Workshop om fejlfindende og -rettende koder

Kjeld Bagger Laursen

October 11, 2005

1 Indledning[1]

Kig på bagsiden af en hvilkensomhelst bog udgivet indenfor de seneste år. Et eller andet sted - ofte i nederste højre hjørne - finder du det såkaldte ISBN, f.eks.

`ISBN 87-502-0440-8`,

tit sammen med en eller flere stregkoder. Hver eneste bog der udgives har sit eget International Standard Book Number, og dette tal spiller en vigtig rolle i hele markedsføringsprocessen: Ved lagerkontrol, ved bestilling, forsendelse, prissætning og betaling.

ISBN-systemet er et eksempel på dagligdags anvendelse af *fejlfindende* og *fejlrettende* koder. Fejlfindende og fejlrettende 'koder' kender vi udmærket fra det almindelige sprog. De virker ved hjælp af såkaldt *redundans*, dvs. ekstra, 'overflødig' information, som tjener som korrektions-mulighed. Masser af trykfejl i avisen giver anledning til et billigt grin, men ikke misforståelser. Hvis vejrudsigten lover 'rogn i morgen' er der tilstrækkelig redundans i situationen (vi læser en vejrudsigt, ikke en madopskrift) til at vi forstår hvad den korrekte udsigt er. Men hvis der står at en vis dansk kvindes fornavn er Annå, kan vi se at der er noget galt (fejlfinding), men vi kan ikke umiddelbart fejlrette. Hedder hun Anne, Anna, Anni?

Det drejer sig altså om pålidelighed af informationsoverførsler, og systemet fungerer principielt sådan her:

- Informationen foreligger i standardiseret numerisk form, f.eks. som et ISBN. Vi kan kalde denne række af cifre et *kodeord.*
- Kodeordet indeholder tillige et eller flere tjekcifre.

Tjekcifret er regnet ud efter ganske bestemte regler og sætter modtageren af kodeordet i stand til at kontrollere om der er fejl i kodeordet; vi siger at vi har en *fejlfindende* kode, - og det er det, vi hovedsagelig skal beskæftige os med her. Visse systemer kan (altid eller sommetider) også konstatere hvad det korrekte kodeord er; i så fald har vi en *fejlrettende* kode - endnu et fascinerende emne (og nyttigt: Det nyder du f.eks. godt af hver gang du spiller en CD!), men det må hvile til en anden god gang.

Hvilke informationer indeholder et ISBN? Som eksemplet viser, er et ISBN et 10-cifret tal $x_1x_2...x_{10}$, som éntydigt bestemmer bogen. Cifrene $x_1..x_9$ kan være hele tal 0,1,...,9. Et ISBN indeholder, fra venstre mod højre, oplysninger om udgivelseslandet, forlaget og bogens 'eget' nummer. F.eks. har Danmark cifrene 87, mens de store engelsksprogede lande USA og England deles om tallet 0. De næste to eller tre cifre angiver forlaget: 502 er Polyteknisk Forlag. Så er der 4-7 cifre til at angive den enkelte bog. Endelig er der det sidste ciffer, x_{10}, som er et *tjek-ciffer.*

Et tjekciffer bruges til fejlfinding. I et ISBN vælges x_{10} sådan at

$$x_1 + 2x_2 + 3x_3 + ... + 9x_9 + 10x_{10}$$

[1] Dette dokument ligger også (i pdf-format) på adressen http://www.math.ku.dk/~laursen/ (nederst på siden)

er deleligt med 11. Vi siger at denne vægtede tværsum skal være 0 *modulo* 11, netop fordi den divideret med 11 giver resten 0, og skriver

$$x_1 + 2x_2 + 3x_3 + ... + 9x_9 + 10x_{10} = 0 \pmod{11}.$$

Tjekcifret x_{10} kan være 0,1,...,9, men kan også være 10. For at fastholde cifferantallet, bruger man romertallet X til at betegne tjekcifferværdien 10.

