Komposition av kvadratiska former – från Gauss till Bhargava

J. Brzeziński

Matematiska Vetenskaper
Göteborgs universitet
412 96 Göteborg
jub@math.chalmers.se

1 Inledning

Manjul Bhargava var 28 år gammal när han år 2002 blev professor i matematik vid Princeton universitet. Ett år tidigare försvarade han sin doktorsavhandling som handlar om generaliseringar av Gauss komposition av heltaliga binära kvadratiska former och helt nya liknande teorier som gäller i 12 ytterligare fall. Gauss teori presenterades år 1801 i "Disquisitiones Arithmeticae" – ett av Gauss mest berömda verk. Bhargavas nya tolkning av Gauss resultat öppnade vägen till nya kompositionsteorier som han tillämpar på algebraiska talkroppar och deras klassgrupper – objekt av fundamental betydelse i talteorin. Tidigare som student vid Harvard universitet, generaliserade Bhargava i sitt examensarbete funktionen $n!$, vilket ledde till lösningar av några kända och länge öppna problem. Under de senaste åren har han lyckats lösa några gamla problem om heltaliga kvadratiska former. I denna artikel tänker jag berätta om Bhargavas generaliseringar av Gauss idéer. När denna artikel skrivs har dessa resultat publicerats i fyra arbeten i en av de mest ansedda matematiska tidskrifterna "Annals of Mathematics".

2 Binära kvadratiska former

Mer än hälften av "Disquisitiones Arithmeticae" ägnar Carl Friedrich Gauss åt heltaliga binära kvadratiska former dvs funktioner

$$f(x,y) = ax^2 + bxy + cy^2 = \frac{1}{2}(x,y)\begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

där a, b, c är heltal. Gauss förutsatte oftast att b är ett jämnt heltal, men vi kommer att tillåta godtyckliga heltaliga koefficienter. Man kallar

$$M_f = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}$$

för *formens matris*.

Intresset i sådana funktioner har mycket gamla rötter. Den mest berömda av alla binära kvadratiska former är onekligen summan av två kvadrater $x^2 + y^2$. Frågan om vilka naturliga tal som kan skrivas som summa av två naturliga kvadrater undersöktes för länge sedan. Fermats "Julsats" (annonserad av Fermat i ett brev till Mersenne från den 25 december 1640) säger att varje primtal p som lämnar resten 1 vid division med 4 är summa av två naturliga kvadrater $p = x^2 + y^2$, medan primtal som lämnar resten 3 vid division med 4 kan inte skrivas på denna form (naturligtvis är det jämna primtalet 2 en summa av två kvadrater). Fermat var också intresserad av andra kvadratiska former och de tal som dessa kan representera t ex $x^2 + 2y^2$ och $x^2 + 3y^2$. I ett brev till Pascal skrev Fermat att ett udda primtal $p = x^2 + 2y^2$ precis då p ger resten 1 eller 3 vid division med 8, medan $p = x^2 + 3y^2$ då och endast då $p = 3$ eller p ger resten 1 vid division med 3.

Möjligheten att lösa ekvationen $n = ax^2 + bxy + cy^2$ i heltal x, y då n är ett givet heltal (ej nödvändigt ett primtal) beror på formen, men det finns binära kvadratiska former som representerar exakt samma heltal. Om $f(x, y)$ är en given form så är

$$g(x, y) = f(rx + sy, tx + uy), \tag{1}$$

där r, s, t, u är heltal och

$$\det \begin{pmatrix} r & s \\ t & u \end{pmatrix} = ru - st = 1,$$

en annan binär kvadratisk form som representerar exakt samma naturliga tal som $f(x, y)$. Detta är klart: om $n = g(x_0, y_0)$ så är $n = f(x_1, y_1)$, där $x_1 = rx_0 + sy_0$ och $y_1 = tx_0 + uy_0$. Omvänt, om $n = f(x_1, y_1)$, så löser man ekvationssystemet $x_1 = rx + sy, y_1 = tx + uy$, för att hitta x_0, y_0 sådana att $g(x_0, y_0) = f(rx_0 + sy_0, tx_0 + uy_0) = f(x_1, y_1) = n$. Ekvationssystemet kan lösas i heltal x_0, y_0 därför att dess determinant $ru - st = 1$.

Två former f och g relaterade av en likhet (1) kallas ($\mathbb{Z}-$)ekvivalenta. Mera exakt är de ekvivalenta med avseende på verkan (1) av matrisgruppen $\mathrm{SL}_2(\mathbb{Z})$ som består av alla heltaliga matriser

$$\gamma = \begin{pmatrix} r & s \\ t & u \end{pmatrix}$$

med determinant 1. Vi skall beteckna ekvivalensklassen av formen f med $[f]$. Ekvivalenta binära kvadratiska former representerar exakt samma heltal.

