne give vurderinger altså at finde uligheder, hvor størrelserne indgår.

Man vil ofte i fysikbøger se stabiliteten af brint forklaret ud fra Heisenbergs (nobelprisen i fysik 1932) berømte usikkerhedsrelation mellem sted og impuls. Med notationen fra forrige afsnit formuleres usikkerhedsrelationen normalt ved uligheden

$$\iint_{\mathbb{R}^3\times\mathbb{R}^3} |\nabla_{\boldsymbol{x}}\psi(\boldsymbol{x},\boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} \iint_{\mathbb{R}^3\times\mathbb{R}^3} |\boldsymbol{x}-\boldsymbol{X}|^2 |\psi(\boldsymbol{x},\boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} \geq \frac{1}{4},$$

som holder for enhver funktion $\psi \in C_0^1(\mathbb{R}^3\times\mathbb{R}^3)$ med $\iint |\psi(\boldsymbol{x},\boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} = 1$. Det første dobbeltintegral ovenfor fortolkes som forventningsværdien af kvadratet på elektronens impuls og det andet dobbeltintegral som forventningsværdien af kvadratet på afstanden mellem elektronen og kernen. Uligheden udtrykker bl.a. det forhold, at elektronen ikke kan side vilkårligt tæt ved kernen og samtidig have vilkårlig lille impuls. Ved hjælp af usikkerhedsuligheden ser man, at energiudtrykket for brint opfylder

$$\mathcal{E}_{\mathrm{H}}(\psi) \geq \frac{1}{8}\left(\iint |\boldsymbol{x}-\boldsymbol{X}|^2 |\psi(\boldsymbol{x},\boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X}\right)^{-1} - \iint \frac{|\psi(\boldsymbol{x},\boldsymbol{X})|^2}{|\boldsymbol{x}-\boldsymbol{X}|}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X}.$$

Denne ulighed forklarer dog *ikke* stabiliteten af brint. Det overlades til læseren at indse dette ved at løse følgende opgave.

[4]Nobelprisens webside: www.nobelprize.org indeholder ofte relativt forståelige forklaringer af den fysik, der ligger til grund for prisen og biografier af prismodtagerne.

Opgave 3. Find en følge af funktioner $\psi_n \in C_0^1(\mathbb{R}^3 \times \mathbb{R}^3)$ med

$$\iint |\psi_n(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} = 1$$

således, at

$$\frac{1}{8}\left(\iint |\boldsymbol{x}-\boldsymbol{X}|^2 |\psi_n(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X}\right)^{-1} - \iint \frac{|\psi_n(\boldsymbol{x}, \boldsymbol{X})|^2}{|\boldsymbol{x}-\boldsymbol{X}|}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} \to -\infty$$

når $n \to \infty$.

Stabiliteten af brint følger altså ikke af den sædvanlige formulering af usikkerhedsrelationen, men den følger af nedenstående stærkere formulering.

Sætning 4 (Hardys usikkerhedsrelation)**.** *For alle funktioner $f \in C_0^1(\mathbb{R}^3)$ gælder*

$$\int_{\mathbb{R}^3} |\nabla_{\boldsymbol{x}} f(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} \geq \frac{1}{4}\int_{\mathbb{R}^3} \frac{1}{|\boldsymbol{x}|^2} |f(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x}.$$

Der findes tilsvarende uligheder i alle dimensioner større end 3 men ikke i lavere dimensioner.