Når man har valgt at regne modulo 11, er det dels fordi 11 er større end de cifre der kan forekomme i et ISBN, og dels fordi 11 er et primtal. Hermed er vi inde på *talteori.* Talteori er en klassisk matematisk disciplin, og den har mest været dyrket for sin egen skyld. De moderne anvendelser af den, vi beskriver her, illustrerer et ofte set fænomen i matematik, nemlig at det man oprindelig anså for den reneste matematik viser sig at have stor nytteværdi senere hen. Vi skal ikke bruge så meget talteori her, men kan dog om lidt konstatere at vi er i stand til at opdage *alle* fejl i et enkelt ciffer og *alle* fejl bestående af ombytning af to cifre. Derved adskiller ISBN sig fra et andet nummersystem, EAN (som vi kommer til om lidt), hvor man har valgt at regne modulo 10; en af konsekvenserne af dette valg (af et ikke-primtal) er at ikke alle ombytningsfejl automatisk opdages.

2 Primtal

Nogle hele tal kan kun deles med sig selv og med 1, mens de øvrige er delelige med andre hele tal. De 'udelelige' tal kalder vi som bekendt *primtal* - de øvrige *sammensatte* tal. Listen over primtal starter sådan: 2,3,5,7,11,13,17,23,....(tallet 1 er specielt, det opfattes som en *enhed*, ikke et primtal) Hvis et positivt helt tal n ikke er et primtal, har det altså andre faktorer end de to trivielle, 1 og n, dvs. der findes et helt tal a, $1 < a < n$, der går op i n.

En berømt talteoretisk sætning, *primtalssætningen*, siger at ethvert helt tal kan skrives som et produkt af primtal. F.eks er

$$1528521 = 3 \cdot 17^2 \cdot 41 \cdot 43$$

Sætningen er næsten selvindlysende: Hvis et helt tal (som vi gerne må antage positivt) foreligger, er det enten et primtal eller et sammensat tal. Hvis det er et primtal, er vi straks færdige med at faktorisere det. Og hvis det er et sammensat tal, kan det jo skrives som et produkt af to mindre tal. Nu kan ræsonnementet gentages på hver af faktorerne. Når vi ramler ind i en primfaktor, går vi blot videre til næste faktor: hvis den er et sammensat tal, kan vi gentage spøgen. De efterhånden opdukkende faktorer er altså enten primtal, eller der forekommer mindre og mindre sammensatte tal. Men vi kan jo ikke komme længere ned end til det mindste naturlige tal 1, så vi må til sidst være nået frem til lutter primtal som faktorer.

Øvelse Er 15163 et primtal?

Men primtalssætningen siger mere: Den siger også at *der er kun én måde* at faktorisere på, altså at de indgående primfaktorer er *entydig bestemt.* Vi kan selvfølgelig ændre på faktorernes rækkefølge, men et ganske bestemt sæt af primtalsfaktorer *skal* forekomme (og vi ser bort fra forekomsten af den inderligt ligegyldige faktor 1). For 1528521 er der altså ingen vej udenom: Faktorerne 3, 41 og 43 indgår nødvendigvis, og faktoren 17 to gange, også nødvendigvis. Det er denne entydighed, der viser sig at være nyttig her.

Bemærk iøvrigt, at der selvfølgelig ikke er mulighed for entydig faktorisering, kun entydig *primtals*faktorisering. For 12 har vi jo $12 = 3{\cdot}4 = 2{\cdot}6$; men primtalsfaktoriseringen er $12 = 2 \cdot 2 \cdot 3$, og den er entydig.

3 ISBN

Tilbage til ISBN: Når vi skal udregne tjekcifret x_{10} i et ISBN, regner vi altså en vægtet tværsum ud

$$x_1 + 2x_2 + 3x_3 + ... + 9x_9 + 10x_{10} = 0 \pmod{11},$$

hvor det i^{te} ciffer får vægten i. Jeg påstod før om ISBN-tjekcifret at vi med det kan konstatere *alle* fejl i et enkelt ciffer. For at indse det, kan vi gøre følgende: Lad os om ISBN $x_1x_2...x_{10}$ forestille os at et af cifrene, det er her ligegyldigt hvilket, så kald det bare x_k, er blevet fejllæst som y_k. Der gælder altså at $y_k = x_k + e$, hvor $e \neq 0$ er fejlen. I den vægtede tværsum bidrager e da med ke, og da det sande ISBNs vægtede tværsum er et multiplum af 11, er det falske ISBNs tværsum et tal af formen (et multiplum af 11)+ke. Fejlen forbliver uopdaget hvis ke er et multiplum af 11. Men husk: k og e er tal mellem 1 og 10, og da 11 er et primtal, er det *ikke* muligt for ke at være et multiplum af 11 (det siger entydighedsdelen af primtalssætningen - se ovenfor). Altså opdages fejlen.