Varje kvadratisk form har en viktig invariant – formens *diskriminant*

$$\Delta = \Delta(f) = b^2 - 4ac = -\det M_f.$$

Man inser mycket lätt att alla former g i klassen $[f]$ har samma diskriminant ty (1) säger att

$$g(x, y) = \frac{1}{2}(x, y) \begin{pmatrix} r & t \\ s & u \end{pmatrix} \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix} \begin{pmatrix} r & s \\ t & u \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

dvs $M_g = \gamma^t M_f \gamma$, där γ är matrisen ovan (i γ^t betecknar t matristransponering).

Alltså är $\Delta(g) = -\det M_g = -(\det A)^2 \det M_f = -\det M_f = \Delta(f)$. Observera att diskriminanten $\Delta = b^2 - 4ac$ lämnar resten 0 vid division med 4 om b är jämnt, och resten 1 vid division med 4 då b är udda (dvs $\Delta \equiv 0$ eller 1 (mod 4)). En binär form $f(x, y) = ax^2 + bxy + cy^2$ är indefinit (antar både positiva och negativa värden) då och endast då $\Delta(f) > 0$, och definit (antar endast icke-negativa eller endast icke-positiva värden) då och endast då $\Delta(f) < 0$. Om formen är definit kommer vi att förutsätta att dess värden är icke-negativa dvs $f(x, y) > 0$ om $(x, y) \neq (0, 0)$. En mycket viktig och gammal sats av Eisenstein och Hermite säger att antalet klasser av binära kvadratiska former med given diskriminant är ändligt. Vi skall formulera denna sats endast för positivt definita former dvs för former som endast antar icke-negativa värden. Sådana former har negativ diskriminant.

Sats 1. *Antalet klasser av binära kvadratiska former med given diskriminant är ändligt. Om diskriminanten är negativ så har varje klass exakt en representant $ax^2 + bxy + cy^2$ vars koefficienter satisfierar*

$$-a < b \leq a < c \quad \textit{eller} \quad 0 \leq b \leq a = c$$

En definit binär kvadratisk form vars koefficienter satisfierar olikheterna ovan kallas *reducerad*. Den sista delen av satsen ger en möjlighet att bestämma alla klasser med given negativ diskriminant Δ. I själva verket har vi $4b^2 \leq 4ac = b^2 + |\Delta|$ så att $3b^2 < |\Delta|$ (observera att $|\Delta| = -\Delta$). Alltså är $|b| \leq \sqrt{|\Delta|/3}$ och $4ac = b^2 - \Delta$, vilket innebär att det finns ändligt många möjligheter för b och följaktligen ändligt många möjliga (positiva) delare a, c till $b^2 - \Delta$.

Exempel 1. Om $\Delta = -56$ så har vi i varje klass exakt en reducerad form. Olikheterna ovan ger $|b| \leq \sqrt{56/3} < 5$ så att $|b| \leq 4$. Genom att testa alla b och motsvarande a, c med $4ac = b^2 + 56$ får vi fyra reducerade former med diskriminanten $\Delta = -56$: $x^2 + 14y^2$, $2x^2 + 7y^2$, $3x^2 - 2xy + 5y^2$, $3x^2 + 2xy + 5y^2$. □

3 Komposition

Om man vill karakterisera alla naturliga tal n som är summor av två kvadrater dvs $n = x^2 + y^2$, så kan man utnyttja identiteten

$$(x^2 + y^2)(z^2 + t^2) = (xz + yt)^2 + (xt - yz)^2.$$

Denna identitet (enkel att direkt kontrollera) säger att produkt av två tal som är summor av två kvadrater också är en summa av två kvadrater. Den ger det enklaste exemplet på *komposition* av kvadratiska former (i detta fall $x^2 + y^2$ med sig själv).

Denna komposition kan användas till att karakterisera alla naturliga tal n som kan skrivas som summor av två naturliga kvadrater. Först skriver vi $n = md^2$, där m är kvadratfritt (dvs produkt av olika primtal) och d är ett heltal (sådan framställning är alltid möjlig). Därefter kan man kontrollera att n är en summa av två kvadrater precis då m har som primfaktorer 2 eller primtal som lämnar resten 1 vid division med 4. I en riktning visas det lätt med hjälp av identiteten ovan: Eftersom 2 och alla primtal som lämnar resten 1 vid division med 4 är summor

av två kvadrater, så är också deras produkt m en sådan summa. Faktorn d^2 är "harmlös" eftersom om $m = x^2 + y^2$, så är $n = md^2 = (dx)^2 + (dy)^2$. Bevis i andra riktningen dvs att förekomsten av en primdelare p till m som lämnar resten 3 vid division med 4 gör att n inte kan skrivas som summa av två kvadrater följer lätt när man studerar rester vid division med p av vänster- och högerled i likheten $n = md^2 = x^2 + y^2$ (vi utelämnar detta argument som man t ex kan hitta i boken av Cox [C] på sidan 10).

