Algebra II

Isac Hedén och Johan Björklund

Innehåll

0 Introduktion

Målet med den här kursen är att introducera abstrakt algebra, främst genom att först studera talteori och därefter undersöka vilka talteoretiska begrepp som går att utvidga och överföra till mer generella matematiska objekt än $\mathbb{Z}$. Vi börjar med att påminna om några grundläggande definitioner och resultat från talteorin, såsom delbarhet, entydig primtalsfaktorisering och kongruensräkning. Vi kommer också att bevisa den kinesiska restsatsen och studera RSA-algoritmen - en mycket viktig metod för kryptering. Sedan inför vi begreppet ring - ett matematiskt objekt där vi, precis som i $\mathbb{Z}$, kan multiplicera och addera elementen enligt vissa naturliga räkneregler. När vi bekantat oss med de grundläggande ringbegreppen går vi vidare till ett mer detaljerat studie av några olika egenskaper som en ring kan ha. Exempel på sådana egenskaper är entydig faktorisering, eller att den saknar nolldelare, eller att vart och ett av dess nollskilda element är inverterbart, eller att det finns en motsvarighet till Euklides algoritm. Vi kommer också att titta på några olika sätt att konstruera nya ringar från sådana vi redan har (till exempel produkter, kvoter och polynomringar), och några egenskaper dessa nya ringar får.

Avsnitten markerade med * är extramaterial och har inget annat syfte än att ge en vink om hur kursinnehållet kan komma till nytta i senare kurser och forskning. Det är alltså inte nödvändigt att ha en fullständig förståelse av dessa när det blir tentadags. För den som vill fördjupa sig ytterligare i talteori rekommenderar vi varmt Lars-Åke Lindahls utmärkta talteorikompendium. Det finns fritt tillgängligt som pdf-fil på hans hemsida vid matematiska institutionen.

1 Talteori

1.1 Rationella tal och decimalrepresentationer

Vi visar ett sätt att definiera de rationella talen $\mathbb{Q}$ utifrån heltalen $\mathbb{Z}$. Vi betraktar den kartesiska produkten $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, och inför en ekvivalensrelation $\sim$ på den genom $(a, b) \sim (c, d)$ om och endast om $ad = bc$.

Anmärkning 1.1. Relationen $\sim$ är en ekvivalensrelation, dvs för varje val av $(a, b), (c, d), (e, f) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ uppfyller den

- $(a, b) \sim (a, b)$ (reflexivitet),
- $(a, b) \sim (c, d) \Leftrightarrow (c, d) \sim (a, b)$ (symmetri), och
- $(a, b) \sim (c, d)$ och $(c, d) \sim (e, f) \Rightarrow (a, b) \sim (e, f)$ (transitivitet).

Som bekant ger en ekvivalensrelation $\sim$ på en mängd M upphov till ekvivalensklasser, dvs delmängder av M på formen $[x] = \{y \in M \mid x \sim y\}$. För två

givna ekvivalensklasser $[x]$ och $[y]$, gäller antingen $[x] = [y]$ eller $[x] \cap [y] = \emptyset$, så givet två ekvivalensklasser är de antingen lika med varandra, eller disjunkta (sedda som delmängder av M). Eller, med andra ord, om två ekvivalensklasser har (minst) ett element gemensamt, så är de lika. Varje element i M ligger i någon av ekvivalensklasserna ($x \in M$ ligger förstås i $[x]$), så om man tar unionen av alla ekvivalensklasserna får man hela mängden M.

Om man å andra sidan har en ändlig familj av disjunkta delmängder $U_1, U_2 \ldots, U_n$ av en mängd M vars union är hela M, så kan man definiera en tillhörande ekvivalensrelation på M. Nämligen den som definieras av att $x \sim y$ om och endast om x och y båda ligger U_i för något $i \in \{1, 2, \ldots, n\}$. Samma definition fungerar även om antalet mängder U_i är oändligt, så länge man har kvar villkoren att $U_i \cap U_j = \emptyset$ för $i \neq j$, och att unionen av alla U_i är M. En sådan samling av disjunkta delmängder vars union är hela mängden, brukar kallas för en *partition* av mängden i fråga. Mängden av ekvivalensklasser i M med avseende på $\sim$ betecknas med $M/\sim$.

Definition 1.2. Ett *rationellt tal* är ett element

$$[(a, b)] \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})/\sim .$$

Här betecknar $\sim$ ekvivalensrelationen som definieras av att $(a, b) \sim (c, d)$ om och endast om $ad = bc$.

Vi definierar en additionsregel på de rationella talen genom

$$[(a, b)] + [(c, d)] = [(ad + bc, bd)]$$

samt en multiplikationsregel genom

$$[(a, b)] \cdot [(c, d)] = [(ac, bd)].$$

För att addera två ekvivalensklasser med varandra måste man alltså börja med att välja en representant för var och en av dem, sedan adderar man dessa representanter enligt regeln $(a, b) + (c, d) = (ad + bc, bd)$. Summan av de två ekvivalensklasserna är sedan, per definition, den ekvivalensklass som de två representanternas summa ligger i. Denna definition är lite problematisk eftersom summan à priori skulle kunna bero på vilka representanter man väljer för termerna. Lyckligtvis gör den inte det. Motsvarande gäller för multiplikationsregeln, så att produkten av två rationella tal enbart beror på faktorerna, och inte på vilka representanter man väljer för dem.

Anmärkning 1.3. Angående räkneregler.

- Additionsregeln är väldefinierad, dvs om $(a, b) \sim (a', b')$ och $(c, d) \sim (c', d')$ så är även $(a, b) + (c, d) \sim (a', b') + (c', d')$.
- Multiplikationsregeln är väldefinierad, dvs om $(a, b) \sim (a', b')$ och $(c, d) \sim (c', d')$ så är även $(a, b) \cdot (c, d) \sim (a', b') \cdot (c', d')$.

Vanligtvis skriver man a/b för ekvivalensklassen $(a, b) \in \mathbb{Q}$.

Detta är ett sätt att skapa de rationella talen utifrån de hela talen. Att ta nästa steg till de reella talen är inte riktigt lika lätt från en algebraisk synvinkel eftersom det kräver någon form av analys. Anledningen till att vi betraktar ekvivalensklasser med avseende på $\sim$ istället för att definiera rationella tal kort och gott som talpar $(a, b) \in \mathbb{Z} \times \mathbb{Z} \setminus \{0\}$, är naturligtvis att en del uttryck a/b skall ses som olika namn på (eller representanter för) ett och samma rationella tal. Till exempel är vi vana vid att

$$\frac{1}{4} = \frac{-16}{-64} = \frac{7}{28},$$

och mycket riktigt har vi även $(1, 4) \sim (-16, -64) \sim (7, 28)$ vilket läsaren enkelt kan kontrollera med hjälp av definitionen. Ekvivalensklasserna representerade av $(0, 1)$ och $(1, 1)$ i $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ motsvarar de rationella talen 0 respektive 1.

Ett annat sätt att representera rationella tal (eller för den delen reella tal) är genom dess decimalutveckling.

Definition 1.4. Vi säger att ett positivt reellt tal r har en decimalutveckling (inte nödvändigtvis unik) $r = R_k R_{k-1} \dots R_0 . r_1 r_2 r_3 ...$ om och endast om $R_k > 0$ samt att $0 \leq R_0, R_1, \dots, R_k, r_1, r_2, \dots \leq 9$ och

$$r = \sum_{i=0}^{k} R_i \cdot 10^i + \sum_{j=1}^{\infty} r_j \cdot 10^{-j}.$$

Om r är ett negativt tal, får man dess decimalutveckling genom att ta decimalutvecklingen för $-r$ och sätta ett minustecken framför. Det reella talet noll, som varken är negativt eller positivt, har förstås decimalutveckling $0.000 \dots$.

I definitionen skrev vi att två olika decimalutvecklingar kan representera samma reella tal. Till exempel har vi

Anmärkning 1.5. $0.999 \dots = 1.000 \dots$

Bevis.

$$\begin{aligned} 0.999 \dots &= \sum_{n=1}^{\infty} 9 \cdot 10^{-n} = 9 \sum_{n=1}^{\infty} \cdot 10^{-n} = 9 \frac{1}{10 - 1} = 1 \\ &= 1.000 \dots \end{aligned}$$

□

Vi är huvudsakligen intresserade av de rationella talen vilka visar sig ha speciella decimalrepresentationer.

Proposition 1.6. *En oändlig decimalutveckling representerar ett rationellt tal om och endast om den förr eller senare blir periodisk.*

Bevis. Beviset för att alla rationella tal har en decimalutveckling som förr eller senare blir periodisk kommer direkt efter Eulers sats (sats 1.55); vi koncentrerar oss här på den andra riktningen, dvs vi antar att talet x har en decimalutveckling som förr eller senare blir periodisk, och visar att det då måste vara rationellt.

