FERMATS SISTE TEOREM

John Rognes

Desember 1994

1. Den pythagoreiske læresetning

La $\triangle ABC$ være en rettvinklet trekant, med kateter a og b, og hypotenus c. Da er

$$a^2 + b^2 = c^2 \tag{1}$$

i følge den pythagoreiske læresetning, og omvendt kan vi konstruere en rettvinklet trekant ved å danne en trekant med sidekanter a, b og c som oppfyller likningen over.

Murere har lenge (?) brukt denne metoden ved å markere $3+4+5 = 12$ like lange biter på et tau, gjerne ved å knyte tretten knuter i jevn avstand etter hverandre. Slik fremkommer en rett vinkel i en trekant med sidekanter i forholdene $3 : 4 : 5$. Dette virker fordi

$$3^2 + 4^2 = 9 + 16 = 25 = 5^2$$

og er spesielt anvendelig fordi alle sidekantene er gitt ved heltallige multipler av en gitt lengde, som gjør det enkelt å måle opp sidelengdene.

Vi kaller tre naturlige tall a, b, c som oppfyller (1) et pythagoreisk trippel. Triplet $3, 4, 5$ er det minste eksempelet på slike, men vi skal se at det er mange andre muligheter. Det er klart at hvis a, b, c er et pythagoreisk trippel, så er også $2a, 2b, 2c$ et slikt trippel, og mer generelt er ma, mb, mc et pythagoreisk trippel for ethvert naturlig tall m. Vi sier at disse triplene er ekvivalente (opp til skalering), og er mest interessert i å finne inekvivalente tripler.

Pythagoras' setning fra den greske antikkens geometri var kjent mye tidligere i Kina (når ?), men disse triplenes opphav strekker seg ennå lenger tilbake. På babylonske leirtavler fra ca. 1500 før Kristus er det funnet cuneiform–inskripsjoner med lister av pythagoreiske tripler som begynner med $3, 4, 5$ og fortsetter opp til $4961, 6480, 8161$. Det siste eksempelet er for stort til å ha blitt funnet ved prøving og feiling, så babylonerne må ha hatt en metode for å finne pythagoreiske talltripler, men vi vet ikke noe om hvordan de gikk frem.

Grekerne ved Diophantus visste hvorledes man kunne finne alle pythagoreiske talltripler. Vi minner om at to hele tall p og q kalles relativt primiske hvis deres største felles faktor er 1, dvs. det finnes ikke noe primtall som deler både p og q.

Hvis d er den største felles faktor til a, b, c vil $A = a/d, B = b/d, C = c/d$ være et nytt pythagoreisk trippel, så ethvert slikt trippel a, b, c er ekvivalent med ett trippel A, B, C med største felles faktor lik 1. Fra likningen (1) kan man så se at da består A, B, C av parvis relativt primiske tall.

Teorem. *La a, b, c være parvis relativt primiske naturlige tall slik at $a^2 + b^2 = c^2$. Ved eventuelt å bytte a med b kan vi anta at b er et partall. Da finnes det relativt primiske naturlige tall p og q slik at*

$$a = q^2 - p^2 \qquad b = 2pq \qquad c = q^2 + p^2. \tag{2}$$

Så ved å liste opp relativt primiske p og q med $p < q$ fremkommer alle pythagoreiske tripler som a, b, c gitt ved formlene ovenfor, opp til skalering. For eksempel er det siste babylonske eksempelet gitt ved $p = 40, q = 81$.

La oss bevise dette teoremet.

Bevis. La a, b, c være et pythagoreisk trippel, og definer to rasjonale tall x og y ved $x = a/c$ og $y = b/c$. Så

$$x^2 + y^2 = (a/c)^2 + (b/c)^2 = (a^2 + b^2)/c^2 = 1, \tag{3}$$

dvs. (x, y) er et punkt på enhetssirkelen i planet med rasjonale koordinater.