Formeln för produkt av summor av två kvadrater generaliseras till godtyckliga binära kvadratiska former, men uttrycken inte längre är så enkla som ovan. Man har t ex:

$$(ax^2 + by^2)(z^2 + abt)^2 = a(xz - byt)^2 + b(axt + yz)^2$$

(ett exempel som kommer från Gauss verk), men det är ett mycket specifikt fall. Rent allmänt då man har två helt godtyckliga former blir uttrycken ganska invecklade. Med ganska stor möda definierade Gauss i "Disquisitiones" komposition av binära kvadratiska former på ett sådant sätt att formklasser $[f]$ bildar en grupp med avseende på denna komposition. Frågan är alltså om att definiera produkt av två godtyckliga klasser $[f_1]$ och $[f_2]$ där f_1, f_2 är binära heltaliga primitiva kvadratiska former med samma diskriminant Δ, så att resultatet också är en klass av sådana former. Denna produkt (komposition) skall resultera i en abelsk (dvs kommutativ) gruppstruktur dvs $[f_2][f_1] = [f_1][f_2]$, $([f_1][f_2])[f_3] = [f_1]([f_2][f_3])$, det skall finnas en neutral klass $[e]$ (dvs $[e][f] = [f][e] = [f]$ for varje klass $[f]$) och till varje klass $[f]$ skall det finna dess invers $[g]$ sådan att $[f][g] = [e]$. Gauss definition av kompositionen $[f_1][f_2]$ var ganska komplicerad (rent tekniskt) och det var Dirichlet som kom på en någorlunda enkel definition. Vi följer just den senare definitionen från mitten av 1800-talet.

Vi börjar med några enkla egenskaper hos kvadratiska former vars motivering lämnar vi som övning. Låt $f(x, y) = ax^2 + bxy + cy^2$ vara en heltalig primitiv form. Om $n = ax^2 + bxy + cy^2$ för heltal x, y, så säger vi att representationen av n är *primitiv* om x, y är relativt prima. T ex har a en primitiv representation då man väljer $(x, y) = (1, 0)$. Men det är rent allmänt så att de heltal n som primitivt representeras av former i klassen av f är precis de tal som förekommer som första koefficienten i någon form g av denna klass. Den andra egenskapen är möjligheten att hitta ett tal n som representeras av f och som är relativt primt med ett godtyckligt på förhand givet heltal.

Låt nu $[f_1(x, y)]$ och $[f_2(x, y)]$ vara två klasser av primitiva heltaliga kvadratiska former med samma diskriminant Δ. Låt $f_1(x, y) = a_1x^2 + b_1xy + c_1y^2$. Vi väljer ett udda heltal a_2 relativt primt med a_1 som representeras av f_2 och därefter väljer vi $f_2(x, y) = a_2x^2 + b_2xy + c_2y^2$ som representant av klassen. Man observerar att b_1 och b_2 har samma paritet därför att $b_1^2 - 4a_1c_1 = \Delta = b_2^2 - 4a_2c_2$. Man väljer nu b så att $b^2 - \Delta$ är delbart med $4a_1a_2$. Den möjligheten finns därför att man alltid kan välja b så att b lämnar resten b_1 vid division med $2a_1$ och resten b_2 vid division med a_2 – eftersom $2a_1$ och a_2 är relativt prima så är valet av b möjligt tack vare Kinesiska restsatsen. Vad mera är finns det enbart en sådan rest b vid division med $2a_1a_2$.

Alltså delar $4a_1a_2$ skillnaden $b^2-\Delta$ (ty $b^2-\Delta = b^2-b_1^2+b_1^2-\Delta$ och på samma sätt för b_2). Nu definierar Dirichlet kompsitionen av $[f_1(x,y)]$ och $[f_2(x,y)]$ som klassen av formen:

$$f_3(x,y) = a_1a_2x^2+bxy+cy^2, \quad \text{där} \quad c=\frac{b^2-\Delta}{4a_1a_2}.$$

Den neutrala klassen bestäms av en av följande former:

$$x^2-\Delta y^2 \ \text{ om } \ \Delta \equiv 0 \pmod 4$$

och

$$x^2+xy+\frac{1-\Delta}{4}y^2 \ \text{ om } \ \Delta \equiv 1 \pmod 4.$$

Inversen till klassen av $f(x,y)=ax^2+bxy+cy^2$ definieras mycket enkelt – det är klassen av $g(x,y)=ax^2-bxy+cy^2$, fast det krävs lite arbete för att visa det. Det krävs också en arbetsinsats för att kontrollera andra detaljer som behövs för bevis att klasserna $[f]$ bildar en grupp med avseende på kompositionen. Vi avstår från denna kontroll eftersom det finns en mycket naturlig tolkning av kompositionen som gör dessa egenskaper nästan självklara. Denna tolkning diskuterar vi i sektion 5. Gaussgruppen av heltaliga primitiva binära kvadratiska former med diskriminanten Δ kommer vi att beteckna med $\mathcal{G}(\Delta)$.