Denne ulighed, som vi ikke beviser her, giver

$$\begin{aligned}\iint_{\mathbb{R}^3\times\mathbb{R}^3} |\nabla_{\boldsymbol{x}}\psi(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} &\geq \frac{1}{4}\iint_{\mathbb{R}^3\times\mathbb{R}^3} |\boldsymbol{x}-\boldsymbol{X}|^{-2} |\psi(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} \\ &\geq \frac{1}{4}\left(\iint_{\mathbb{R}^3\times\mathbb{R}^3} |\boldsymbol{x}-\boldsymbol{X}|^{-1} |\psi(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X}\right)^2.\end{aligned}$$

for alle $\psi \in C_0^1(\mathbb{R}^3\times\mathbb{R}^3)$ med $\iint |\psi(\boldsymbol{x}, \boldsymbol{X})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{X} = 1$. Den sidste ulighed ovenfor er en konsekvens af Cauchy-Schwarz ulighed (for integraler) eller Jensens ulighed. Ved brug af ovenstående vurdering er det nu nemt at bevise stabiliteten af brint, d.v.s. at, $E_{\mathrm{H}} > -\infty$, men man får ikke den eksakte værdi fra Sætning 2.

En på mange måder mere interessant ulighed, der også udtrykker usikkerhedsrelationen og også giver stabiliteten af brint, er Sobolevuligheden, som jeg igen nævner uden bevis.

Sætning 5 (Sobolevs ulighed)**.** *For alle funktioner $f \in C_0^1(\mathbb{R}^3)$ er*

$$\int_{\mathbb{R}^3} |\nabla_{\boldsymbol{x}} f(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} \geq \frac{3}{4}(2\pi^2)^{2/3}\left(\int_{\mathbb{R}^3} |f(\boldsymbol{x})|^6 \mathrm{d}\boldsymbol{x}\right)^{1/3}. \tag{3}$$

Der findes tilsvarende uligheder i alle dimensioner større end 3 men ikke i lavere dimensioner.

Makroskopisk stabilitet

Stabiliteten af brint generaliserer nemt til stabilitet af andre atomer og molekyler. Men hvad med makroskopiske objekter der nemt indeholder over 10^{23} partikler? Fra et matematisk synspunkt er vi interesserede i at forstå energien, når antallet af partikler går mod uendelig.

Lad os betragte et system bestående af $2N$ partikler ($N \in \mathbb{N}$), som for overskuelighedens skyld antages alle at have samme masse ($= 1$). Lad os også antage, at halvdelen af partiklerne har ladning $+1$ og den anden halvdel ladning -1. Hvis vi numererer partiklerne $i = 1, 2, \ldots, 2N$ lader vi de første N partikler være positivt ladet, dvs. de har ladningerne $q_i = 1$ for $i = 1, \ldots, N$ og derfor $q_i = -1$ for $i = N+1, \ldots, 2N$. Positionerne kalder vi $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{2N} \in \mathbb{R}^3$. Bølgefuntionen for dette system er en C_0^1 funktion af alle disse variable altså $\psi \in C_0^1(\mathbb{R}^{6N})$. Energiudtrykket er givet ved

$$\begin{aligned}\mathcal{E}_{2N}(\psi) &= \frac{1}{2}\sum_{i=1}^{2N}\int_{\mathbb{R}^{6N}} |\nabla_{\boldsymbol{x}_i}\psi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{2N})|^2 \mathrm{d}\boldsymbol{x}_1 \cdots \mathrm{d}\boldsymbol{x}_{2N} \\ &\quad + \int_{R^{6N}} \sum_{1\le i<j\le 2N} \frac{q_i q_j}{|\boldsymbol{x}_i - \boldsymbol{x}_j|} |\psi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{2N})|^2 \mathrm{d}\boldsymbol{x}_1 \cdots \mathrm{d}\boldsymbol{x}_{2N}.\end{aligned}$$

Systemets laveste energi (grundtilstandsenergien) er

$$E(2N) = \inf\left\{ \mathcal{E}_{2N}(\psi) \;\middle|\; \psi \in C_0^1(\mathbb{R}^{6N}), \int_{\mathbb{R}^{6N}} |\psi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{2N})|^2 \mathrm{d}\boldsymbol{x}_1 \cdots \mathrm{d}\boldsymbol{x}_{2N} = 1 \right\}.$$

Resultatet (2) om brintatomets energi betydergiver, at $E(2) = -\frac{1}{4}$. Ved brug af resultatet om brints stabilitet viser man nemt, at der for alle $N \in \mathbb{N}$ gælder $E(2N) > -\infty$.