Øvelse En bestillingsliste indeholder ISBN 0-201-82753-4. Er dette ISBN korrekt?

Hvis et ISBN er blevet beskadiget, så et ciffer ikke kan læses, kan man rekonstruere det manglende ciffer.

Kig f.eks. på følgende tal: 0-19-859665-0. Vi har sat streger i det, men kun for at øge læsbarheden. Det er et ISBN - hvilket vi kan tjekke. Lad os nu forestille os at der er kradset i tallet, så femte ciffer ikke kan læses: Vi ser altså 0-19-8x9665-0. Kan vi rekonstruere tallet, altså finde x? Regner vi den vægtede tværsum ud, får vi

$$1 \cdot 0 + 2 \cdot 1 + 3 \cdot 9 + 4 \cdot 8 + 5 \cdot x + 6 \cdot 9 + 7 \cdot 6 + 8 \cdot 6 + 9 \cdot 5 + 10 \cdot 0$$

$$= 250 + 5x.$$

Dette tal må være et multiplum af 11. Da $250 = 22 \cdot 11 + 8$, ser vi at vi får samme rest, som $8 + 5x$ giver, ved division med 11. Altså er $250 + 5x = 8 + 5x \pmod{11}$. Vi skal altså finde et multiplum af 5, der ved addition af 8 giver et multiplum af 11. Klart nok er $x = 5$ en løsning, idet $5 \cdot 5 + 8 = 33$. Intet andet tal mellem 0 og 9 løser denne ligning.

Øvelse En bestillingsliste er blevet tygget igennem af en stor hund, og indeholder derfor det ufuldstændige ISBN 0-669-3x787-5. Find ud af hvad der skulle have stået.

Forklaringen på at vi i eksemplet får præcis ét svar, altså konstaterer at det rekonstruerede ciffer er entydigt bestemt, får vi af primtalssætningen: Hvis k betegner vægten (= plads nummeret), hvis x_1 og x_2 er to kandidater til ciffer nr. k i vores ISBN, og hvis vi bruger betegnelsen r for alle de *øvrige* cifres samlede bidrag til den vægtede tværsum, skal der jo gælde at både $r + kx_1$ og $r + kx_2$ er multipla af 11, og derfor at $(r + kx_1) - (r + kx_2)$ er et multiplum af 11. Symbolsk: $kx_1 - kx_2 = 0 \pmod{11}$. Erstatter vi $x_1 - x_2$ med x, står vi altså med ligningen $kx = 0 \pmod{11}$, eller i ord: kx er et multiplum af 11. Men hvis $kx = 11 \cdot c$ for et eller andet helt tal c, må kx jo være et tal hvori 11 er en primfaktor. Da k er et helt tal mellem 1 og 10, og x ligger mellem 0 og 10, er den eneste mulighed at $x = 0$. Altså er $x_1 = x_2$.

Øvelse I den foregående øvelse kunne ISBN rekonstrueres. Kan man det i øvelsen før? Hvorfor ikke?

En anden fejlmulighed ligger i ombytning af to cifre. Denne fejl opstår typisk ved fejl-kopiering af et ISBN. I et ISBN vil en ombytningsfejl blive fanget af tværsumstjekket. Det kan man indse på f.eks. følgende måde; læg mærke til at vi ikke indskrænker os til at betragte fejlagtig ombytning af *nabo*cifre.