Vi kan utan inskränkning anta att $0 \leq x < 1$ eftersom ett tal q är rationellt om och endast om $q + n$ är rationellt, där n är ett godtyckligt heltal. Då har vi alltså ett tal på formen

$$x = 0.a_1a_2 \dots a_k b_1 b_2 \dots b_p b_1 b_2 \dots b_p \dots$$

som så småningom blir periodiskt med period $b_1 b_2 \dots b_p$. Det kan förstås hända att $k = 0$, så att perioden börjar direkt efter decimalpunkten. I vilket fall har vi

$$10^k x = a_1 a_2 \dots a_k.b_1 b_2 \dots b_p b_1 b_2 \dots$$

och

$$10^{k+p} x = a_1 a_2 \dots a_k b_1 b_2 \dots b_p.b_1 b_2 \dots b_p b_1 b_2 \dots,$$

så om vi bildar differensen av dessa får vi

$$(10^{k+p} - 10^k)x = a_1 \dots a_k b_1 \dots b_k - a_1 \dots a_k.000 \dots$$

Om vi kallar heltalet med denna decimalutveckling för m, så har vi

$$x = \frac{m}{10^{k+p} - 10^k},$$

så x är ett rationellt tal. Varför är $10^{k+p} - 10^k \neq 0$? □

Exempel 1.7. Låt $x = 0.171717 \dots$. Då har vi

$$(100 - 1)x = 100x - x = 17.1717 \dots - 0.1717 \dots = 17,$$

så $x = 17/99$.

Exempel 1.8. Låt $x = 1.23171717 \dots$. Då har vi

$$(10^4 - 10^2)x = 12317.1717 \dots - 123.1717 \dots = 12194,$$

så $x = 12194/9900 = 6097/4950$.

Resten av det här kapitlet så kommer vi huvudsakligen att studera heltalen modulo n även om vi kommer att återvända till konstruktionen av de rationella talen långt senare.

1.2 Delbarhet

Definition 1.9. Om a och b är heltal, så sägs b vara *delbart* med a om det finns ett heltal x sådant att $ax = b$. Vi kommer även att säga att b är en *multipel* av a, eller att a är en *delare* till b för att mena samma sak. Vi skriver $a|b$ i så fall. Varje heltal a är delbart med $\pm a$ och ± 1; dessa delare kallas *triviala*.

En lämplig övning för att vänja sig vid delbarhetsbegreppet är att bevisa några av dess egenskaper:

Proposition 1.10. *Låt a, b, och c vara heltal.*

1. *Om $a|b$ och $b \neq 0$, så följer det att $|a| \leq |b|$·*
2. *Om $a|b$, så följer det att $a|bc$.*
3. *Om $a|b$ och $b|c$ så följer det att $a|c$.*
4. *Om $c|a$ och $c|b$ så följer det att $c|(ax + by)$ för godtyckliga heltal x och y.*
5. *Om $a|b$ och $b|a$ så följer det att $a = \pm b$.*
6. *Om $c \neq 0$, så gäller $a|b$ om och endast om $ac|bc$.*

Nollskilda tal har endast ändligt många delare, så givet två heltal a, b varav minst ett är nollskilt, kan de endast ha ändligt många *gemensamma* delare.

Definition 1.11. Vi inför beteckningen (a, b) för den största gemensamma delaren till a och b, dvs (a, b) är det största heltalet som delar både a och b. Vi låter $(0, 0) = 0$, och vi säger att a och b är relativt prima om $(a, b) = 1$.

Exempel 1.12. Heltalet 102 har de positiva delarna $1, 2, 3, 6, 17, 34, 51$, och 102, medan -170 har de positiva delarna $1, 2, 5, 10, 17, 34, 85$, och 170. De gemensamma positiva delarna är således $1, 2, 17$, och 34, så den största gemensamma delaren till 102 och -170 är 34.

Denna metod för att bestämma den största gemensamma delaren till två givna heltal kan bli ganska mödosam ifall talen är stora, så vi ska beskriva ett effektivare sätt, nämligen Euklides algoritm.

Proposition 1.13. *För alla heltal n gäller att $(a, b) = (a - nb, b)$.*

Bevis. Vi visar att c är en gemensam delare till a och b om och endast om c är en gemensam delare till $a - nb$ och b. Det betyder att a och b har precis samma gemensamma delare som $a - nb$ och b, så i synnerhet gäller då att $(a, b) = (a - nb, b)$.

Antag först att $c|a$ och $c|b$. Då har vi även $c|a-nb$ enligt proposition 1.10(4). Alltså är c en gemensam delare till $a - nb$ och b. Om å andra sidan $c|a - nb$ och $c|b$, så har vi $cx = a - nb$ och $cy = b$ för några väl valda heltal x och y. Det följer att $a = cx + nb = cx + ncy = c(x + ny)$, så $c|a$. Alltså är c en gemensam delare till a och b, och beviset är klart. □

Sats 1.14. *[Divisionsalgoritmen] Givet heltal a och b, med $a > 0$, finns det två entydigt bestämda heltal q och r sådana att $b = aq + r$ och $0 \leq r < a$. Talen q och r kallas för kvot respektive rest då b divideras med a.*

Bevis. I (den aritmetiska) heltalsföljden

$$\ldots, b - 3a, b - 2a, b - a, b, b + a, b + 2a, b + 3a, \ldots$$

finns det ett minsta icke-negativt heltal. Vi kallar det talet, som är på formen $b - qa$ för något väl valt q, för r. Då har vi $b = aq + r$, och vårt val av r medför att $0 \leq r < a$. Det visar existensen av heltal q och r sådana att $b = aq + r$ och $0 \leq r < a$, så nu återstår det bara att visa att de är entydigt bestämda. Antag därför att det finns ett till par av heltal q', r' sådana att $b = aq' + r'$ och $0 \leq r' < a$. Då har vi $aq + r = aq' + r'$, varav följer att $a(q - q') = r' - r$ där dessutom $-a < r' - r < a$ eftersom $0 \leq r, r' < a$. Heltalet $r' - r$ har alltså både egenskapen att den har a som en faktor och att dess absolutbelopp är mindre än a, alltså måste $r' - r = 0$, så $r = r'$. Följdaktligen måste $aq = aq'$ och eftersom $a \neq 0$ följer det att $q = q'$. □

Exempel 1.15. Om $a = 13$ och $b = 37$ får vi $q = 2$ och $r = 11$, eftersom $37 = 13 \cdot 2 + 11$.

Definition 1.16. En icke-tom mängd A av heltal kallas för ett *ideal* om den är sluten under addition och under multiplikation med godtyckliga heltal, dvs om den har följande två egenskaper:

1. $x, y \in A \Rightarrow x + y \in A$
2. $x \in A, n \in \mathbb{Z} \Rightarrow nx \in A$.

Exempel 1.17. Mängderna $\{0\}, \mathbb{Z}$, och $\{0, \pm 3, \pm 6, \pm 9, \ldots\}$ är ideal. Mer generellt kan man säga att alla mängder på formen $\{ng \mid n \in \mathbb{Z}\}$ bestående av heltalsmultiplar av ett heltal g, är ideal. Ett sådant ideal sägs vara genererat av g, och betecknas $g\mathbb{Z}$. Med den notationen har vi $3\mathbb{Z} = \{0, \pm 3, \pm 6, \pm 9, \ldots\}$. Lägg märke till att nollidealet $\{0\}$ genereras av 0, och idealet $\mathbb{Z}$ genereras av 1.

Följande proposition visar att de enda ideal som finns är just dessa, som genereras av ett enda element.

Proposition 1.18. *För varje ideal A, finns det ett entydigt ickenegativt heltal g sådant att $A = g\mathbb{Z}$. Generatorn g karakteriseras av att vara det minsta ickenegativa heltal som ingår i A.*

Bevis. Eftersom idealet $\{0\}$ är på formen $g\mathbb{Z}$ (välj $g = 0$), kan vi fortsättningsvis anta att A inte är nollidealet. Låt g vara det minsta positiva heltal som ligger i A (för att förvissa sig om att det finns positiva heltal i A, kan man välja något element $a \in A \setminus \{0\}$. Då ligger även $(-1) \cdot a$ i A, och antingen är a eller $-a$ positivt). Vi visar att $A = g\mathbb{Z}$. Den ena inklusionen, $g\mathbb{Z} \subset A$, följer direkt av

att A är multiplikativt sluten: $g \in A$ medför att $ng \in A$ för varje $n \in \mathbb{Z}$. Då återstår det att visa $A \subset g\mathbb{Z}$. Låt $b \in A$ vara ett godtyckligt element i A. Enligt divisionsalgoritmen finns det heltal q och r sådana att $b = gq + r$ och $0 \leq r < g$. Eftersom A är sluten under multiplikation så har vi $qg \in A$. Men A är även sluten under subtraktion, så då b och gq tillhör A, måste även $r = b - gq$ göra det. Men g är det minsta positiva heltalet i A, och $0 \leq r < g$, så enda möjligheten är att $r = 0$, och det betyder att $b = gq$. Då har vi visat att $b \in A \Rightarrow b \in g\mathbb{Z}$, dvs $A \subset g\mathbb{Z}$. Å ena sidan har vi alltså $g\mathbb{Z} \subset A$, och å andra sidan har vi $A \subset g\mathbb{Z}$. Det visar att $A = g\mathbb{Z}$. □

Låt a och b vara två heltal, och bilda mängden av linjärkombinationer av a och b med koefficienter i $\mathbb{Z}$, dvs

$$A = \{ax + by \mid x, y \in \mathbb{Z}\}.$$

Den är sluten under addition, ty om $ax_0 + by_0$, och $ax_1 + by_1$ tillhör A, så gör även $(ax_0 + by_0) + (ax_1 + by_1) = a(x_0 + x_1) + b(y_0 + y_1)$ det. Den är även sluten under multiplikation med godtyckligt heltal, ty om $ax_0 + by_0 \in A$, och $m \in \mathbb{Z}$, så har vi $m(ax_0 + by_0) = a(mx_0) + b(my_0)$. Alltså är A ett ideal, och kan därmed skrivas som $A = g\mathbb{Z}$ för något väl valt $g \in \mathbb{Z}$, enligt föregående proposition. Följande resultat visar att g kan väljas som den största gemensamma delaren till a och b.