For at forstå makroskopiske objekter er vi interesserede i, hvad der sker med $E(2N)$ når $N \to \infty$. Jeg har for ganske nylig (se [5, 10]) bevist, at grænseværdien

$$\lim_{N\to\infty} \frac{E(2N)}{(2N)^{7/5}}$$

eksisterer og er strengt negativ. Faktisk finder jeg et udtryk for denne grænseværdi.

Sætning 6 (Dysons formel)**.**

$$\lim_{N\to\infty} \frac{E(2N)}{(2N)^{7/5}} = \inf\left\{ \mathcal{E}(u) \mid u \in C_0^1(\mathbb{R}^3),\ \int_{\mathbb{R}^3} |u(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} = 1 \right\}, \tag{4}$$

hvor

$$\mathcal{E}(u) = \frac{1}{2}\int_{\mathbb{R}^3} |\nabla_{\boldsymbol{x}} u(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} - \frac{4^{5/4}\Gamma(\frac{3}{4})}{5\pi^{1/4}\Gamma(\frac{5}{4})} \int |u(\boldsymbol{x})|^{5/2} \mathrm{d}\boldsymbol{x}.$$

At højresiden af (4) ikke er $-\infty$ følger af Sobolevuligheden (3) og Hölders ulighed (som omtales i de fleste bøger om mål og integralteori).

Det kunne umiddelbart se ud som om højresiden af (4) er lige så svær at udregne som $E(2N)$ selv. Det er langt fra tilfældet. Det er nemt at finde højresiden af (4) numerisk på en computer. Antallet N skal ikke være særlig stort (mindre end 10), før selv den hurtigste computer, man teoretisk kan forestille sig, ikke vil kunne udregne $E(2N)$ indenfor en rimelig tid.

Inden jeg tager hele æren for at have bevist (4), skal jeg understrege, at formlen blev gættet af den berømte matematiske fysiker F. Dyson i 1967 [1] og at den ene halvdel af beviset er et fælles arbejde med E. Lieb fra Princeton University. Det samlede bevis er på over 70 sider, så jeg kan ikke komme ind på det her.

Formlen (4) viser, at der må være noget galt med vores beskrivelse af systemet af ladede partikler. Potensen 7/5 i formlen har nemlig katastrofale følger for eksistensen af en makroskopisk verden. Da $7/5 > 1$ vil en konsekvens af formlen være, at alt stof vil have en voldsom tiltrækningskraft og alt vil kollapse. Bemærk nemlig at to makroskopiske objekter med henholdsvis $2N_1$ og $2N_2$ partikler adskilt vil have en energi, der (hvis jeg ser bort fra en positiv multiplikativ konstant) er ca. $-N_1^{7/5} - N_2^{7/5}$. Hvis vi på den anden side lader dem kollapse til ét objekt, bliver energien $-(N_1+N_2)^{7/5} \ll -N_1^{7/5} - N_2^{7/5}$. De to objekter kan derfor kollapse og udløse en enorm energimængde. Man kan ved en nærmere analyse se, at et makroskopisk objekt med 10^{23} partikler vil få en udstrækning mindre end en atomkerne altså langt mindre end et atom bestående af blot to partikler. Objekter bestående af mange partikler vil altså være langt mindre end objekter med få partikler. Med andre ord formlen (4) betyder, at en makroskpisk verden *ikke* kan eksistere.