Vi forestiller os altså at det 10-cifrede tal $x_1x_2...x_k...x_j...x_{10}$ er et ISBN, som er blevet fejlskrevet som $x_1x_2...x_j...x_k...x_{10}$. Vi regner den vægtede tværsum af dette tal ud:

$$x_1 + 2x_2 + ... + kx_j + ... + jx_k + ... + 10x_{10}$$

og trækker den fra det rigtige ISBNs tværsum, som jo er $x_1 + 2x_2 + ... + kx_k + ... + jx_j + ... + 10x_{10}$. Resultatet af denne subtraktion er åbenbart

$$kx_k + jx_j - (kx_j + jx_k)$$

(alle andre led går jo ud med hinanden), så den numeriske afvigelse mellem de to tværsummer er

$$(j-k)\,|x_j - x_k|\,.$$

Fejlen opdages netop hvis dette tal *ikke* er et multiplum af 11. Men både $j-k$ og $|x_j - x_k|$ er tal mellem 1 og 9, og vi ser igen - fordi 11 er et primtal (altså fordi vi har primtalssætningen!) - at $(j-k)\,|x_j - x_k|$ *ikke* er et multiplum af 11. Altså opdages ombytningsfejlen.

4 Galois

I det foregående har vi jo regnet modulo 11, og betragter man mængden af alle hele tal modulo 11, har man det matematiske objekt der hedder et *endeligt legeme*. Dets traditionelle betegnelse er $\mathbf{Z}_{11}$. Disse endelige legemer kaldes også *Galois legemer*, efter den franske matematiker Evariste Galois. Galois døde allerede som 20 årig, i en duel i 1832 (duellen havde dog ikke noget med hans matematik at gøre). Han gjorde adskillige banebrydende matematiske opdagelser og er, trods sit meget korte liv, en af den abstrakte algebras absolut mest markante skikkelser.

Ordet legeme ser måske besynderligt ud, i denne sammenhæng. Det er kommet ind i dansk som en oversættelse af det tyske Körper. Det lyder ikke umiddelbart meget bedre, som navnet på et matematisk begreb, men forklaringen skal nok søges i dets latinske rod *corpus*, dvs. det tilsvarende danske korps (og franske corps), i betydningen 'en organisation med en velspecificeret og sammentømret struktur'. Den tyske matematiker Richard Dedekind (1831-1916) var vist den første der brugte ordet Körper; på norsk og svensk blev det iøvrigt til krop!

5 EAN

EAN står for European Article Number. Sådan et tal står i den stregkode, som alle supermarkedvarer er mærket med. En vares stregkode læses ved kassen af en scanner, der aflæser stregkoden ved hjælp af en laser-lysstråle. Denne stråle aflæser *forholdene* mellem de hvide og de mørke stregers *bredder* (det er derfor en scanner kan læse en stregkode under 'skæve' vinkler). Kasseapparatets computer tjekker det læste nummer, bl.a. ved at slå det op i sin database, og styrer kassebonsudprintningen af varenavn og -pris. Hvordan stregkoden tjekkes, i princippet, skal jeg kort beskrive nu.

Et EAN er et 13-cifret tal, og hver af cifrene kan være 0, 1,..., 9. Ligesom med ISBN indeholder et EAN oplysninger om varens oprindelsesland, producenten, varenummeret. Det sidste ciffer x_{13} er et tjekciffer. Det regnes ud fra kravet at

$$x_1 + x_3 + x_5 + x_7 + x_9 + x_{11} + x_{13} + 3(x_2 + x_4 + x_6 + x_8 + x_{10} + x_{12}) = 0 \pmod{10},$$

altså at den angivne vægtede tværsum skal være et multiplum af 10. Vi noterer os straks at dette tjek vil opdage *alle* fejl i et enkelt ciffer: Hvis et enkelt ciffer er forkert, vil denne afvigelse fra det korrekte blive ganget med 1 eller med 3, når tværsummen regnes ud, og resultatet kan ikke være et multiplum af 10 (kan du se det? - læg mærke til at hvis man i stedet for vægtene 1 og 3 havde brugt vægte der har faktorer fælles med 10, f.eks. 1 og 2, så ville visse fejl forblive uopdaget. Brugte man vægten 2, i stedet for 3, på cifrene med lige index, ville en fejl på f.eks. 5 jo blive ganget med 2, og resultatet ville stadig være et multiplum af 10 - altså tilsyneladende et korrekt EAN.)

Øvelse Scanneren vil ikke læse tandpastatubens EAN, så kassedamen indtaster 5-720000-052363. Hvad tilbagemelding giver systemet: Accepterer det indtastningen?