Proposition 1.19. *Givet två heltal a och b, så genereras idealet $\{ax + by \mid x, y \in \mathbb{Z}\}$ av $g = (a, b)$. Eller med andra ord:*

1. *Det finns heltal x_0, y_0 sådana att $ax_0 + by_0 = (a, b)$.*
2. *$ax + by$ är en multipel av (a, b) för varje val av heltal x och y.*

Bevis. Enligt proposition 1.18 har vi $A = g\mathbb{Z}$ för något positivt heltal g, och vi behöver visa att g inte är något annat än just den största gemensamma delaren till a och b. Genom att välja (x, y) som $(1, 0)$ respektive $(0, 1)$ ser vi att både a och b ligger i $A = g\mathbb{Z}$. Det finns alltså $m, n \in \mathbb{Z}$ sådana att $gm = a$ och $gn = b$, vilket betyder att g är en gemensam delare till a och b. Men är g verkligen den största gemensamma delaren? Låt d vara en gemensam delare till a och b. Eftersom $g \in A$ finns det heltal x_0, y_0 sådana att $ax_0 + by_0 = g$, och det följer att $d|g$ (proposition 1.10(4)). Det medför i sin tur, eftersom $g > 0$, att $d \leq g$. Alltså finns det inte några större gemensam delare till a och b än g. □

Som specialfall av proposition 1.19 kan vi dra slutsatsen att heltalet 1 kan skrivas som en linjärkombination med heltalskoefficienter av två givna heltal a och b om och endast om a och b är relativt prima. Nämligen, om $(a, b) = 1$ så har vi

$$\{ax + by \mid x, y \in \mathbb{Z}\} = \mathbb{Z}$$

så i synnerhet tillhör heltalet 1 idealet $\{ax + by \mid x, y \in \mathbb{Z}\}$. Om å andra sidan 1 kan skrivas som linjärkombination av a och b så ligger 1 i idealet $\{ax + by \mid x, y \in \mathbb{Z}\}$ och då måste det ju vara det minsta positiva heltalet i idealet.

Korollarium 1.20. *Varje gemensam delare till a och b delar den största gemensamma delaren till a och b. Alltså, om $c|a$ och $c|b$ så följer det att $c|(a,b)$.*

Bevis. Eftersom $\{ax+by \mid x,y \in \mathbb{Z}\} = (a,b)\mathbb{Z}$, har vi $ax_0+by_0 = (a,b)$ för några väl valda heltal x_0, y_0. Nu tillämpar vi proposition 1.10(4). □

Korollarium 1.21. *1. (ca,cb)=c(a,b) för varje icke-negativt heltal c.*

2. Om $d = (a,b) \neq 0$, så gäller $\left(\frac{a}{d}, \frac{b}{d}\right) = 1$.

Bevis. 1. Enligt proposition 1.19 genereras idealen $A = \{ax + by \mid x,y \in \mathbb{Z}\}$ och $B = \{cax + cby \mid x,y \in \mathbb{Z}\}$ av (a,b) respektive (ca,cb). Vi visar att

$$B = c(a,b)\mathbb{Z};$$

av detta följer att $c(a,b) = (ca,cb)$, eftersom de båda är icke-negativa generatorer till B. Vi har

$$\begin{aligned} n \in B &\Leftrightarrow n = cax_0 + cby_0 \text{ för några heltal } x_0 \text{ och } y_0 \\ &\Leftrightarrow n = c(ax_0 + by_0) \Leftrightarrow n = c(a,b)k \text{ för något heltal } k \\ &\Leftrightarrow n \in c(a,b)\mathbb{Z}. \end{aligned}$$

I mittenraden använde vi att $A = (a,b)\mathbb{Z}$, så att vi kunde skriva $ax_0 + by_0 = (a,b)k$ för något $k \in \mathbb{Z}$.

2. Av (1) följer att $d\left(\frac{a}{d}, \frac{b}{d}\right) = (a,b) = d$. Division med d ger nu resultatet.

□

Proposition 1.22. *Om $(a,b) = 1$, och $a|bc$, så följer det att $a|c$.*

Bevis. Antag att $(a,b) = 1$ och att $a|bc$. Naturligtvis har vi $a|ac$, så a är en gemensam delare till ac och bc. Följdsats 1.20 ger då att $a|(ac,bc)$, och vidare ger följdsats 1.21 att $(ac,bc) = c(a,b) = c$, så $a|c$. □

Proposition 1.23. *Om $a|c$, $b|c$ och $(a,b) = 1$ så följer det att $ab|c$.*

Bevis. Enligt antagande finns det heltal x,y sådana att $ax = c$ och $by = c$. Alltså är a en faktor i by (eftersom $ax = by$), och eftersom $(a,b) = 1$, ger proposition 1.22 att $a|y$, så det finns ett heltal z sådant att $az = y$. Men då har vi $abz = by = c$, så $ab|c$. □

Proposition 1.24. *Om $(a,b) = (a,c) = 1$, så följer det att $(a,bc) = 1$.*

Bevis. Vi använder kommentaren efter proposition 1.19, för att hitta heltal x, y, z, w sådana att $ax + by = 1$ och $az + cw = 1$. Då har vi $bcyw = (1 - ax)(1 - az) = 1 - an$ där $n = z + x - axz$, och det följer att

$$an + bcyw = 1.$$

Alltså kan heltalet 1 skrivas som en linjärkombination av a och bc med heltalskoefficienter, så a och bc är relativt prima. □

I ett exempel ovan beräknade vi (a, b) för två heltal genom att skriva upp alla delare till a och alla delare till b var för sig, och sade att (a, b) var det största tal som förekom både bland delarna a och delarna till b. Vi återvänder nu till problemet att beräkna (a, b), för att presentera ett effektivare sätt. Vi kan förstås anta att både a och b är icke-negativa, och att $a \geq b$.

Om $b = 0$, så har vi $(a, b) = (a, 0) = a$, och vi är klara. I annat fall använder vi proposition 1.13 som säger att $(a, b) = (a - nb, b)$ för alla heltal n. Speciellt har vi alltså att

$$(a, b) = (a - qb, b) = (r, b) = (b, r),$$

där q och r är kvoten respektive resten då a divideras med b. Eftersom $0 \leq r < b$ har vi ersatt paret (a, b) med ett mindre par (b, r) där $r < b < a$, och vi kan upprepa hela proceduren om och om igen. Eftersom vi i varje steg får ett par av ickenegativa heltal som är mindre än i föregående steg, måste vi till slut få ett par av heltal varav det ena är 0. Vi sammanfattar diskussionen i

Sats 1.25. *(Euklides algoritm) Låt a och b vara heltal med $a \geq b \geq 0$. Låt $a_0 = a$ och $b_0 = b$.*

1. *Om $b_0 = 0$, så är $(a, b) = a_0$.*
2. *Om $b_0 \neq 0$ så använder vi divisionsalgoritmen för att hitta q och r sådana att $a_0 = qb_0 + r$ med $0 \leq r < b_0$.*
3. *Sätt $a_0 = b_0$, $b_0 = r$ och gå till (1).*

Algoritmen måste förr eller senare ta slut eftersom följden av b_0 som uppstår är en avtagande följd av icke-negativa heltal.

Exempel 1.26. Vi beräknar $(247, 91)$. Divisionsalgoritmen ger

$$\begin{aligned} 247 &= 2 \cdot 91 + 65 \quad (*) \\ 91 &= 1 \cdot 65 + 26 \quad (**) \\ 65 &= 2 \cdot 26 + 13 \quad (***) \\ 26 &= 2 \cdot 13. \end{aligned}$$

Vi får $(247, 91) = (91, 65) = (65, 26) = (26, 13) = (13, 0) = 13$.

Enligt proposition 1.19 har ekvationen

$$ax + by = (a, b)$$

en heltalslösning x_0, y_0 (vi kommer senare att se att ekvationen till och med har oändligt många heltalslösningar). Genom att nysta upp ovanstående divisionsalgoritm kan vi hitta en lösning till ekvationen $247x + 91y = 13$. Vi börjar med att skriva 13 som linjärkombination av heltalen 65 och 26 med hjälp av $(***)$, sedan använder vi $(**)$ för att skriva 26 som linjärkombination av 91 och 65 - så att även 13 kan skrivas som linjärkombination av 91 och 65. Slutligen använder vi $(*)$ för att skriva 65 (och därmed 13) som linjärkombination av 247 och 91. Att skriva 13 som linjärkombination av dessa är samma sak som att hitta en lösning till ekvationen $247x + 91y = 13$. Beräkningarna ser ut som följer:

$$\begin{aligned} 13 &= 65 - 2 \cdot 26 = 65 - 2(91 - 1 \cdot 65) = \\ &= 3 \cdot 65 - 2 \cdot 91 = 3(247 - 2 \cdot 91) - 2 \cdot 91 = \\ &= -8 \cdot 91 + 3 \cdot 247. \end{aligned}$$

En heltalslösning till ekvationen $247x + 91y = 13$ är alltså $x = 3, y = -8$.