4 Kvadratiska kroppar och ideal

Låt $D \neq 1$ beteckna ett kvadratfritt heltal. Då är $K=\mathbb{Q}(\sqrt{D})$ en kvadratisk kroppsutvidgning av de rationella talen $\mathbb{Q}$. Dess element är alla tal $a+b\sqrt{D}$, där $a,b\in\mathbb{Q}$. Vi betecknar med $\bar{x}$ bilden av $x\in K$ i den icke triviala automorfismen av K dvs om $x=a+b\sqrt{D}$, så $\bar{x}=a-b\sqrt{D}$. Då är *spåret* $\mathrm{Tr}(x)=x+\bar{x}$ och *normen* $\mathrm{Nr}(x)=x\bar{x}$ tal tillhörande $\mathbb{Q}$.

Vi betecknar med $R(K)$ ringen av heltalen i K dvs alla tal $x=a+b\sqrt{D}$ för vilka ekvationen $X^2-\mathrm{Tr}(x)X-\mathrm{Nr}(x)=X^2-2aX+(a^2-Db^2)=0$ har heltaliga koefficienter. Det är väl-känt att

$$R(K)=\mathbb{Z}+\frac{\Delta_K+\sqrt{\Delta_K}}{2}\mathbb{Z},$$

där $\Delta_K=D$ om $D\equiv 1 \pmod 4$ och $\Delta_K=4D$ för övriga Δ. Δ_K kallas *diskriminanten* av K. Man visar att en godtycklig delring till $R(K)$ är

$$R(K)_f=\mathbb{Z}+f\frac{\Delta_K+\sqrt{\Delta_K}}{2}\mathbb{Z},$$

där f är ett positivt heltal. Talet $\Delta_K f^2$ kallas *diskriminanten* av ringen $R(K)_f$. Låt $I=\mathbb{Z}\alpha+\mathbb{Z}\beta$ vara en godtycklig $\mathbb{Z}$-modul i K, där α,β är linjärt oberoende över $\mathbb{Q}$. Då

$$\mathcal{O}(I)=\{x\in K \mid xI\subseteq I\}$$

är en delring till K. Därför, är den en av ringarna $R(K)_f$. Om $\mathcal{O}(I) = R(K)_f$, så säger vi att I tillhör $R(K)_f$. Självklart är då I en modul över $R(K)_f$. Alla $R(K)_f-$moduler tillhörande denna ring bildar en grupp med avseende på vanlig multiplikation av moduler för vilken $R(K)_f$ är identiteten. I själva verket, om I tillhör $R(K)_f$ och $\bar{I} = \{\bar{x} \,|\, x \in I\}$, så är också $\bar{I}$ ett ideal tillhörande $R(K)_f$. Vad mera är $I\bar{I} = \mathrm{Nr}(I)R(K)_f$ för ett entydigt rationellt tal $\mathrm{Nr}(I)$ som man kallar för *normen* av I. Därför är $\frac{1}{\mathrm{Nr}(I)}\bar{I}$ inversen till I. (Observera att normen av I kan också definieras som den positiva generatorn av det $\mathbb{Z}-$ideal i $\mathbb{Q}$ som genereras av alla normer $\mathrm{Nr}(x)$ för $x \in I$. Om $I \subseteq R(K)_f$, så är normen av I lika antalet element i kvotringen $R(K)_f/I$). Vi säger att två moduler I och J på K tillhör samma *klass*[1] om $J = \alpha I$, där $\mathrm{Nr}(\alpha) > 0$. Vi betecknar med $[I]$ klassen av I.

Alla idealklasser som man får med hjälp av ideal tillhörande $R(K)_f$ bildar en abelsk grupp med avssende på multiplikationen: $[I][J] = [IJ]$. Detta påstående är mycket lätt att kontrollera därför att vi redan vet att alla ideal tillhörande $R(K)_f$ bildar en grupp och produkt av klasserna inte beror på valet av dess representanter dvs om I hör till samma klass som I' och J till samma klass som J' så är $[IJ] = [I'J']$. Gruppen av alla idealklasser tillhörande $R(K)_f$ betecknas med $Cl(R(K)_f)$. Ett mycket viktigt resultat om talkroppar säger att denna grupp är ändlig. Den kallas *klassgruppen* av $R(K)_f$. Om $f = 1$ har vi $R(K)_f = R(K)$ och gruppen kallas ofta klassgruppen av K. Dess ordning kallas *klasstalet* av K och betecknas med h_K. Det är en av de viktigaste invarianterna av talkroppen som är intressant att beräkna i många olika sammanhang. I nästa sektion visar vi hur en sådan beräkning kan göras mycket enkelt med hjälp av binära kvadratiska former.

5 Korrespondens mellan ideal och former.

Låt i fortsättningen $\sqrt{D}$ beteckna ett en gång för alla fixerat värde av kvadratroten ur D. Ett talpar (α, β), där $\alpha, \beta \in \mathbb{Q}(\sqrt{D})$, $\alpha = a + b\sqrt{D}, \beta = c + d\sqrt{D}$, kallas *positivt orienterat* om $ad - bc > 0$, och *negativt orienterat* om $ad - bc < 0$. Notera att om paret (α, β) är positivt orienterat, så är paret (β, α) negativt orienterat och omvänt. Därför är det alltid lätt att välja ett positivt orienterat talpar.