Problemet er, at vi har set bort fra, at de negativt ladede elektroner opfylder det meget vigtige *eksklusions-* eller *Pauliprincip*. Dette princip, som blev formuleret af Wolfgang Pauli og for hvilket han fik nobelprisen i fysik i 1945, siger løst sagt, at to elektroner ikke kan være i samme tilstand. Den mere matematiske formulering af Pauliprincippet er, at bølgefunktionen ψ skal være antisymmetrisk i de variable der svarer til elektroner. Dvs. vi skal antage, at ψ opfylder

$$\begin{aligned}&\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{x}_{N+1},\boldsymbol{x}_{N+2},\ldots,\boldsymbol{x}_{N+j-1},\boldsymbol{x}_{N+j},\boldsymbol{x}_{N+j+1},\boldsymbol{x}_{2N}) = \\ &-\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{x}_{N+j},\boldsymbol{x}_{N+2},\ldots,\boldsymbol{x}_{N+j-1},\boldsymbol{x}_{N+1},\boldsymbol{x}_{N+j+1},\boldsymbol{x}_{2N}),\end{aligned}$$

for alle $j = 1,\ldots,N$. Lad os kalde underummet af $C_0^1(\mathbb{R}^{6N})$ bestående af funktioner med denne egenskab for $C_{\mathrm{F}}^1(\mathbb{R}^{6N})$. Den fysisk korrekte energi er derfor

$$E_F(2N) = \inf\left\{\mathcal{E}_{2N}(\psi) \;\middle|\; \psi \in C_{\mathrm{F}}^1(\mathbb{R}^{6N}), \int_{\mathbb{R}^{6N}} |\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{2N})|^2 \mathrm{d}\boldsymbol{x}_1\cdots\mathrm{d}\boldsymbol{x}_{2N} = 1\right\}.$$

For denne energi kan man vise følgende meget fundamentale resultat.

Sætning 7 (Eksistensen af den makroskopiske grænse)**.** *Grænseværdien*

$$\lim_{N\to\infty} \frac{E_F(2N)}{2N} \tag{5}$$

eksisterer.

Bemærk at hvor vi tidligere havde potensen 7/5 optræder nu potensen 1. Grænseværdien udtrykker energien per partikel for makroskopiske objekter. Man kan fortolke eksistensen af denne grænse som eksistensen af en makroskopisk verden. Grænsen omtales i litteraturen normalt som *den termodynamiske grænse* (termodynamik er beskrivelsen af makroskopiske objekter). Vi har her kun omtalt en meget simpel variant af den termodynamiske grænse (se [3]). Værdien af grænsen i (5) kendes ikke.

Eksistensen af (5) er en nem konsekvens af følgende to egenskaber ved $E_F(2N)$.

Lemma 8 (Subadditivitet af energien)**.** *Energien E_F er subadditiv, d.v.s. for alle $N_1, N_2 \in \mathbb{N}$ er*

$$E_F(2(N_1+N_2)) \leq E_F(2N_1) + E_F(2N_2). \tag{6}$$

(Det samme gælder i øvrigt også for energien $E(2N)$ defineret uden Pauliprincippet.)

Sætning 9 (Stabilitet of makroskopisk stof)**.** *Der eksisterer en positiv konstant C så*

$$-CN \leq E_F(2N) \leq 0, \tag{7}$$

for alle $N \in \mathbb{N}$.

Opgave 10. Vis at Sætning 7 følger af Lemma 8 og Sætning 9.

Opgave 11 (Udfordringen). I den form Lemma 8 er formuleret her, tror jeg ikke, den findes i litteraturen. Der findes tilsvarende og endda vanskeligere resultater. Læseren udfordres til at finde et bevis for Lemma 8. Den, der tager udfordringen op, kan overveje at udbygge det til et bachelor- eller fagprojekt i emnet.

Sætning 9 om stabilitet af makroskopisk stof er meget svær at vise og har en lang historie, som jeg kort vil beskrive nedenfor. Det første bevis blev givet af Dyson og Lenard i [2].