Unionen $I \cup J$ av två ideal $I = a\mathbb{Z}$ och $J = b\mathbb{Z}$ behöver inte nödvändigtvis vara ett ideal. I själva verket är $I \cup J$ ett ideal om och endast om det ena av idealen är helt inneslutet i det andra, dvs om och endast om en av generatorerna a och b är delbar med den andra. Hur som helst så finns det ett *minsta ideal* som innehåller $I \cup J$, nämligen idealet $(a, b)\mathbb{Z} = \{ax + by \mid x, y \in \mathbb{Z}\}$. Man kan alltså karakterisera den största gemensamma delaren (a, b) som den icke-negativa generatorn till det minsta ideal som innehåller unionen $a\mathbb{Z} \cup b\mathbb{Z}$.

Å andra sidan följer det direkt från definitionen av ideal att *snittet* $I \cap J$ av två ideal $I = a\mathbb{Z}$ och $J = b\mathbb{Z}$ är ett ideal. Ett heltal x ligger i $I \cap J$ om och endast om $a|x$ och $b|x$, dvs om x är en *gemensam multipel* till a och b. Idealet $a\mathbb{Z} \cap b\mathbb{Z}$ är således mängden av alla gemensamma multilpler till a och b, och denna observation leder till ett begrepp som kan ses som en slags motsats till största gemensamma delare.

Definition 1.27. Låt a och b vara två heltal. Den icke-negativa generatorn till idealet $a\mathbb{Z} \cap b\mathbb{Z}$ kallas för den *minsta gemensamma (positiva) multipeln* av de två talen, och betecknas med $[a, b]$.

Lägg märke till att $[a, b] = 0$ om $a = 0$ eller $b = 0$, eftersom snittet $a\mathbb{Z} \cap b\mathbb{Z}$ då är lika med nollidealet. Om a och b båda är nollskilda, så är $a\mathbb{Z} \cap b\mathbb{Z}$ ett nollskilt ideal, eftersom det åtminstone innehåller talet ab. Alltså har a och b positiva gemensamma multipler, och $[a, b]$ är den minsta av dessa.

Exempel 1.28. $[30, 42] = 210$, eftersom i följden $30, 60, 90, 120, 150, 180, 210, \ldots$ av multipler av 30, är 210 det första talet som dessutom är en multipel av 42.

Proposition 1.29. *Om c är ett icke-negativt heltal, så gäller* $[ca, cb] = c[a, b]$.

Bevis. Vi visar att

$$[ca, cb]\mathbb{Z} = c[a, b]\mathbb{Z},$$

dvs att de icke-negativa heltalen $[ca, cb]$ och $c[a, b]$ genererar samma ideal. Från det följer förstås att de är lika.

$$[ca, cb]\mathbb{Z} = ca\mathbb{Z} \cap cb\mathbb{Z} = c(a\mathbb{Z} \cap b\mathbb{Z}) = c[a, b]\mathbb{Z}.$$

□

Proposition 1.30. *Om a och b är icke-negativa heltal, så har vi $[a, b](a, b) = ab$.*

Bevis. Om antingen a eller b är noll, så är $[a, b] = 0$, så likheten stämmer uppenbarligen. Därför kan vi anta att a och b båda är positiva, och vi börjar med att visa likheten i fallet då de är relativt prima. Om $(a, b) = 1$, så måste varje gemensam multipel till a och b vara en multipel även av ab, enligt proposition 1.23. Speciellt så kan den minsta gemensamma multipeln inte vara mindre än ab, och därför råder likhet. Vi övergår till det allmänna fallet, där $(a, b) = d$. Då har vi $(\frac{a}{d}, \frac{b}{d}) = 1$, så

$$[\frac{a}{d}, \frac{b}{d}] = \frac{ab}{d^2}$$

på grund av vad vi just bevisat. Multiplikation med d^2 ger då

$$ab = d^2[\frac{a}{d}, \frac{b}{d}] = d[a, b] = (a, b)[a, b].$$

□

1.3 Primtal

Definition 1.31. Ett heltal > 1 kallas för *primtal* om det bara har triviala delare, och annars kallas det sammansatt.

Således är $p > 1$ ett primtal om och endast om $1 < x < p \longrightarrow x \nmid p$

Proposition 1.32. *Låt p vara ett primtal. Om p delar ab, så delar p antingen a eller b.*

Bevis. Antag att $p|ab$ men att $p \nmid a$. Eftersom p bara har triviala delare följer det att $(p, a) = 1$, så $p|b$ enligt proposition 1.22 □

Korollarium 1.33. *Låt p vara ett primtal. Om $p|b_1b_2 \dots b_n$ så gäller $p|b_i$ för något $i = 1, \dots, n$.*

Bevis. Enligt proposition 1.32 har vi $p|b_1b_2 \dots b_n \longrightarrow p|b_1 \vee p|b_2 \dots b_n$. Resultatet följer med hjälp av induktion. □

Sats 1.34 (Aritmetikens fundamentalsats)**.** *Varje heltal $n > 1$ kan skrivas som en produkt av primtal på ett entydigt sätt så när som på faktorernas ordning.*

Bevis. Att en sådan faktorisering överhuvudtaget är möjlig visar man med induktion på följande vis. Antag att varje heltal större än ett som är mindre än n kan skrivas som en produkt av primtal. Heltalet 2 kan uppenbart skrivas som en produkt av primtal på ett entydigt sätt vilket ger en bas för induktionen. Om n är ett primtal, så är n en produkt av primtal, bestående av en enda faktor. Annars är n sammansatt, och då kan vi skriva $n = n_1 n_2$ med $1 < n_1, n_2 < n$, och det följer av induktionsantagandet att både n_1 och n_2 är en produkt av primtal. Alltså är n också det.

För att visa entydigheten i en sådan faktorisering (så när som på faktorernas ordning), antar vi att det finns ett heltal som kan skrivas som produkt av primtal på två olika sätt. I så fall finns det även ett minsta tal n med den egenskapen. Vi har alltså $n = p_1 p_2 \dots p_r = q_1 q_2 \dots q_s$ där p_i och q_i är primtal och de båda sätten att faktorisera n på är olika. Eftersom $p_1 | q_1 q_2 \dots q_s$ har vi att $p_1 | q_i$ för något $i = 1, \dots, s$ (enligt Följdsats 1.33), och genom att numrera om q_i kan vi anta att $p_1 | q_1$, vilket förstås betyder att $p_1 = q_1$. Men då betraktar vi talet

$$\frac{n}{p_1} = p_2 \dots p_r = q_2 \dots q_s;$$

det är mindre än n och kan skrivas som en produkt av primtal på två olika sätt, men det strider mot att n var det minsta talet med den egenskapen, så vi drar slutsatsen att det inte finns några heltal med olika primtalsfaktoriseringar. □

Det är lätt att bestämma (a, b) och $[a, b]$ om vi lyckas skriva a och b som produkt av primtal.

Proposition 1.35. *Låt a och b vara två positiva heltal. Skriv $a = p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}$, $b = p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$ där $p_1, \dots, p_k$ är olika primtal och $m_1, \dots, m_k, n_1, \dots, n_k$ är olika icke-negativa heltal. Låt $d_j = \min(m_j, n_j)$ och $D_j = \max(m_j, n_j)$. Då har vi $(a, b) = p_1^{d_1} p_2^{d_2} \dots p_k^{d_k}$, och $[a, b] = p_1^{D_1} p_2^{D_2} \dots p_k^{D_k}$*

Bevis. Övning. □

I formuleringen av proposition 1.35 ser det ut som om precis samma primfaktorer förekommer både som faktorer i a och i b. Anledningen till att vi kan skriva på det sättet är att vi tillåter exponenterna att vara noll. Till exempel om $a = 12$ och $b = 21$, så har vi primfaktorerna $p_1 = 2, p_2 = 3$ och $p_3 = 7$ exponenterna blir $(m_1, m_2, m_3) = (2, 1, 0)$ och $(n_1, n_2, n_3) = (0, 1, 1)$.

Proposition 1.36. *Det finns oändligt många primtal.*

Bevis. Vi visar att det för varje ändlig samling $\mathcal{P} = \{p_1, \dots, p_n\}$ av primtal går att hitta ett primtal som inte ligger i $\mathcal{P}$. Givet en sådan samling $\mathcal{P}$, låter vi

$N = p_1 p_2 \dots p_n + 1$. Enligt aritmetikens fundamentalsats har N en primfaktor q (som skulle kunna vara lika med N), men eftersom $(N, p_j) = (1, p_j) = 1$ för varje j, medan $(N, q) = q$, kan q omöjligen vara ett av primtalen i $\mathcal{P}$. □

Däremot har vi följande resultat.