Sats 2. *Det finns en–entydig korrespondens mellan alla* SL_2-*klasser av heltaliga primitiva binära kvadratiska former med diskriminant* Δ *och idealklasser tillhörande den kvadratiska ringen med diskriminanten* Δ. *Korrespondensen är given på följande sätt:*

Om $f(x, y) = ax^2 + bxy + cy^2$ *är en form med diskriminanten* $\Delta = b^2 - 4ac$ *så är idealklassen som svarar mot klassen av formen* f *klassen av idealet* $I = 2a\mathbb{Z} + (-b + \sqrt{\Delta})\mathbb{Z}$.

Om $I = \mathbb{Z}\alpha + \mathbb{Z}\beta$ *är ett ideal tillhörande delringen av* $\mathbb{Q}(\sqrt{D})$ *med diskriminanten* Δ *och paret* (α, β) *är positivt orienterat, så är klassen av binära kvadratiska former motsvarande klassen av* I*:*

$$\frac{\mathrm{Nr}(\alpha x + \beta y)}{\mathrm{Nr}(I)}$$

[1] Egentligen kallas klasserna definierade på detta sätt för *smala* klasser, medan med klassen av I menas alla $J = \alpha I$, där α är ett godtyckligt nollskilt element i K. Men vi sysslar här endast med smala klasser.

Exempel 2. Vi kunde konstatera att det finns 4 klasser med diskriminant $\Delta = -56$ (se Exempel 1): x^2+14y^2, $2x^2+7y^2$, $3x^2+2xy+5y^2$, $3x^2-2xy+5y^2$. Alltså finns det också 4 idealklasser i kroppen $\mathbb{Q}(\sqrt{-14})$ tillhörande ringen $\mathbb{Z}[\sqrt{-14}]$. Enligt satsen ovan kan dessa 4 klasser representeras av följande 4 ideal:

$$I_1 = 2\mathbb{Z} + \sqrt{-56}\mathbb{Z}, \qquad I_2 = 4\mathbb{Z} + \sqrt{-56}\mathbb{Z},$$
$$I_3 = 6\mathbb{Z} + (-2+\sqrt{-56})\mathbb{Z}, \qquad I_4 = 6\mathbb{Z} + (2+\sqrt{-56})\mathbb{Z}.$$

Se vidare Exempel 3. □

Satsen ovan om korrespondensen mellan klasser av heltaliga primitiva binära kvadratiska former och idealklasser tillhörande en kvadratisk ring (bägge med samma diskriminant) förklarar enkelt gruppstrukturen som introducerades av Gauss. Det är nämligen så att om man har två former f_1 och f_2 så kan man först gå över till motsvarande ideal I_1 och I_2, multiplicera dessa och bilda produkten I_1I_2. Därefter kan man gå tillbaka till den form (eller rättare sagt dess klass) som svarar mot produkten. På det sättet får man precis den form som Gauss definierade som komposition av f_1 och f_2. Det är dock klart att det kräver lite arbete att kontrollera dessa påståenden.

Exempel 3. Vi återkommer till Exempel 2 för att visa hur man komponerar två formklasser. Betrakta formerna $f_2(x,y) = 2x^2+7y^2$ och $f_3(x,y) = 3x^2-2xy+5y^2$. Dessa definierar ideal $I_2 = 4\mathbb{Z}+\sqrt{-56}\mathbb{Z} = [4, \sqrt{-56}]$ och $I_3 = 6\mathbb{Z}+(-2+\sqrt{-56})\mathbb{Z} = [6, -2+\sqrt{-56}]$ (vi använder en kortare beteckning för ideal för att inte upprepa $\mathbb{Z}$ och lättare multiplicera). Vi har

$$\begin{aligned} I_3I_4 &= [6, 2+\sqrt{-56}][6, -2+\sqrt{-56}] = [36, -12+6\sqrt{-56}, 12+6\sqrt{-56}, -60] = \\ &= [12, 12\sqrt{-56}, 12+6\sqrt{-56}] = 6[2, 2\sqrt{-56}, 2+\sqrt{-56}] = \\ &= 6[2, \sqrt{-56}] = 6I_1 \end{aligned}$$

ty $12 = 2\cdot 36 - 60$ och $(2+\sqrt{-56})-2 = \sqrt{-56}$. Detta visar att $[I_2][I_3] = [I_1]$. På liknande sätt visar man att t ex $[I_2]^2 = [I_1]$, $[I_3]^2 = [I_4]^2 = [I_2]$. Man inser att idealklasserna bildar en cyklisk grupp med 4 element. Dessa likheter säger hur man komponerar motsvarande former. Så t ex är kompositionen av klasserna av f_3 med f_4 lika med klassen av f_1. Detta finner uttryck i en ganska lång formel. Vi skriver i stället en kortare identitet som motsvarar $[I_2]^2 = [I_1]$:

$$(2x^2+7y^2)(2z^2+7t^2) = (2xz+7yt)^2 + 14(xt-yz)^2.$$

Naturligtvis vet vi utan denna identitet (enbart tack vare Sats 2) att mot $[I_2]^2$ svarar formen f_1, ty denna form svarar mot $[I_1]$. □