Mere om stabilitetsproblemets historie

Vi har set, at forklaringen af atomers stabilitet var en af kvantemekanikkens tidlige triumfer. Den numeriske værdi for E_H givet i Sætning 2 passer i øvrigt fantastisk godt med eksperimenter. Vi har også set, at stabilitet af atomer kan forstås ud fra passende formuleringer af usikkerhedsrelationen, dog ikke ud fra den sædvanlige formulering givet af Heisenberg.

Det overraskende er, at spørgsmålet om stabilitet af makroskopiske systemer er næsten helt overset i fysiklitteraturen. Den første gang, der bliver gjort opmærksom på problematikken, er i en artikel [8] i 1939 af den norske kemiker Lars Onsager (som for et helt andet arbejde fik nobelprisen i kemi i 1968). Onsager[5] giver i sin artikel et stort set matematisk korrekt argument for stabilitet under visse simplificerende antagelser. Onsagers hovedargument har ligget til grund for mange senere beviser for stabilitet. (Den første artikel jeg selv skrev [9] var en kort bemærkning om Onsagers argument.)

Resultatet i Sætning 6 viser, at stabiliteten af makroskopiske objekter ikke kan forklares alene ud fra usikkerhedsrelationen alene. Man er nødt til også at tage Pauliprincippet i betragtning. Denne vitale konsekvens af Pauliprincippet bliver stort set aldrig fremhævet i fysiklitteraturen, hvor der ellers ofte lægges stor vægt på princippets betydning for stabilitet af stjerner, f.eks. de såkaldte hvide dværge. Dyson [1] er den første, der indser betydningen af Pauliprincippet for eksistensen af en makroskopisk verden.

Som allerede nævnt er det også Dyson, der i samarbejde med A. Lenard først viser Sætning 9 om stabiliteten af makroskopisk stof. E. Lieb og J. Lebowitz [3] er de første til at vise eksistensen af den termodynamiske grænse i en betydelig mere generel situation end her.

Som nævnt i indledningen har vi set bort fra magnetiske vekselvirkninger og tyngdekraften. Jeg har selv været med til at vise [4], at man kan inkludere de magnetiske vekselvirkninger i stabilitetsbetragtningerne (se også næste afsnit om uløste problemer). Ved betragtning af magnetiske kræfter kommer lysets hastighed c naturligt ind som en parameter. I de enheder vi har valgt, får lysets hastighed en værdi, der ofte benævnes α^{-1}, hvor α kaldes *finstrukturkonstanten*. Finstrukturkonstantens værdi er $\alpha \approx 1/137$. Fordi α er relativt lille, har de magnetiske kræfter kun lille indflydelse på f.eks. atomare spektre. De ændringer i spektrene, der skyldes de magnetiske kræfter, kaldes for finstrukturen. Det viser sig (se [4]), at man for at få makroskopisk stabilitet må kræve, at finstrukturkonstanten er lille nok. Heldigvis er $1/137$ lille nok!

Tyngdekraften er betydelig mere kompliceret. For denne vekselvirkning gælder makroskopisk stabilitet ikke, men systemer skal være ganske store, før de bliver ustabile. Den indiske fysiker S. Chandrasekhar udviklede teorien for gravitationel stabilitet og instabilitet og fik for det nobelprisen i fysik i 1983. En matematisk behandling af Chandrasekhars teori kan findes i [6, 7]. Når gravitationel instabilitet indtræffer, kan der formodentlig opstå sorte huller. Det er også muligt, at hele universet en gang med tiden vil kollapse. Der findes ikke i dag nogen fysisk teori, der kan forklare de forhold, der opstår under gravitationelt kollapse. Det vil kræve en teori der omfatter både kvantemekanik og generel relativitetsteori.

[5]Selvom Lars Onsager var kemiker, har han haft stor indflydelse på moderne matematik f.eks. indenfor kvantegrupper og topologiske feltteorier.