I modsætning til situationen for ISBN kan ombytning af visse nabocifre slippe igennem. For at give et konkret eksempel kan vi se på de to første cifre i et EAN x_1 og x_2. Vi kan

selvfølgelig antage at $x_1 \neq x_2$, ellers *er* der jo ingen synlig fejl. Disse to cifre bidrager til den vægtede tværsum med beløbet $x_1 + 3x_2$. Hvis de er blevet byttet om, giver de i stedet bidraget $x_2 + 3x_1$, så den vægtede tværsum er altså blevet ændret med $2x_2 - 2x_1 = 2(x_1 - x_2)$. Det betyder jo at hvis x_1 og x_2 afviger 5 fra hinanden, vil den vægtede tværsum være et multiplum af 10, både med og uden ombytningen. Denne misere opstår selvfølgelig fordi tallet 10 ikke er et primtal.

Øvelse En vare har EAN 7-311570-658908. Angiv en cifferombytningsfejl, som fanges. Angiv en, der ikke fanges.

Hvorfor bruger man da regning modulo 10 ved EAN? Formodentlig fordi det er væsentligst at benytte regning modulo et tal der er større end de forekommende cifre (og dermed ikke selv kan forekomme som ciffer, eller som fejl). Og så slipper man for den skønhedsplet, som ISBN har, nemlig at tjekcifret 10 kan forekomme, og det nødvendiggør brugen af et ekstra symbol X. Hvordan dette undgås i CPR-numre, der også tjekker ved regning modulo 11, skal vi nu se.

6 CPR

CPR numre kender vi jo alle: Et CPR nummer er et 10-cifret tal, og hver person i Danmark har et. De første 6 cifre angiver personens fødselsdag, og de sidste fire er et løbenummer. I løbenummeret indbygges en del information: Af sidste ciffer kan man se personens køn, af første ciffer (i løbenummeret) om fødselsdagen ligger i det 20. eller det 19. århundrede (hvis det nævnte ciffer er mindre end 5, er personen født i det 20. århundrede, mens et ciffer ≥ 5 angiver en fødselsdato før år 1900). Endvidere er der en tværsumsalgoritme for CPR numre: Cifrene c_1, c_2, ... c_{10} tildeles i rækkefølge vægtene $4, 3, 2, 7, 6, 5, 4, 3, 2, 1$ og tallet $4c_1 + 3c_2 + 2c_3 + 7c_4 + 6c_5 + 5c_6 + 4c_7 + 3c_8 + 2c_9 + 1c_{10}$ skal opfylde at

$$4c_1 + 3c_2 + 2c_3 + 7c_4 + 6c_5 + 5c_6 + 4c_7 + 3c_8 + 2c_9 + 1c_{10} = 0 \pmod{11}.$$

Der regnes altså modulo 11, ligesom ved ISBN, så fejl i et enkelt ciffer fanges altid. Der er selvfølgelig andre tjekmuligheder: De første seks cifre skal jo være en dato. Men bemærk at ikke alle ombytningsfejl afsløres (omend nok de mest sandsynlige gør). Bemærk også at fordi man i løbenummeret har adskillige cifre at 'råde over', er det ikke nødvendigt at benytte et 'ekstra' symbol såsom X. For at opnå et multiplum af 11 i den vægtede tværsum, kan man jo justere på andet end blot det sidste ciffer i CPR nummeret.

Øvelse Kør tværsumstjekket på dit eget CPR nummer. Lav også, med udgangspunkt i dags dato, et CPR-nummer for en nyfødt pige. Lav et lovligt CPR-nummer for en mand, født i dag i 1897.

7 Lidt mere litteratur om emnet:

Stephen Barnett: *Some Modern Applications of Mathematics*, Ellis Horwood, Hemel Hempstead 1995
Jens Carstensen: *Talteori*, Systime, Herning 1993
Joseph A Gallian, Steven Winters: Modular Arithmetic in the Marketplace, American Mathematical Monthly 95 (1988), 548-551.
Her har vi mest beskæftiget os med fejlfindende koder. Fejl*rettende* koder kan du læse mere om, f.eks. i H. Elbrønd Jensen og T. Høholdt: *Fejlkorrigerende koder, en introduktion* (Mat-Pr nr. 8, marts 1994, 27 sider), som kan fås ved henvendelse til Matematisk Institut, DTU. Vi har slet ikke været inde på *hemmelige* koder, der behandles i *kryptologien*. Emnet har ikke bare med spioner og sligt at gøre - i alle situationer hvor fortrolighed og databeskyttelse er væsentlig, f.eks. banktransaktioner (tænk på Dankort), spiller kryptologi en rolle. Du kan læse mere herom i Peter Landrock og Knud Nissen: *Kryptologi*, Abacus, Vejle, 1990.