Proposition 1.37. *För varje positivt heltal n finns det en följd av n stycken på varandra följande heltal som alla är sammansatta, dvs det finns godtyckligt långa luckor mellan primtalen.*

Bevis. Heltalen $(n+1)! + 2, (n+1)! + 3, \dots, (n+1)! + (n+1)$ är alla sammansatta, eftersom de är delbara med $2, 3, \dots, n+1$ respektive samt inte lika med $2, 3, \dots, n+1$ respektive. □

1.4 Kongruenser

Definition 1.38. Låt m vara ett positivt heltal. Om $m|(a-b)$ så säger vi att a är *kongruent* med b *modulo* m, och skriver $a \equiv b \pmod{m}$. I annat fall säger vi att a och b inte är kongruenta modulo m, och skriver då istället $a \not\equiv b \pmod{m}$.

Villkoret för att $a \equiv b \pmod{m}$ är förstås ekvivalent med att $a = b + mq$ för något heltal q.

Proposition 1.39. *Kongruens modulo m är en ekvivalensrelation, eller med andra ord:*

1. $a \equiv a \pmod{m}$ *för alla* a.
2. *Om* $a \equiv b \pmod{m}$, *så gäller även* $b \equiv a \pmod{m}$.
3. *Om* $a \equiv b \pmod{m}$ *och* $b \equiv c \pmod{m}$, *så gäller även* $a \equiv c \pmod{m}$.

Bevis. Beviset är inte svårt, och lämnas som övning. □

Proposition 1.40. *Låt a, b, c, och d vara heltal.*

1. *Om* $a \equiv b \pmod{m}$ *och* $c \equiv d \pmod{m}$, *så är* $a + c \equiv b + d \pmod{m}$
2. *Om* $a \equiv b \pmod{m}$ *och* $c \equiv d \pmod{m}$, *så är* $ac \equiv bd \pmod{m}$
3. *Om* $a \equiv b \pmod{m}$, *så är* $a^k \equiv b^k \pmod{m}$ *för alla icke-negativa heltal* k.
4. *Låt $f(x)$ vara ett polynom med heltalskoefficienter. Om* $a \equiv b \pmod{m}$, *så är* $f(a) \equiv f(b) \pmod{m}$.

Bevis. Övning. □

I olika tillämpningar kan det vara nödvändigt att beräkna höga potenser a^k modulo m. Om k är litet, kan man förstås utan större problem beräkna detta på det naiva sättet, dvs att utföra $k-1$ multiplikationer, men då k är stort är detta sätt väldigt tidsödande. Istället bör man utföra beräkningen rekursivt, genom att använda att

$$a^k = \begin{cases} (a^{k/2})^2 & \text{om } k \text{ är jämnt.} \\ a \cdot (a^{(k-1)/2})^2 & \text{om } k \text{ är udda.} \end{cases}$$

Vi illustrerar med ett exempel:

Exempel 1.41. För att beräkna $3^{1304} \pmod{121}$ gör vi först en lista av avtagande tal, varav det första är 1304, enligt regeln att efterföljaren till ett tal n är $n/2$ om n är jämnt, och $n-1$ annars. I det här fallet får vi $1304, 652, 326, 163, 162, 81, 80, 40, 20, 10, 5, 4, 2, 1$. Sedan beräknar vi i ordning $3^1, 3^2, 3^4, 3^5, 3^{10}, \ldots$ ända tills vi når 3^{1304}. I varje steg krävs alltså antingen en multiplikation med 3, eller en kvadrering (beroende på om om vi vill dubbla exponenten, eller bara höja den med ett). Resultatet blir som följer:

k	1304	652	326	163	162	81	80	40	20	10	5	4	2	1
$3^k \pmod{121}$	81	9	3	27	9	3	1	1	1	1	1	81	9	3

Definition 1.42. Låt a vara ett heltal. Mängden $\overline{a} = \{x \in \mathbb{Z} \mid x \equiv a \pmod{m}\}$ av heltal som är kongruenta med a modulo m, kallas för a:s restklass modulo m.

Proposition 1.43. *Det finns exakt m stycken restklasser modulo m, nämligen $\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{m-1}$.*

Definition 1.44. Varje mängd $\{x_1, x_2, \ldots, x_m \mid x_i \in \mathbb{Z}\}$ som innehåller exakt en representant för varje restklass modulo m kallas för ett fullständigt restklassystem modulo m.

Exempel 1.45. Mängden $\{0, 1, 2, \ldots, m-1\}$ är ett exempel på ett fullständigt restklassystem modulo m.

Exempel 1.46. Mängden $\{4, -7, 14, 7\}$ är ett fullständigt restklassystem modulo 4.

Givet ett positivt heltal n, brukar mängden $\{0, 1, \ldots, n-1\}$ kallas för *heltalen modulo n*, och betecknas $\mathbb{Z}_n$. Det är då underförstått att man utför addition och multiplikation av dessa element modulo n. Till exempel har vi $4 + 5 = 2$ beräknat i $\mathbb{Z}_7$.

Lemma 1.47. *Om x och y tillhör samma restklass modulo m, så har vi $(x, m) = (y, m)$.*

Bevis. Om $x \equiv y \pmod{m}$, så är $x = y + qm$ för något heltal q, och det följer från proposition 1.13 att $(x, m) = (y, m)$. □

Detta resultat gör att följande definition blir meningsfull.

Definition 1.48. En restklass $\overline{a}$ modulo m kallas relativt prim med m om $(a, m) = 1$.

Definition 1.49. Låt $\varphi(m)$ beteckna antalet restklasser modulo m som är relativt prima med m. Funktionen φ definierad av detta kallas för Eulers φ-funktion. En mängd $\{r_1, r_2, \ldots, r_{\varphi(m)}\}$ kallas för ett reducerat restklassystem om de $\varphi(m)$ heltalen som den innehåller är valda parvis icke-kongruenta och relativt prima med m.

På grund av lemma 1.47 så är $\varphi(m)$ även lika med antalet heltal i mängden $\{0, 1, 2, \ldots, m-1\}$ som är relativt prima med m. Vidare ser vi att $\{y_1, y_2, \ldots, y_{\varphi(m)}\}$ är ett reducerat restklassystem om och endast om $y_i \not\equiv y_j \pmod{m}$ för $i \neq j$ och y_i är relativt prima med m för $i = 1, 2, \ldots, \varphi(m)$.

Exempel 1.50. Bland de positiva heltalen som är mindre än 8, finns det fyra stycken som är relativt prima med 8, nämligen $1, 3, 5, 7$; det följer att $\varphi(8) = 4$, och att $\{1, 3, 5, 7\}$ är ett reducerat restklassystem modulo 8.

Exempel 1.51. Om p är ett primtal, så är vart och ett av talen $1, 2, \ldots, p-1$ relativt prima med p. Det följer att $\varphi(p) = p - 1$, och att $\{1, 2, \ldots, p-1\}$ är ett reducerat restklassystem modulo p.

Exempel 1.52. Låt p vara ett primtal, och k ett positivt heltal. För ett heltal $a > 1$ ser vi att a och p^k är relativt prima om och endast om a inte innehåller någon faktor p (de positiva delarna till p^k är ju $\{1, p, p^2, \ldots, p^k\}$). Bland de positiva heltalen som är mindre än p^k, är alltså samtliga relativt prima med p^k utom talen np där $n = 0, 1, 2, \ldots, p^{k-1} - 1$. Alltså är $\varphi(p^k) = p^k - p^{k-1}$, och ett reducerat restklassystem får man t.ex. genom att ta $\{0, 1, 2, \ldots, p^k - 1\} \setminus \{np \mid n = 0, 1, 2, \ldots, p^{k-1} - 1\}$.

Exempel 1.53. Vi har $\varphi(3^3) = 3^3 - 3^2 = 18$, och för att få ett reducerat restklassystem kan man utgå från mängden $\{0, 1, 2, \ldots, 26\}$ och ta bort talen $\{0, 3, 6, 9, 12, 15, 18, 21, 24\}$.

Proposition 1.54. *Låt $(a, m) = 1$. Låt vidare $\{r_1, r_2, \ldots, r_m\}$ vara ett fullständigt-, och $\{s_1, s_2, \ldots, s_{\varphi(m)}\}$ vara ett reducerat restklassystem modulo m. Då är även $\{ar_1, ar_2, \ldots, ar_m\}$ och $\{as_1, as_2, \ldots, as_{\varphi(m)}\}$ ett fullständigt- respektive reducerat restklassystem modulo m.*

Bevis. För att visa att $\{ar_1, ar_2, \ldots, ar_m\}$ är ett fullständigt restklassystem, räcker det att visa att $ar_i \equiv ar_j \pmod{m} \longrightarrow i = j$, dvs att talen $ar_1, ar_2, \ldots, ar_m$ är parvis icke-kongruenta modulo m. Antag därför att $ar_i \equiv ar_j \pmod{m}$. Då får vi $m|(ar_i - ar_j)$, så $m|a(r_i - r_j)$. Men $(a, m) = 1$ och då ger proposition 1.22 att $m|(r_i - r_j)$, dvs $r_i \equiv r_j \pmod{m}$, vilket endast är möjligt om $i = j$, eftersom $r_1, \ldots, r_m$ är ett fullständigt restklassystem.