Det store problem: Stabilitet i kvantefeltteori

Jeg har diskuteret stabilitetsproblemer, der involverer elektromagnetiske kræfter og tyngdekraften. I hele diskussionen indtil videre har jeg stiltiende antaget, at disse kræfter kunne beskrives ved, hvad man vil kalde klassiske felter beskrevet ved funktioner eller vektorfelter på det tre-dimensionale rum. Fysikken fortæller os, at dette ikke er en korrekt beskrivelse. Vekselvirkninger skal beskrives kvantemekanisk. Det vil sige, at felterne skal beskrives ved operatorer på Hilbertrum. Sådanne teorier, hvor felterne er beskrevet kvantemekanisk, kaldes *kvantefeltteorier*.

Et eksempel på en kvantefeltteori er teorien for ladede partiklers vekselvirning med det kvantiserede elektromagnetise felt. Denne teori for lys og stof kaldes *kvanteelektrodynamikken*. Feynman, Schwinger og Tomonaga fik i 1965 nobelprisen for (relativt uafhængigt af hindanden) at have formuleret kvanteelektrodynamikken. Man kunne håbe, at det var muligt at formulere og bevise stabilitet i kvanteelektrodynamik. Det er desværre langt fra tilfældet. Kvantelektrodynamik er en besynderlig teori, hvor fysiske størrelser kun kan udtrykkes som potensrækker i finstrukturkonstanten. D.v.s. finstrukturkonstanten er den variable i potensrækken. Selve opskrivningen af disse potensrækker er ganske vanskelig og man har i dag højst udregnet de første 5 eller 6 led i rækkerne. Dyson var den første, der beskrev, hvordan man systematisk kan finde disse rækker. Dyson er vist også den første, der indser, at disse rækker formodentlig har konvergensradius 0 (jeg ved ikke om der findes et fuldstændigt bevis for dette).

I en teori, der kun er beskrevet ved potensrækker med konvergensradius 0, giver det ingen mening at undersøge stabilitet. Faktisk er der meget lidt, der giver matematisk mening i en sådan teori. Det skal retfærdigvis nævnes, at man, ved at udregne de første få led i kvantelektrodynamikkens potensrækker, får en helt forrygende overenstemmelse med eksperimetelle værdier (ofte overenstemmelse udover 10 betydende cifre).

En måde, at forsøge at studere stabilitet i kvanteelektrodynamik, er ved at indføre en konvergensfaktor. Det er en helt almindelig teknik i matematik. Man indfører en lille ekstra parameter $\varepsilon > 0$, som sikrer at alt konvergerer. Der foregår for tiden en del forskning af stabilitet for kvanteelektrodynamik med en konvergensfaktor inkluderet. Man ville selvfølgelig gerne kunne lade ε gå mod nul, men det er ikke muligt.

Et håb er, at problemerne med kvantelektrodynamik skyldes, at man ikke har den fulde teori. Et langsigtet håb er, at man ved at inkludere de kræfter, vi indtil videre har ignoreret, nemlig den svage og den stærke vekselvirkning, vil få en teori, der ikke har de samme problemer som kvanteelektrodynamik. Man er selvfølgelig mest interesseret i en teori, der inkluderer alle fire kendte kræfter: elektromagnetisme, tyngdekraften og den svage og stærke vekselvirkning. Det har

nok meget lange udsigter.

Litteraturliste

[1] Dyson, Freeman J., Ground state energy of a finite system of charged particles, *Jour. Math. Phys.* **8**, 1538–1545 (1967).

[2] Dyson, Freeman J. and Lenard, Andrew, Stability of matter. I , *Jour. Math. Phys.* **8**, 423–434, (1967) and II *Jour. Math. Phys.* **9**, 698–711 (1968).

[3] Lieb, Elliott H. and Lebowitz, Joel L. , The constitution of matter: Existence of thermodynamics for systems composed of electrons and nuclei. Advances in Math. **9**, 316–398 (1972).