8 Workshopopgaver

Øvelse Er 15163 et primtal? [2]

Øvelse En bestillingsliste indeholder ISBN 0-201-82753-4. Er dette ISBN korrekt?

Øvelse En bestillingsliste er blevet tygget igennem af en stor hund, og indeholder derfor det ufuldstændige ISBN 0-669-3x787-5. Find ud af hvad der skulle have stået.
Øvelse I den foregående øvelse kunne ISBN rekonstrueres. Kan man det i øvelsen før? Hvorfor ikke?

Øvelse Scanneren vil ikke læse tandpastatubens EAN, så kassedamen indtaster 5-720000-052363. Hvad tilbagemelding giver systemet: Accepterer det indtastningen?

Øvelse En vare har EAN 7-311570-658908. Angiv en cifferombytningsfejl, som fanges. Angiv en, der ikke fanges.

Øvelse Kør tværsumstjekket på dit eget CPR nummer. Lav også, med udgangspunkt i dags dato, et CPR-nummer for en nyfødt pige. Lav et lovligt CPR-nummer for en mand, født i dag i 1897.

9 Appendix

Her er et bevis for primtalssætningen. Det går tilbage til Euklid.[3]

Theorem 1 *Ethvert helt tal kan skrives som et produkt af primtal og de forekommende primfaktorer er entydigt bestemt.*

BEVIS Lad det hele tal m være givet. Vi kan antage at $m > 0$ (fortegnsskifte, hhv. trivialtilfældet 0). Hvis m er et primtal, er vi færdige. Og hvis m er sammensat, kan det faktoriseres. De fundne faktorer er enten alle primtal, ellers gentages argumentet. Hver ny-funden faktor er mindre end det først faktoriserede tal. Før eller siden stopper processen.

Vi mangler at vise at faktoriseringen er entydig. Vi trækker igen på Euklid:

Lemma 2 *Hvis p er et primtal der går op i produktet $a \cdot b$, så går p op i enten a eller b.*

BEVIS (Euklid) Når to hele tal c og d ikke har nogen fælles faktorer, så findes der hele tal e og f for hvilke $ce + df = 1$. For at se det, dividerer vi det mindste tal c op i det største d

$$\begin{aligned} d &= a_1 c + r_1 \\ a_1 &= a_2 r_1 + r_2 \\ r_1 &= a_3 r_2 + r_3 \end{aligned}$$

osv. Læg mærke til at ingen af resterne er 0 (hvis en rest var 0, kunne vi gå tilbage op igennem regnestykkerne og få en fælles faktor). Da resterne bliver mindre og mindre, må vi ende med en linie, der ser sådan ud:

$$r_{s-1} = a_{s+1} r_s + 1$$

I den linie kan vi indsætte de foregående og få vores mellemresultat. □

Så kan vi til gengæld let vise at hvis p går op i $a \cdot b$, men ikke i a, så går p op i b :

[2] Prøv at google vendingen 'table of primes' så får du adgang til mange lister over primtal. Du kan også finde info om de største kendte primtal

[3] Euklid (ca. 300 fvt) levede og virkede i Alexandria. Hans 'Elementer' siges at være den næstmest læste bog gennem tiderne. I dens 13 bøger samlede han 465 sætninger, som ikke alle kan have været hans opdagelse. Men dens aksiomatiske opbygning har øvet enorm indflydelse.

Vi kan jo skrive $as + pt = 1$ og derfor at $abs + pbt = b$. Her står så direkte at læse at p går op i venstresiden, og derfor i højresiden.