Då återstår att visa att $\{as_1, \ldots, as_{\varphi(m)}\}$ är ett reducerat restklassystem. För detta krävs två saker; dels måste talen $as_1, \ldots, as_{\varphi(m)}$ vara parvis icke-kongruenta modulo m, dels måste de vara relativt prima med m. Att de är

parvis icke-kongruenta visas på exakt samma sätt som vi gjorde i detta bevis första del, så det återstår bara att visa att de är relativt prima med m. För detta använder vi proposition 1.24; den säger att om m är relativt primt med både a och s_i, så är m relativt primt även med deras produkt as_i. □

Sats 1.55 (Eulers sats)**.** *Om $(a, m) = 1$, så har vi*

$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

Bevis. Låt $\{s_1, \ldots, s_{\varphi(m)}\}$ vara ett reducerat restklassystem modulo m. Eftersom $(a, m) = 1$, så är även $\{as_1, \ldots, as_{\varphi(m)}\}$ ett reducerat resklassystem enligt proposition 1.54. Till varje i finns det alltså ett (och endast ett) j sådant att $s_i \equiv as_j \pmod{m}$. Nu bildar vi två olika produkter; den första bildar vi genom att multiplicera samtliga faktorer från det reducerade restklassystemet $\{s_1, \ldots, s_{\varphi(m)}\}$, och den genom att multiplicera samtliga faktorer från det reducerade restklassystemet $\{as_1, \ldots, as_{\varphi(m)}\}$. Enligt proposition 1.40(2) blir dessa produkter kongruenta med varandra modulo m:

$$\prod_{j=1}^{\varphi(m)} (as_j) \equiv \prod_{i=1}^{\varphi(m)} s_i \pmod{m}.$$

Vi bryter ut $a^{\varphi(m)}$:

$$a^{\varphi(m)} \prod_{j=1}^{\varphi(m)} s_j \equiv \prod_{i=1}^{\varphi(m)} s_i \pmod{m}.$$

Detta är ekvivalent med att m delar differensen

$$\left(a^{\varphi(m)} \prod_{j=1}^{\varphi(m)} s_j \right) - \left(\prod_{i=1}^{\varphi(m)} s_i \right) = (a^{\varphi(m)} - 1) \prod_{i=1}^{\varphi(m)} s_i.$$

Men eftersom $(m, s_i) = 1$ för $i = 1, \ldots, \varphi(m)$, så kan vi genom upprepad användning av proposition 1.22 på det högra ledet dra slutsatsen att

$$m | (a^{\varphi(m)} - 1),$$

eller med andra ord att $a^{\varphi(m)} \equiv 1 \pmod{m}$. □

En viktig konsekvens av Eulers sats är följande resultat, som brukar kallas för Fermats lilla sats. Det finns även andra bevis för Fermats lilla sats, som inte bygger på Eulers sats.

Sats 1.56 (Fermats lilla sats)**.** *Om p är ett primtal, och $p \nmid a$, så har vi*

$$a^{p-1} \equiv 1 \pmod{p}.$$

För varje heltal a gäller $a^p \equiv a \pmod{p}$.

Bevis. Om $p \nmid a$, så är $(a, p) = 1$. Eftersom $\varphi(p) = p - 1$, så följer satsens första del omedelbart från sats 1.55. Vi multiplicerar denna kongruens med a för att erhålla $a^p \equiv a \pmod{p}$, och detta håller uppenbarligen även om $a \equiv 0 \pmod{p}$, således för alla heltal a. □

Exempel 1.57. Om vi räknar modulo 7 får vi $3^1 \equiv 3, 3^2 \equiv 2, 3^3 \equiv 6, 3^4 \equiv 4, 3^5 \equiv 5$, och slutligen $3^6 \equiv 1$, i enlighet med sats 1.56. Beräkna gärna 2^6 för att försäkra dig om att även det är kongruent med 1 modulo 7.

1.5 Rationella tals decimalutveckling*

Nu har stunden kommit att ta oss an beviset av den andra implikationen i sats 1.6. I samband med satsens formulering bevisade vi ju bara ena riktningen, att om ett reellt tal har en decimalutveckling som förr eller senare blir periodisk, så är talet i fråga rationellt. Det som behövs för att beviset ska bli fullständigt är alltså att se varför alla rationella tal har en decimalutveckling som förr eller senare blir periodisk. Eulers sats är en viktig ingrediens i detta.

Bevis. Låt q vara ett rationellt tal. Eftersom heltal uppenbarligen har en decimalutveckling som förr eller senare är periodisk, kan vi anta att q inte är ett heltal. Vi skriver

$$q = \frac{A}{2^r 5^s b}$$

för några heltal A och b med $r, s \geq 0$; $b > 0$; och $(b, 10) = 1$. Allt detta utan inskränkning: I nämnaren har vi bara faktoriserat ut största möjliga 2-, respektive 5-potens (r och s behöver inte vara strikt större än 0), och så har vi förlängt med (-1) ifall b var negativ från början. Låt $t = \max\{r, s\}$, då har vi

$$q = \frac{A}{2^r 5^s b} = \frac{2^{t-r} 5^{t-s} A}{2^t 5^t b} = 10^{-t} \frac{2^{t-r} 5^{t-s} A}{b} = 10^{-t} \frac{B}{b},$$

där $B = 2^{t-r} 5^{t-s} A$ är ett heltal. Vi behöver alltså visa att

$$10^{-t} \frac{B}{b}$$

har en decimalutveckling som förr eller senare blir periodisk, och detta är fallet om och endast om B/b har en decimalutveckling som förr eller senare blir periodisk, eftersom faktorn 10^{-t} endast bidrar med att skifta decimalerna t steg åt höger. Vi har

$$\frac{B}{b} = N + \frac{a}{b},$$

för något heltal N och ett heltal a sådant att $0 < a/b < 1$. Eftersom N har en decimalutveckling med enbart 0:or efter decimalkommat, räcker det förstås att visa att a/b har en decimalutveckling som förr eller senare är periodisk. Nu ger Eulers sats, eftersom $(10, b) = 1$, att

$$10^k \equiv 1 \pmod{b},$$

där $k = \varphi(b)$. Det betyder att det finns ett (positivt) heltal d sådant att

$$bd = 10^k - 1.$$

Men då har vi

$$\begin{aligned} \frac{a}{b} &= \frac{ad}{bd} = \frac{ad}{10^k - 1} = \frac{ad}{10^k} \times \frac{1}{(1 - 10^{-k})} = \frac{ad}{10^k}\left(1 + 10^{-k} + 10^{-2k} + \ldots\right) = \\ &= ad\left(10^{-k} + 10^{-2k} + \ldots\right). \end{aligned}$$

I näst sista steget använder vi standardformeln för en geometrisk summa med kvot 10^{-k}. Om vi nu kan visa att ad har högst k siffror, så är vi klara, eftersom det då framgår av ovanstående att a/b har en periodisk decimalutveckling med en period som har längd högst k.

Men eftersom $a/b < 1$ så är $ad < bd$, och $bd < 10^k$, så $ad < 10^k$ och vi är klara. □

Av beviset framgår inte bara att q har en decimalutveckling som så småningom blir periodisk, utan även att denna period har en längd som delar $\varphi(b)$.

1.6 Kinesiska restsatsen

Lemma 1.58. *Låt a, b vara heltal som är relativt prima, och antag att den så kallade diofantiska ekvationen $ax + by = c$ har en heltalslösning $(x, y) = (x_0, y_0)$. Då ges samtliga lösningar av*

$$\begin{cases} x = x_0 + nb \\ y = y_0 - an \end{cases}, \quad n \in \mathbb{Z}$$

Bevis. Vi sätter in $(x, y) = (x_0 + nb, y_0 - an)$ i den diofantiska ekvationen för att se att det är en lösning:

$$a(x_0 + nb) + b(y_0 - an) = ax_0 + abn + by_0 - abn = c.$$

I det sista steget använde vi att $(x, y) = (x_0, y_0)$ är en lösning. Nu behöver vi visa att det inte finns några andra lösningar, dvs att givet en lösning (x', y') så finns det ett $n \in \mathbb{Z}$ sådant att $(x', y') = (x_0 + nb, y_0 - an)$. Antag därför att $ax' + by' = c$. Då får vi

$$ax_0 + by_0 = ax' + by'. \tag{1}$$

Efter en omskrivning får vi $a(x_0 - x') = b(y' - y_0)$. Eftersom a är en delare i vänsterledet, så delar a även högerledet. Men a och b är relativt prima, så vi får $a|(y' - y_0)$, alltså att $an = y' - y_0$ för något $n \in \mathbb{Z}$. Det är förstås ekvivalent med att $y' = y_0 + an$ för något $n \in \mathbb{Z}$. Vi sätter in detta i ekvation (1); det ger att $ax' = ax_0 + by_0 - b(y_0 + an)$, så $ax' = ax_0 - abn$. Division med a ger att $x' = x_0 - bn$, och det visar att det inte finns några andra lösningar än just de som beskrivs i lemmat. □

Korollarium 1.59. *Kongruensen $ax \equiv 1 \pmod{m}$ är lösbar om och endast om $(a, m) = 1$, och i så fall är dess lösningar parvis kongruenta med varandra modulo m. Dvs om x_0, x_1 är två lösningar, så är $x_0 \equiv x_1 \pmod{m}$.*

Bevis. Kongruensen $ax \equiv 1 \pmod{m}$ har lösningar om och endast om det finns heltal x, och n sådana att $ax+nm = 1$, och enligt proposition 1.19 är detta fallet om och endast om $(a, m) = 1$. I så fall ges samtliga lösningar, enligt lemma 1.58, av

$$\begin{cases} x = x_0 + km \\ n = n_0 - ka \end{cases}, \quad k \in \mathbb{Z},$$

där $(x, n) = (x_0, n_0)$ är en lösning. Det visar att två olika lösningar x, x_0 är kongruenta med varandra modulo m. □

Nu ska vi titta på en metod för att lösa flera kongruenser samtidigt.