[4] Lieb, Elliott H. and Loss, Michael and Solovej, Jan Philip, Stability of matter in magnetic fields, *Phys. Rev. Lett.* **75**, 985–989 (1995). 22–116, (1977).

[5] Lieb, Elliott H. and Solovej, Jan Philip, Ground state energy of the two-component charged Bose gas, *Commun. Math. Phys.* **252**, 485 – 534, (2004).

[6] Lieb, Elliott H. and Thirring, Walter E. , Gravitational Collapse in Quantum Mechanics with Relativistic Kinetic Energy, *Annals of Phys. (N.Y.)* **155**, 494–512 (1984).

[7] Lieb, Elliott H. and Yau, Horng-Tzer, The Chandrasekhar Theory of Stellar Collapse as the Limit of Quantum Mechanics, *Commun. Math. Phys.* **112**, 147–174 (1987).

[8] Onsager, Lars, Electrostatic Interaction of Molecules, *Jour. Phys. Chem.* **43**, 189–196 (1939).

[9] Solovej, Jan Philip, Some simple counter examples to the Onsager lemma on metric space, *Reports on Math. Phys.*, 28, 1989.

[10] Solovej, Jan Philip, Upper Bounds to the Ground State Energies of the One- and Two-Component Charged Bose Gases. To appear in *Comunications in Mathematical Physics* 2006.

FAMØS april 2006.
Fagblad for Aktuar-, Matematik-,
Økonomi- og Statistikstuderende ved
Københavns Universitet.

Redaktionsgruppe:
Mikkel Abrahamsen
Nikolaj Strands
Sara Arklint (ansvh.)
Tarje Bargheer
Taus Brock-Nannestad
Ulrik Buchholtz

Tegnere:
Anne Vinkel
Martin Damhus aka Damskur

Deadline for næste nummer:
Tirsdag den 6. juni 2006

Indlæg modtages gerne og bedes sendt
til `famos@math.ku.dk` – meget gerne
skrevet i LaTeX.

FAMØS er et internt fagblad.

Eftertryk tilladt med kildeangivelse.

Fagbladet FAMØS
c/o Institut for matematiske fag
Matematisk Afdeling
Universitetsparken 5
2100 København Ø
`http://www.math.ku.dk/famos/`

Oplag: 500 stk.
ISSN 1395-2145

Glædeligt forår!

Kig ud ad vinduet. Solen er brudt frem og foråret langt om længe kommet til København. Der er mildhed og varme i luften, samtlige persienner i E-bygningen er straks blevet trukket ned, og den sidste martssne er smeltet.

Knopper sidder i hobetal på alle træer og buske, og i E-bygningens kælder bugner det med specialestuderende som snart springer ud som kandidater.

Alle disse specialestuderende har brug for formidlingsaktiviteter for at blive færdige, så vi har på FAMØS modtaget i alt fire artikler denne gang. Det er vi selvfølgelig glade for, men fire artikler er alt for meget. Vi har derfor udskudt de to af artiklerne til juninummeret, og hvis vi modtager flere artikler bliver de tidligst bragt til oktober.

Det er selvfølgelig lidt hårde vilkår. Især når formidlingsaktiviteter er en obligatorisk del af kandidatuddannelsen. Men hvis vi på den anden side bragte alle fire artikler på en gang, ville artiklerne formodentlig slet ikke blive læst.

Så der er ikke tale om at vi går rundt med en flammekaster og forkuller alle de nye skud. Der er derimod tale om at vi luer ud og prikler, så der bliver ordentlig plads til alle de nye spæde planter.

Jeg er vist ved selv at blive kvalt i dette vildskud af metaforer. Og min leder er vist allerede afgået ved kvælningdøden. Så jeg må hellere stoppe her.

På vegne af redaktionen vil jeg blot ønske alle veloverståede eksaminer og en kåd påske.

Vi ses til revyen.