Betrakt en figur med enhetssirkelen hvor det er trukket en rett linje fra $(-1, 0)$ til (x, y) som skjærer sirkelen i disse to punktene. Linjen har stigningstall $t = y/(x+1)$, som igjen er et rasjonalt tall. Vi kan skrive

$$y = t(x + 1), \tag{4}$$

og setter inn for y i (3), og får en annengradslikning i x som bestemmer x-koordinaten til skjæringspunktene uttrykt ved t:

$$x^2 + t^2(x + 1)^2 = 1$$

eller

$$x^2(1 + t^2) + x(2t^2) + (t^2 - 1) = 0.$$

Vi kjenner en rot $x = -1$ til denne annengradslikningen, og kan dermed faktorisere venstresiden som

$$(x + 1)(x(1 + t^2) + (t^2 - 1)) = 0,$$

dvs. $x = -1$ eller

$$a/c = x = \frac{1 - t^2}{1 + t^2}.$$

Fra (4) får vi

$$b/c = y = \frac{2t}{1 + t^2}.$$

Stigningstallet $t \in (0, 1)$ er et rasjonalt tall, så kan skrives på formen $t = p/q$ med $p < q$ relativt primiske naturlige tall. Vi setter inn for t og ganger opp med q^2 i teller og nevner, og får:

$$a/c = x = \frac{q^2 - p^2}{q^2 + p^2}$$
$$b/c = y = \frac{2pq}{q^2 + p^2}$$

Siden a, b, c er relativt primiske må da $a = q^2 - p^2$, $b = 2pq$ og $c = q^2 + p^2$, som ønsket. □

Fermats siste teorem

Et av de siste matematiske verk fra antikken som ble oversatt fra gresk til latin var grekeren Diophantus' bok *Arithmetic*, som opprinnelig ble skrevet omkring år 250 etter Kristus. På slutten av 1630–tallet studerte den franske juristen Pierre de Fermat den da nylig oversatte teksten, og der hvor Diophantus diskuterer hvordan man kan skrive et rasjonalt tall som en sum av kvadrater av to andre rasjonale tall, kommenterte Fermat i margen (på latin):

> "On the other hand, it is impossible to separate a cube into two cubes, or a biquadrate into two biquadrates, og generally any power except a square into two powers with the same exponent. I have discovered a truly marvellous proof of this, which however the margin is not large enough to contain."

I moderne formulering:

Fermats siste teorem. *Det finnes ingen naturlige tall x, y, z og n med $n \geq 3$ slik at*

$$x^n + y^n = z^n.$$

Fermat publiserte selv nesten ingen bevis (kun ett), og i sine brev til andre matematikere presenterte han gjerne bare sine resultater og overlot til mottageren å rekonstruere et bevis. Så det er uklart hvilket bevis han hadde i tankene i dette tilfellet. I sin korrespondanse refererer han til alltid bare til tilfellene $n = 3$ og $n = 4$, så det er sannsynlig at Fermat ikke hadde noe bevis for $n \geq 5$. Etter Fermats død i 1665 publiserte hans sønn hans marg–notater, og slik ble mange av Fermats formodninger bevart og studert i ettertiden.

Navnet "Fermats siste teorem" kommer antageligvis fra at det ved begynnelsen av 18–hundretallet var blitt gitt bevis for alle hans mange andre formodninger, og at dette da var den siste gjenværende uoppklarte formodningen.

La oss se på Fermats bevis i tilfellet $n = 4$.

Bevis. Fermat beviser at likningen $x^4 + y^4 = z^4$ ikke har noen løsninger i naturlige tall x, y, z ved å vise at likningen

$$x^4 + y^4 = u^2 \tag{5}$$

ikke har noen slike løsninger. Resultatet følger ved å sette $u = z^2$.

La x, y, u være en løsning til (5). Vi skal vise at det da finnes en mindre løsning X, Y, U til (5) i den forstand at $U^2 < u^2$.

Vi skriver (5) som $(x^2)^2 + (y^2)^2 = u^2$ og kan finne relativt primiske naturlige tall p, q slik at

$$x^2 = q^2 - p^2 \qquad y^2 = 2pq \qquad u = q^2 + p^2 \tag{6}$$

ved Diophantus' formler (3).