Sats 1.60 (Kinesiska restsatsen)**.** *Systemet av kongruenser*

$$\begin{cases} x \equiv a_1 \pmod{m_1} \\ x \equiv a_2 \pmod{m_2} \\ \vdots \\ x \equiv a_r \pmod{m_r} \end{cases}$$

där $m_1, m_2, \ldots, m_r$ är parvis relativt prima, har en unik lösning modulo $m_1 m_2 \ldots m_r$, dvs det finns en lösning, och om x_0, x_1 är lösningar så är $x_0 \equiv x_1 \pmod{m_1 m_2 \ldots m_r}$.

Bevis. Beviset består av två delar. Vi visar först att det existerar en lösning överhuvudtaget och sedan att den lösningen är unik modulo $m_1 m_2 \ldots m_r$. Vi ansätter lösningen

$$x = b_1 m_2 m_3 \cdots m_r + m_1 b_2 m_3 \cdots m_r + \ldots + m_1 m_2 m_3 \cdots b_r$$

och försöker hitta b_j som löser detta. Ifall vi stoppar in vår ansats i ekvationssystemet så får vi efter förenkling systemet

$$\begin{cases} b_1 m_2 m_3 \cdots m_r \equiv a_1 \pmod{m_1} \\ m_1 b_2 m_3 \cdots m_r \equiv a_2 \pmod{m_2} \\ \vdots \\ m_1 m_2 m_3 \cdots b_r \equiv a_r \pmod{m_r} \end{cases}$$

Var och en av dessa ekvationer har en lösning eftersom $(m_1 m_2 \cdots m_{i-1} m_{i+1} \cdots m_r, m_i) = 1$. Då har vi alltså lösningar $c_j = b_j$ som vi sedan sätter in i vår ansats och får en lösning

$$x = c_1 m_2 m_3 \cdots m_r + m_1 c_2 m_3 \cdots m_r + \ldots + m_1 m_2 m_3 \cdots c_r.$$

För att visa entydighet antar vi att vi har två lösningar x_0 och x_1. Att $x_0 - x_1 \equiv 0 \pmod{m_j}$ för alla j, betyder att $m_j | x_0 - x_1$ för alla j; men eftersom $m_1, m_2, \ldots, m_r$ är parvis relativt prima så medför detta att $m_1 m_2 \ldots m_r | x_0 - x_1$, vilket är precis detsamma som att säga att $x_0 \equiv x_1 \pmod{m_1 m_2 \ldots m_r}$. □

Vi avslutar det här avsnittet med ett användbart resultat om Eulers φ-funktion.

Proposition 1.61. *Om $m_1, m_2 \in \mathbb{Z}$ är relativt prima och $m = m_1 m_2$, så gäller*

$$\varphi(m) = \varphi(m_1)\varphi(m_2).$$

Bevis. Givet ett positivt heltal n, låter vi $\mathcal{F}(n)$ beteckna det fullständiga restklassystemet $\{0, 1, \ldots, n-1\}$ bestående av n tal, och $\mathcal{R}(n)$ delmängden av $\mathcal{F}(n)$ bestående av tal som är relativt prima med n; $\mathcal{R}(n)$ innehåller $\varphi(n)$ element. Både $\mathcal{F}(m)$ och den kartesiska produkten $\mathcal{F}(m_1) \times \mathcal{F}(m_2)$ innehåller m element, och vi definierar en funktion genom

$$\begin{array}{rcl} \tau : \mathcal{F}(m) & \longrightarrow & \mathcal{F}(m_1) \times \mathcal{F}(m_2) \\ x & \mapsto & (x_1, x_2), \end{array}$$

där $(x_1, x_2) \in \mathcal{F}(m_1) \times \mathcal{F}(m_2)$ definieras av villkoren $x \equiv x_1 \pmod{m_1}$, och $x \equiv x_2 \pmod{m_2}$. Vi vill nu visa att denna funktion τ är bijektiv. Eftersom $\mathcal{F}(m)$ och $\mathcal{F}(m_1) \times \mathcal{F}(m_2)$ innehåller lika många element så följer surjektivitet per automatik om vi visar att τ är injektiv. Men detta är precis innehållet i sats 1.60, för givet $(x_1, x_2) \in \mathcal{F}(m_1) \times \mathcal{F}(m_2)$ säger den att det finns ett (unikt) $x \in \mathcal{F}(m)$ sådant att

$$\left\{ \begin{array}{ll} x \equiv x_1 & \pmod{m_1} \\ x \equiv x_2 & \pmod{m_2} \end{array} \right. .$$

Vidare påstår vi att $\tau(x) \in \mathcal{R}(m_1) \times \mathcal{R}(m_2)$ om och endast om $x \in \mathcal{R}(m)$. Antag först att $x \in \mathcal{R}(m)$, dvs $(x, m) = 1$. Då gäller, eftersom $m = m_1 m_2$ att $(x, m_1) = (x, m_2) = 1$. Men då får vi även att $(x_1, m_1) = (x_2, m_2) = 1$, där $\tau(x) = (x_1, x_2)$. Vi använder proposition 1.13 för att dra den slutsatsen. Det betyder att $\tau(x) \in \mathcal{R}(m_1) \times \mathcal{R}(m_2)$. Om å andra sidan $\tau(x) \in \mathcal{R}(m_1) \times \mathcal{R}(m_2)$, så att $(x, m_1) = (x, m_2) = 1$, så har vi även $(x, m_1 m_2) = 1$ enligt proposition 1.24. Med andra ord $x \in \mathcal{R}(m)$.

Detta visar att restriktionen av τ till $\mathcal{R}(m)$ ger en bijektion mellan mängderna $\mathcal{R}(m)$ och $\mathcal{R}(m_1) \times \mathcal{R}(m_2)$. Alltså måste deras antal element $\varphi(m)$ respektive $\varphi(m_1)\varphi(m_2)$ vara desamma. Det bevisar satsen. □

Exempel 1.62. Vi tar en titt på följande system av kongruenser.

$$\left\{ \begin{array}{rcll} x & \equiv & 1 & \pmod{3} \\ x & \equiv & 2 & \pmod{4} \\ x & \equiv & 5 & \pmod{7} \end{array} \right.$$

Med beteckningar som i beviset för sats 1.60, behöver vi hitta tal b_1, b_2, b_3 exempelvis med $0 \leq b_1 < 3$, $0 \leq b_2 < 4$, $0 \leq b_3 < 7$ sådana att

$$\left\{ \begin{array}{rcll} 4 \cdot 7b_1 & \equiv & 1 & \pmod{3} \\ 3 \cdot 7b_2 & \equiv & 2 & \pmod{4} \\ 3 \cdot 4b_3 & \equiv & 5 & \pmod{7} \end{array} \right. .$$

Var och en av dessa ekvationer är på formen $ax \equiv y \pmod{m}$ med $(a, m) = 1$, och motsvarar en diofantisk ekvation som kan lösas med hjälp av Euklides algoritm. Men i det här fallet är de inblandade talen så små att det går snabbare att prova sig fram till rätt lösning. Man upptäcker att $(b_1, b_2, b_3) = (1, 2, 1)$ fungerar, och vi får $x = 4 \cdot 7 \cdot 1 + 3 \cdot 7 \cdot 2 + 3 \cdot 4 \cdot 1 = 82$. Enligt sats 1.60, är detta den enda lösningen i intervallet $[0, 83]$, och samtliga lösningar ges av $x = 82 + 84n$ där $n \in \mathbb{Z}$.

1.7 RSA-kryptering

År 1977 uppfann R.L. Rivest, A. Shamir och L.M. Adleman en krypteringsmetod som sedan dess har kallats för RSA-algoritmen. Den använder kongruensräkning, och svårigheten med att knäcka denna kryptering bygger på svårigheten i att primfaktorisera stora sammansatta tal och att beräkna e:te rötter modulo ett sammansatt tal för ett givet heltal e. RSA-algoritmen bygger på följande sats.

Sats 1.63. *Antag att m är ett positivt kvadratfritt heltal, dvs att varje primtal i en faktorisering $m = p_1 p_2 \dots p_r$ av m bara förekommer en gång, och låt e och d vara positiva heltal sådana att $ed \equiv 1 \pmod{\varphi(m)}$. Då är*

$$a^{ed} \equiv a \pmod{m}$$

för varje heltal a.