Korrespondensen mellan klasser av heltaliga primitiva binära kvadratiska former och idealklasser tillhörande en kvadratisk ring (bägge med samma diskriminant) är en mycket kraftfull numerisk metod att bestämma alla idealklasser – det är mycket lätt att lista ut alla formklasser och på så sätt alla idealklasser. Därför var det ingen tillfällighet att Bhargavas resultat om nya typer av liknande korrespondenser presenterades för första gången på en konferens ägnat åt numeriska beräkningar i talteori som hölls i Sydney år 2002. En översikt av Bhargavas resultat publicerades därefter i [B1] som en artikel i konferensrapporten.

6 Bhargavas komposition

En kvadratisk form $f(x,y) = ax^2+bxy+cy^2$ kan ses från ett bredare perspektiv som en funktion av vektorer (x,y) från $V = \mathbb{Z} \times \mathbb{Z}$ till $\mathbb{Z}$. Denna funktion är besläktad med en bilinjär funktion

$$F((x,y),(z,t)) = f(x+z,y+t) - f(x,y) - f(z,t) = 2axz + bxt + byz + 2cyt$$

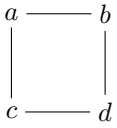
dvs funktionen F är linjär både med avseende på variabeln $v = (x,y)$ och variabeln $w = (z,t)$. Samtidigt är

$$f(x,y) = \frac{1}{2}F((x,y),(x,y))$$

så att f är bestämd av F. Rent allmänt har en bilinjär funktion formen:

$$F((x,y),(z,t)) = (ax+by)z + (cx+dy)t = axz + byz + cxt + dyt,$$

där a,b,c,d är talkoefficienter. Sådan funktion kan beskrivas av en $2 \times 2-$matris, eller en "kvadrat":

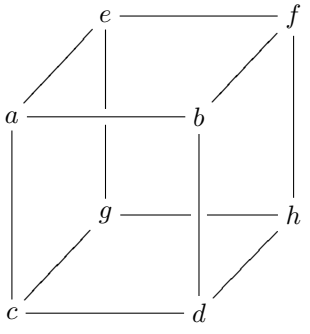


Bhargava studerar trelinjära funktioner från $V \times V \times V$ till $\mathbb{Z}$. Sådana funktioner kan uttryckas på följande sätt:

$$F((x,y),(z,t),(u,v)) = (axz + bxt + cyz + dyt)u + (exz + fxt + gyz + hyt)v.$$

En sådan funktion kan beskrivas med hjälp av en $2 \times 2 \times 2-$matris, eller en "kub":

$\mathcal{C}$:



Man kan naturligtvis betrakta liknande avbildningar F med större antal variabler och försöka tänka på dessa som flerdimensionella "kuber" eller "rätvinkliga parallelepipeder". Vi skall begränsa oss till vanliga kuber. En sådan kub har tre

symmetriplan som delar den i två kvadrater. Man beskriver dessa tre kvadratpar med hjälp av matriser:

$$M_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ N_1 = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$

$$M_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ N_1 = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$

$$M_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ N_1 = \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

Mot varje par ordnar nu Bhargava en kvadratisk form:

$$Q_i(x,y) = -\det(M_i x + N_i y) \tag{2}$$

där $i = 1, 2, 3$. T ex

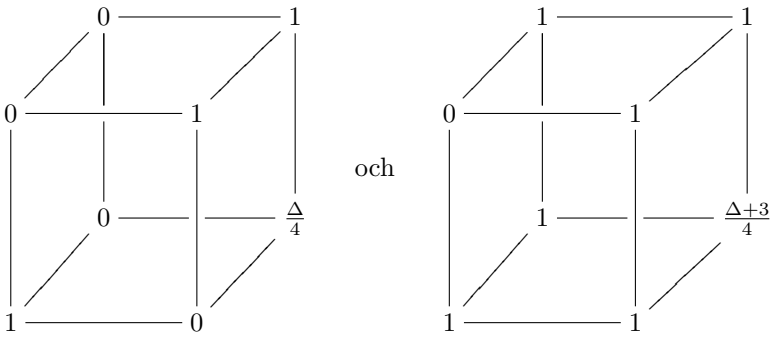
$$\begin{aligned} Q_1(x,y) &= -\det \begin{pmatrix} ax+ey & bx+fy \\ cx+gy & dx+hy \end{pmatrix} = \\ &= (bc-ad)x^2 + (bg+cf-ah-de)xy + (fg-eh)y^2 \end{aligned}$$

Diskriminanten av denna form är lika med

$$\begin{aligned} \Delta(q_1) &= (bg+cf-ah-de)^2 - 4(bc-ad)(fg-eh) = \\ &= a^2h^2 + b^2g^2 + c^2f^2 + d^2e^2 - 2(abgh + cdef + acfh + bdeg + aedh + bfcg) + \\ &\quad + 4(adfg + bceh). \end{aligned}$$

Det intressanta är att samma diskriminant har formerna Q_2 och Q_3. Man kallar denna diskriminant för *kubens* $\mathcal{C}$ *diskriminant* som skall betecknas $\Delta(\mathcal{C})$. Ibland när man vill understryka att formerna Q_i kommer från kuben $\mathcal{C}$ betecknas dessa med $Q_i^{\mathcal{C}}$.