Likningen for x^2 gir oss $x^2 + p^2 = q^2$, så det finnes relativt primiske naturlige tall a, b slik at

$$x = b^2 - a^2 \qquad p = 2ab \qquad q = b^2 + a^2. \tag{7}$$

Derfor er

$$y^2 = 2pq = 2(2ab)(b^2 + a^2) = 4ab(a^2 + b^2)$$

et kvadrattall. Her er ab og a^2+b^2 relativt primiske, for et primtall som deler a men ikke b deler ikke a^2+b^2, og tilsvarende med a og b byttet om.

Hvis et produkt av relativt primiske tall er et kvadrattall, så er hvert av de to tallene er kvadrattall. Dette følger lett ved å se på den entydige primtallsfaktoriseringen av tallene.

Så både a, b og a^2+b^2 er kvadrattall, og vi kan skrive

$$a = X^2 \qquad b = Y^2 \qquad a^2+b^2 = U^2.$$

Da er

$$X^4 + Y^4 = a^2 + b^2 = U^2$$

så X, Y, U er en løsning av (5), med

$$U^2 = a^2 + b^2 = q < q^2 + p^2 = u < u^2$$

fra (6) og (7).

Altså har vi funnet en ekte mindre løsning X, Y, U enn x, y, u. Ved å gjenta konstruksjonen får vi en uendelig nedadstigende følge av løsninger, som er absurd fordi den nedadstigende følgen av naturlige tall $u^2 > U^2 > \dots$ må stoppe opp en gang.

Dette gir en motstrid til eksistensen av en løsning til (5), og vi er fremme. □

Tallteori

Det er klart at for å vise teoremet for alle $n \geq 3$ er det tilstrekkelig å vise det når n er et ulike primtall, og i tilfellet $n = 4$. For hvis $n = pq$ og $x^n + y^n = z^n$ vil $(x^p)^q + (y^p)^q = (z^p)^q$, så x^p, y^p, z^p utgjør en løsning for eksponenten q.

På 1700–tallet gir Euler et (ufullstendig) bevis for tilfellet $n = 3$ av Fermats siste teorem.

1816: Det franske akademiet utlover en pris for en løsning på Fermats siste teorem. En ny pris utloves i 1850, men trekkes tilbake i 1856 på Cauchys anbefaling, og en medalje gis til Kummer.

1820–årene: Sophie Germain viser at hvis p og $2p+1$ er primtall, så har $x^p+y^p = z^p$ ingen løsning med $p \nmid xyz$.

Fra 1825 til 1839 viser Dirichlet, Legendre og Lamé tilfellene $n = 5$ og $n = 7$ av Fermats siste teorem. Så i 1847 presenterer Lamé og Cauchy feilaktige bevis for generelle n.

Begynnende tre år før, i årene 1844–47 publiserte Kummer sine grunnleggende bidrag til den moderne algebraiske tallteorien. Kummer (og Lamé og Cauchy) tok utgangspunkt i faktoriseringen

$$z^p = x^p + y^p = (x+y)(x+\zeta y)\dots(x+\zeta^{p-1}) \tag{8}$$

med $\zeta = e^{2\pi i/p}$ en primitiv $p$te–rot av enheten. (p er her et odde primtall.) Her ligger $x + \zeta^i y$ i et tallsystem (en ring) av komplekse tall på formen

$$a_0 + a_1\zeta + \dots + a_{p-1}\zeta^{p-1}$$

med a_i hele tall. Slike tall kalles syklotomiske (sirkeldelende) heltall. Det gir mening å snakke om primtall i dette tallsystemet, og ethvert syklotomisk heltall

kan faktoriseres som et produkt av primtall i dette tallsystemet. Men generelt vil en slik faktorisering ikke være entydig for alle p. Hvis dette tallsystemet hadde entydig faktorisering ville hver faktor $(x + \zeta^i y)$ være en $p$te–potens, siden disse faktorene er relativt primiske og har produkt lik z^p, fra (8). Dette var Lamés utgangspunkt for sitt feilaktige bevis, men Kummer hadde allerede vist i 1844 at for $p \geq 23$ har det syklotomiske tallsystemet ikke entydig faktorisering.