Vi kan nu vende tilbage til primtalssætningen:

Vi skal vise entydigheden af primtalsfaktoriseringen. Hvis $p_1 \cdot ... \cdot p_n = q_1 \cdot ...q_m$, og begge sider er ordnet, så $p_i \leq p_{i+1}$, hhv. $q_i \leq q_{i+1}$, kan vi først bortforkorte alle fælles faktorer. Kig så på p_1. Dette tal går op i venstresiden, og derfor i højresiden. Det går ikke op i q_1. Derfor (lemmaet!) går det op i produktet af de øvrige. Forkort så p_1 ud. Gentag spøgen med p_2. Efter tur bliver alle faktorerne på venstresiden fjernet. Vi ender i en modstrid. □

Jeg kan ikke nære mig for at vise jer en flot konsekvens af disse ting:

Theorem 3 *Der er uendelig mange primtal*

BEVIS (Euklid) Antag at påstanden ikke er sand. Så er her en fuldstændig liste over alle primtallene

$$\{2, 3, 5, 7, ..., P\}$$

Lad så

$$M := 2 \cdot 3 \cdot 5 \cdot 7 \cdot ... \cdot P + 1$$

Det er klart at $M > P$, så M er ikke med på listen. Altså er M ikke et primtal. Så er M sammensat og må derfor kunne faktoriseres. Det betyder at M har mindst én primtalsfaktor p, altså p går op i M.Dette p står efter vores antagelse på listen, så det er klart at p går op i $2 \cdot 3 \cdot 5 \cdot 7 \cdot ... \cdot P$. Derfor går p også op i $M - 2 \cdot 3 \cdot 5 \cdot 7 \cdot ... \cdot P$. Men det betyder jo at går p op i 1. Det er absurd! □

Bemærkning Iøvrigt er det ikke tilfældigt at der skal en vis anstrengelse til for at vise at primtalsfaktoriseringen er entydig. For der findes talmængder, hvor det har fin mening at tale om 'primtal', men hvor faktorisering *ikke* er entydig. For at følge med i resten af denne bemærkning, skal du vide at hvad *de komplekse tal* er, og specielt huske at de *komplekse* tal i er karakteriseret ved at $i^2 = -1$.

Betragt mængden af alle tal af formen

$$a + ib\sqrt{3},$$

hvor a og b kan være vilkårlige hele tal. Eksempler er

$$1, i\sqrt{3}, -17, 2 - 2i\sqrt{3}.$$

Det er klart at summen af to sådanne tal igen er af denne form. Og tilsvarende for produkt:

$$(a_1 + ib_1\sqrt{3})(a_2 + ib_2\sqrt{3} = a_1a_2 - 3b_1b_2 + i(a_1b_2 + a_2b_1)\sqrt{3}.$$

Specielt kan man altså begynde at se på faktorisering af elementer indenfor denne klasse. Et simpelt eksempel på faktorisering inden for klassen er

$$4 = 2 \cdot 2,$$

men man kunne jo også skrive at

$$4 = (1 + i\sqrt{3})(1 - i\sqrt{3}).$$

'Prim'tal i denne mængde må rimeligvis være de tal der har 'prim'egenskaben, at der ikke er andre faktorer (i den mængde vi betragter) end tallet selv (og 1).

Og prøv så at afgøre om 2 er et primtal i denne mængde. Ifølge gange-reglen ovenfor skal vi altså afgøre om vi kan finde hele tal a_1, b_1, a_2, b_2, der opfylder at

$$a_1a_2 - 3b_1b_2 + i(a_1b_2 + a_2b_1)\sqrt{3} = 2,$$

eller skrevet ud

$$\begin{aligned} a_1a_2 - 3b_1b_2 &= 2 \\ a_1b_2 + a_2b_1 &= 0 \end{aligned}$$

Jeg overlader sagen til dig! (men kan godt afsløre at der er *ingen* løsninger.)

Prøv så derefter at vise at både $1 + i\sqrt{3}$ og $1 - i\sqrt{3}$ er 'prim'tal.

Med den viden i hus har vi altså fundet et eksempel på en talmængde, der 'svarer til' de hele tal, men *ikke* har entydig primtalsopløsning.

PS I er velkomne til at melde tilbage til mig, spørgsmål, opgaveløsninger, kommentarer til noterne, til seancen osv. Det er lettest at emaile til mig: laursen@math.ku.dk