Exempel 4. De kuber som ger de neutrala formerna: $Q_\Delta = x^2 - \Delta y^2$ om $\Delta \equiv 0 \pmod 4$ och $Q_\Delta = x^2 + xy + \frac{1-\Delta}{4} y^2$ om $\Delta \equiv 1 \pmod 4$ är



□

Nu kan vi definiera Gauss grupp av primitiva formklasser med diskriminant Δ och Gauss komposition i enlighet med Bhargavas idé. Denna definition är något abstrakt, men samtidigt mycket elegant. Dessutom kan den ges för flera andra månglinjära funktioner F.

Man startar med gruppen $\mathcal{F}(\Delta)$ bestående av alla summor $\sum n_i[Q]$, där n_i är heltal och symbolerna $[Q]$ svarar mot alla heltaliga primitiva binära kvadratiska former med diskriminanten Δ. Vi förutsätter att nästan alla n_i är lika med 0 dvs summor är ändliga. Man adderar sådana summor genom att addera koefficienterna n_i framför samma $[Q]$. Det är klart att man får en grupp (den är oändlig). Därefter bildar vi kvotgruppen av denna grupp genom att kräva två egenskaper:

"**Kublagen.**" Varje summa $[Q_1] + [Q_2] + [Q_3]$ är lika med 0 om det finns en kub $\mathcal{C}$ sådan att $Q_i = Q_i^{\mathcal{C}}$ (dvs de tre formerna kommer från en kub).

"**Neutrallagen.** Klassen av $[Q_\Delta]$ är 0.

Låt oss påminna om att Q_D definierades i Exempel 4.

Vi betecknar med $\mathcal{H}(\Delta)$ den delgrupp till $\mathcal{F}(\Delta)$ som genereras av alla summor $[Q_1^{\mathcal{C}}] + [Q_2^{\mathcal{C}}] + [Q_3^{\mathcal{C}}]$ och $[Q_\Delta]$. Den grupp som vi får genom att bilda kvoten $\mathcal{B}(\Delta) = \mathcal{F}(\Delta)/\mathcal{H}(\Delta)$ kallar vi *Bhargavagruppen.* Nu kommer överraskningen:

Sats 3. (*Bhargava*) *Gruppen* $\mathcal{B}(\Delta)$ *är isomorf med Gaussgruppen av formklasser* $\mathcal{G}(\Delta)$ *dvs med den idealklassgruppen av den kvadratiska kroppen* $\mathbb{Q}(\sqrt{\Delta})$.

Kompositionen av två former Q_1 och Q_2 sker genom övergång till Bhargavas klasser av dessa former i gruppen $\mathcal{B}(\Delta)$. Därefter adderar man klasserna i denna grupp och summan visar sig vara klassen av samma form som Gauss ordnar mot formerna Q_1 och Q_2 då dessa komponeras i enlighet med hans definition. Detta är långt ifrån självklart och bevisas i [B2]. Bhargavas definition förenklar inte de praktiska beräkningarna (när de behövs), men den förklarar att Gauss komposition har en tolkning som är generaliserbar till andra situationer. På så sätt föds nya tankar och nya möjligheter att hantera andra matematiska problem.

Bhargavs definition är mycket elegant och för direkt tankarna till möjliga generaliseringar. Först och främst definierar Bhargava komposition av kuber dvs en gruppoperation som gör det möjligt att addera två godtyckliga kuber $\mathcal{C}_1$ och $\mathcal{C}_2$ eller snarare deras klasser $[\mathcal{C}_1]$ och $[\mathcal{C}_2]$ under verkan av gruppen $\mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z})$. För att definiera verkan av denna grupp på kuber betraktar man kuber $\mathcal{C}$ som "vektor av tre kolonnvektorer":

$$\mathcal{C} = \left(\begin{pmatrix} M_1 \\ N_1 \end{pmatrix}, \begin{pmatrix} M_2 \\ N_2 \end{pmatrix}, \begin{pmatrix} M_3 \\ N_3 \end{pmatrix} \right).$$

Man låter en matris $\gamma = \begin{pmatrix} r & s \\ t & u \end{pmatrix}$ verka på $\begin{pmatrix} M \\ N \end{pmatrix}$ genom matrismultiplikation dvs

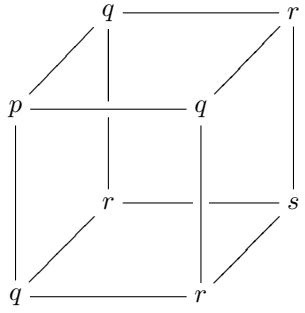
$$\gamma \begin{pmatrix} M \\ N \end{pmatrix} = \begin{pmatrix} rM + sN \\ tM + uN \end{pmatrix}.$$

Alltså om $\gamma = (\gamma_1, \gamma_2, \gamma_3) \in \mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z})$, så är verkan av γ på en kub $\mathcal{C}$ definierad koordinatvis dvs $\begin{pmatrix} M_i \\ N_i \end{pmatrix}$ multipliceras med γ_i för $i = 1, 2, 3$.