For å avlaste dette innførte Kummer et generalisert tallbegrep, såkalte ideelle tall, som har gitt opphav til ideal–begrepet i abstrakt algebra. Kummer viste at disse ideelle tallene har entydig faktoriserings–egenskapen. Videre innførte han klassetallet $h = h_p$ som måler i hvilken grad entydig faktorisering slår feil. $h_p > 1$ for $p \geq 23$.

I 1847 viste Kummer at hvis p er et odde primtall slik at $p \not| h_p$ så holder Fermats siste teorem for eksponenten p. Primtallet p kalles da et regulært primtall. Ca. 60% av alle primtall forventes å være regulære, men det er ukjent om det virkelig finnes uendelig mange slike. De eneste irregulære (ikke regulære) primtallene mindre enn 100 er $37, 59, 67$.

Med dette vokste algebraisk tallteori som et moderne felt, og frem til 1992 hadde forfininger av Kummers metode vist at et moteksempel til Fermats siste teorem måtte involvere en eksponent p større enn fire millioner, og heltall x, y, z større enn $4,5 \times 10^{1.911.370}$. (Å skrive ned et moteksempel x, y, z i titallssystemet ville kreve nesten seks millioner siffer, eller flere tusen A4–ark.)

Faltings' bevis av Mordell–formodningen gir at for en gitt $n \geq 4$ er det bare endelig mange inekvivalente løsninger.

$K_4(\mathbb{Z}) = 0$ gir at gitt et odde primtall p finnes det en eksplisitt e_0 slik at det ikke er noen løsninger for $n = p^e$ med $e \geq e_0$.

De siste ti årene

Ved en forelesning i Oberwolfach i 1985 skisserte Gerhard Frey hvordan en løsning a, b, c til likningen $a^p + b^p = c^p$ ville gi opphav til en meget spesiell kurve. Denne Frey–kurven er gitt ved

$$y^2 = x(x - a^p)(x + b^p). \tag{9}$$

Dette er en glatt kurve i planet. Hvis vi i stedet oppfatter x og y som komplekse tall bestemmer denne likningen en figur i $\mathbb{C}^2 = \mathbb{C} \times \mathbb{C}$ med to reelle frihetsgrader, dvs. en reell flate. Likevel kalles dette en kompleks kurve, fordi den er av kompleks dimensjon 1. Ved å føye til punkter ute ved uendelig på en passende måte viser det seg at denne komplekse kurven er topologisk ekvivalent med en torus, dvs. overflaten til en smultring, og Frey–kurven kalles derfor en *elliptisk* kurve.

Frey skisserte hvordan den såkalte Taniyama–Shimura formodningen om elliptiske kurver (definert over $\mathbb{Q}$) ville medføre Fermats siste teorem. Jean–Pierre Serre viste så hvordan en del av hans formodning om "level reduction for modular Galois representations" ville fullføre Freys bevisskisse. Denne delen av Serres formodning ble så vist av Ken Ribet i 1986, og dermed var det bevist at Fermats siste teorem for alle $n \geq 3$ ville følge fra Taniyama–Shimura formodningen.

La oss nå forklare denne formodningen.

La H være det øvre halvplanet $H = \{z = x + iy \mid y > 0\}$. Hvis $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ er en heltallig 2×2-matrise med determinant $ad - bc = 1$ definerer

$$z \mapsto \frac{az + b}{cz + d}$$

en avbildning av H til seg selv. Hvis N er et naturlig tall kalles mengden $\Gamma(N)$ av matriser med $a \equiv d \equiv 1 \mod N$ og $b \equiv c \equiv 0 \mod N$ en kongruensgruppe.