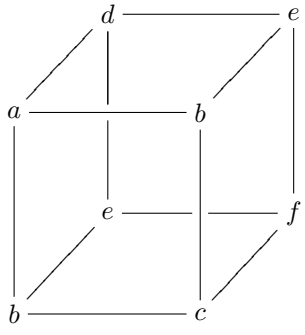
Nu när klasser av kuber $[\mathcal{C}]$ är definierade betraktar Bhargava alla kuber med samma diskriminant $\Delta = \Delta(\mathcal{C})$ som definierar primitiva kvadratiska former (i enlighet med (2)) – sådana kuber kallar Bhargava för *projektiva*. Definitionen av summan av två kuber $[\mathcal{C}_1]$, $[\mathcal{C}_2]$ kräver några tekniska detaljer som leder till att det finns en kub $[\mathcal{C}_3]$ som på ett naturligt sätt är summan av de två givna: $[\mathcal{C}_1] + [\mathcal{C}_2] = [\mathcal{C}_3]$. Därefter är det lätt att kontrollera att denna operation definierar en grupp. Bhargava beskriver gruppen av kubklasser och uttrycker den genom Gaussgruppen $\mathcal{G}(\Delta)$ av alla idealklasser i kroppen $\mathbb{Q}(\sqrt{D})$ (mera exakt, gruppen av alla kuber är isomorf med produkten $\mathcal{G}(\Delta) \times \mathcal{G}(\Delta)$).

Med Bhargavas kuber och deras grupp associeras flera andra intressanta matematiska objekt som ”ärver” gruppstrukturen från kubernas addition. Låt mig nämna två enklaste exemplen.

Det första är kubiska binära former $f(x, y) = px^3 + 3qx^2y + 3rxy^2 + sy^3$, p, q, r, s heltal. Dessa associeras med kuber av formen:



Genom att använda sig av addition av kubklasser får Bhargava en helt ny komposition (gruppoperation) i mängden av binära kubiska former. En annan ny komposition får han för par av kvadratiska former $(ax^2 + 2bxy + cy^2, dx^2 + 2exy + fy^2)$ som associeras med kuber



– dvs den första formen i paret svarar mot den främre kvadraten, och den andra mot den bakre kvadraten.

Bhargava definierar komposition (gruppoperation) för 13 nya objekt (multilinjära funktioner) som är av intresse i olika sammanhang. Hans nya syn på komposition

ger en möjlighet att se struktur (en gruppoperation) i dessa mängder. Den ger en möjlighet att klassificera dessa objekt. Det gör Bhargava när det gäller kroppsutvidgningar av grader 2 (detta är klassiska fallet av kvadratiska talkroppar), 3 (kubiska utvidgningar av de rationella talen), 4 (kvartiska utvidgningar) och 5 (kvintiska utvidgningar). Dessa möjligheter att klassificera olika kroppsutvidgniar leder vidare till satser om viktiga invarianter av sådana kroppar som t ex diskriminanter och klasstal ([B5]). Precis som i det klassiska fallet av binära kvadratiska former och kvadratiska kroppar får man kraftfulla metoder att klassificera och räkna upp antalet objekt med givna egenskaper. Det gäller kubiska, kvartiska och kvintiska kroppsutvidgningar som behandlas i olika delar av Bhargavas avhandling (se [B2], [B3], [B4]; det kvintiska fallet väntar fortfarande för publicering).

Referenser

[B1] Bhargava, M.: *Gauss Composition and Generalizations, in Algorithmic Number Theory* (Sydney 2002), Lecture Notes in Computational Sciences **vol. 2369**, 1–8, Springer Verlag, Berlin (2002).

[B2] Bhargava, M.: Higher composition laws I: A new view on Gauss composition, and quadratic generalizations, *Annals of Mathematics* **159**, 217–250 (2004).

[B3] Bhargava, M.: Higher composition laws II: On cubic analogues of Gauss composition, *Annals of Mathematics* **159**, 865–886 (2004).

[B4] Bhargava, M.: Higher composition laws III: The parametrization of quartic rings, *Annals of Mathematics* **159**, 1329–1360 (2004).

[B5] Bhargava, M.: The density of discriminants of quartic rings and fields, *Annals of Mathematics* **162**, 1031–1063 (2005).

[C] Cox, M.: *Primes of the form x^2+ny^2, Fermat, Class Field Theory, and Complex Multiplication*, John-Wiley &Sons, Inc., New York 1989.

[G] Gauss, C. F.: *Disquisitiones Arithmeticae*, Yale University Press i översättning av A.A. Clarke, 1965, Springer Verlag, 1986.