En kompleks kurve (som spesielt er en reell flate) E kalles *modulær* hvis den kan parametriseres ved en (surjektiv holomorf) funksjon $f : H \to E$, som er periodisk for virkningen av en kongruensgruppe på H. Det vil si, for alle $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ i $\Gamma(N)$ og $z \in H$ er

$$f(\frac{az+b}{cz+d}) = f(z).$$

Taniyama–Shimura formodningen. *Enhver elliptisk kurve definert over $\mathbb{Q}$ er modulær.*

Frey, Serre og Ribets resultater viser at Frey–kurven, hvis den eksisterer, ikke kan være modulær. Merk at Frey–kurven bare eksisterer hvis det finnes en løsning i naturlige tall a, b, c til likningen $a^p + b^p = c^p$.

Om morgenen den 23. juni 1993, mot slutten på en forelesningsrekke ved Isaac Newton–instituttet i Cambridge, England, annonserte så den britiske matematikeren Andrew Wiles at han hadde bevist en stor del av Taniyama–Shimura formodningen, nemlig at enhver såkalt semistabil elliptisk kurve definert over $\mathbb{Q}$ er modulær. Frey–kurven er semistabil, og må derfor være modulær, hvis den eksisterer. Denne motsigelsen med Frey, Serre og Ribets resultater viser at Frey–kurven ikke kan eksistere, og at det derfor ikke finnes noen løsning a, b, c til likningen $a^p + b^p = c^p$ for p et odde primtall. Dermed er Fermats siste teorem bevist.

Samme dag spredde nyheten seg via Internettet. Dagen etter nådde resultatet forsidene på flere av verdens største aviser (New York Times, Guardian, Le Monde).

Noen uker senere ble et 200–siders manuskript innlevert til tidsskriftet *Inventiones Mathematicae*, som fordelte oppgaven med å kontrollere beviset mellom fem "referees". Ekspertenes reaksjoner var først positive. Wiles var allerede kjent som en pålitelig matematiker, f. eks. gjennom sitt bevis av den såkalte Iwasawas main conjecture, publisert i 1990.

Så i oktober 1993 spredte det seg rykter om at det var en feil i beviset, knyttet til konstruksjonen av et såkalt Euler–Kolyvagin–system i en Selmer–gruppe. Og i desember sirkulerte et brev fra Wiles hvor han sier at det er et gap i argumentet, men at han ville forklare hvordan det skulle repareres i sitt kurs ved Princeton University, New Jersey, i løpet av våren 1994.

Våren 1994 takker Wiles nei til et tilbud om å reklamere for klesforretningskjeden The Gap i USA.

Ved den internasjonale matematikk–kongressen i Zürich i august 1994 står Wiles frem og sier klart fra at han ikke har et bevis for Fermats siste teorem.

Så, den 7. oktober 1994 frigis to nye manuskripter fra Princeton University. Ett kortere fellesarbeide mellom Wiles og hans tidligere student Richard Taylor, med tittelen "Ring theoretic properties of certain Hecke algebras," og et 136–siders arbeide av Wiles, med tittelen "Modular elliptic curves and Fermats last theorem."

Konstruksjonen av Euler–Kolyvagin–systemet som tidligere skapte problemer er nå skiftet ut med en Hecke–algebra, som vises å være et "komplett snitt" i det første arbeidet. Dette var en eldre ide av Wiles, som han i første omgang forlot da han ble kjent med Euler–Kolyvagin–systemer i en artikkel av Flach fra 1992. Det var i forbindelse med denne nye ingrediensen at Wiles viste seg å ha trukket forhastede slutninger, som kanskje er forståelig, gitt spenningen knyttet til å være så nær et bevis for Fermats siste teorem.

Gerd Faltings, også ved Princeton, er en av flere eksperter som nå har gått god for Wiles' nye bevis. Artiklene er inne til vurdering, men ekspertene har nå hatt mer tid til å sette seg inn i Wiles' konstruksjoner, og det er en meget sterk følelse av at problemet omkring Fermats siste teorem nå endelig er løst, drøyt 350 år etter dets formulering, og kanskje 3.500 år etter dets matematiske opphav på de babylonske leirtavlene.