Bevis. Vi behöver visa att $m|(a^{ed} - a)$ för varje heltal a. Eftersom m är kvadratfritt så följer detta om vi bara visar att $a^{ed} - a$ är delbart med varje primfaktor i m eller, med andra ord, att vi för varje primfaktor p i m har $a^{ed} \equiv a \pmod{p}$. Låt därför p vara en godtycklig primfaktor i m. Om $a \equiv 0 \pmod{p}$, så gäller det uppenbarligen att $a^{ed} \equiv a \pmod{p}$, så vi kan anta att $a \not\equiv 0 \pmod{p}$.

Enligt antagande har vi $ed = 1 + n\varphi(m)$ för något icke-negativt heltal n, och

$$\varphi(m) = \varphi(p \cdot m/p) = \varphi(p)\varphi(m/p) = (p-1)\varphi(m/p).$$

Alltså har vi $ed = 1 + N(p-1)$ för något icke-negativt heltal N. Reslutatet följer nu av sats 1.56:

$$a^{ed} = a^{1+(p-1)N} = a \cdot (a^{p-1})^N \equiv a \cdot 1^N = a \pmod{p}.$$

□

En offentlig RSA-nyckel består av ett par (m, e) av heltal, där m är *modulen* och e är den offentliga exponenten. Talet m är produkten av två olika primtal p och q (för att krypteringen ska anses någorlunda säker trots dagens snabba datorer, bör man välja primtal större än 2^{512}). Exponenten e måste vara relativt primt med $\varphi(m)$, alltså med $p-1$ och $q-1$, och ofta väljer man den som

något litet primtal, som till exempel $3 = 2+1$, $17 = 2^4+1$, eller $65537 = 2^{16}+1$, eftersom $a^e \pmod m$ kan beräknas väldigt snabbt för dessa särskilda val av e. I praktiken, när man tillämpar RSA-kryptering, så börjar man med att välja en exponent e. Sedan genererar man två slumpmässiga, rejält stora primtal p och q sådana att $(p-1, e) = (q-1, e) = 1$, och låter slutligen $m = pq$.

Den privata nyckeln består av paret (m, d), där d är det unika positiva heltalet mindre än $\varphi(m)$ som uppfyller $ed \equiv 1 \pmod{\varphi(m)}$. Talet d, såväl som talen p, q och $\varphi(m)$ hålls hemliga av innehavaren av den privata nyckeln.

Om nu en person, säg Adam, vill skicka ett hemligt meddelande till Bertil, ägaren av den privata nyckeln, så börjar han med att konvertera sitt meddelande till ett tal på något standardsätt. Till exempel kan han använda den vanliga ASCII-koden. Vill han till exempel skicka meddelandet "Hello!", blir $a = 072101108108111033$, eftersom H,e,l,o,! har ASCII-koder 072,101,108,111,033 respektive. Om meddelandet är för långt, så att $a > m - 1$, så får man dela upp det i flera kortare meddelanden och kryptera var och en av dem för sig. Sedan använder Adam den offentliga nyckeln och beräknar $a^e \pmod m$, och väljer en representant $0 \leq b \leq m-1$ för denna restklass (så att $b \equiv a^e \pmod m$). Talet b är den kryptotext som Adam skickar till Bertil. När Bertil får talet b använder han sin privata nyckel, och beräknar $c = b^d \pmod m$. Enligt sats 1.63 så är $b^d \equiv a^{ed} \equiv a \pmod m$, så Bertil kan återfinna talet a genom att välja den representant för $c \pmod m$ som ligger i intervallet $0 \leq a \leq m-1$.

Om någon utomstående skulle få tillgång till kryptotexten b, så skulle han behöva dra den e:te roten ur b för att återfinna talet a. Det tycks inte finnas någon bra metod för att göra detta förutom just genom att ha tillgång till talet d. Men för att få reda på det så måste man först beräkna $\varphi(m)$, och för att beräkna $\varphi(m)$ behöver man kunna faktorisera m. Anledningen till att RSA-kryptering anses vara en väldigt säker metod, är att det är utom räckhåll för dagens datorer och faktoriseringsalgoritmer att faktorisera 1000-siffriga tal utan att det krävs orimligt lång tid.

Det är viktigt att a inte är för litet i förhållande till m, så att $a^e < m$. För om det vore fallet, skulle man ju helt enkelt kunna beräkna den vanliga e:te roten av b och således återfinna det krypterade meddelandet. Det finns speciella tekniker för att lösa detta probelm, men det är inget som vi går in på i den här framställningen.

En poäng med att ha en publik nyckel är att ingen speciell kommunikation krävs mellan parterna som vill skicka krypterade meddelanden mellan sig innan de kan börja. Om de hade varit tvugna att först kommunicera med varandra för att komma överens om en krypteringsmetod, kunde det ju hända att någon utomstående avlyssnade denna kommunikation. Och i så fall vore ju själva krypteringen helt poänglös efter det, eftersom inte bara mottagaren av kryptot skulle kunna läsa det, utan även den utomstående som fick reda på krypteringsmetoden. En annan fördel med RSA-algoritmen är att vem som helst kan utföra själva krypteringen, men bara ägaren av den privata nyckeln kan dekryptera deras meddelanden.

1.8 Analytisk talteori*

En annan kanske mer överraskande metod att angripa talteoretiska problem är med hjälp av analys, huvudsakligen komplex sådan. Genom att se samband mellan dels analytiska uttryck och dels talteoretiska objekt så kan man få svar på exempelvis frågor om ungefär hur många primtal det finns mellan 1 och n för stora n. Ett av de mest klassiska exemplen på sådana samband är Riemanns ζ funktion.

Definition 1.64. Vi definierar *Riemanns ζ-funktion* som $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ då $Re(s) > 1$.

Då gäller det bland annat att

$$\prod_{p \in P} (1 - \frac{1}{p^s})^{-1} = \zeta(s)$$

där P är mängden av alla primtal. Genom detta samband så kan man snabbt dra vissa slutsatser om primtalen, exempelvis så hade ju produkten på vänster sida varit ändlig och alltså begränsad i närheten av $s = 1$ om det fanns ändligt många primtal. Vi vet att då s närmar sig 1 så närmar sig ζ-funktionen den harmoniska serien vilken divergerar, alltså så kan den inte vara begränsad nära 1 vilket gör att det måste finnas oändligt många primtal.

Ett något kraftfullare resultat som nog är betydligt svårare att få fram med traditionella algebraiska metoder är följande sats. Beviset kommer bara att skissas så eventuella konvergensfrågor och liknande lämnas åt läsaren att visa.

Sats 1.65. *Låt $p(n)$ vara sannolikheten att ett slumpmässigt plockat tal x från mängden av alla positiva heltal mindre än eller lika med n är kvadratfri, dvs att $p^2 \not| x$ för alla primtal p. Då gäller det att $\lim_{n\to\infty} p(n) = \frac{6}{\pi^2}$.*

Bevis. Vi låter n gå mot ∞. Sannolikheten för att ett godtyckligt tal är delbart med något givet primtal p:s kvadrat är då mycket nära $1/p^2$. Sannolikheten att den inte är delbar med p^2 blir då förstås $1 - 1/p^2$. För att ta reda på sannolikheten att den inte är delbart med något primtals kvadrat så måste vi då alltså multiplicera sannolikheterna att den inte är delbar med varje givet primtal. Då får vi

$$\lim_{n\to\infty} p(n) = \prod_{p \in P} (1 - \frac{1}{p^2}) = \frac{1}{\zeta(2)}.$$

Men då man med envariabelanalys kan visa att $\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ så följer det att $\lim_{n\to\infty} p(n) = \frac{6}{\pi^2}$. □

Det bör nämnas att ett av de absolut största problemen inom matematiken, Riemannhypotesen handlar just om ζ-funktionen. Funktionen utvidgas

till hela komplexa talplanet förutom $s = 1$ på ett speciellt sätt. Riemannhypotesen säger då att alla nollställen finns koncentrerade till $s = \frac{1}{2} + it$ där $t \in \mathbb{R}$ eller $s = -2, -4, -6, ...$ Hittills så har alla beräkningar tytt på att detta stämmer, men ingen har ännu lyckats hitta något bevis.

Ifall man är intresserad av analytisk talteori så är det huvudsakligen kurserna "Komplex analys" samt "Analytisk talteori" som man bör läsa.

2 Ringar

Heltal, rationella tal, reella tal och komplexa tal har alla flera intressanta egenskaper. Man kan exempelvis addera och multiplicera dem, det finns ett speciellt tal (0) som inte ändrar något då man adderar det och ett speciellt tal (1) som inte ändrar något då man multiplicerar med det. Även andra matematiska objekt har liknande egenskaper, exempelvis så kan man addera och multiplicera kvadratiska matriser, polynom och reellvärda kontinuerliga funktioner. Det finns även motsvarande ettor och nollor hos dessa objekt. Det gör det mycket lockande att istället för att undersöka dem var och en separat formulera en mer generell definition för deras gemensamma egenskaper och utifrån det härleda satser som kan appliceras på många olika objekt.

2.1 Definitionen av begreppet ring

Definition 2.1. En *ring* är en mängd R tillsammans med två funktioner f, g från $R \times R$ till R kallade multiplikation och addition som skrivs så att $a+b = f(a,b)$, $a \cdot b = g(a,b)$ ($a \cdot b$ skrivs vanligen ab) med följande egenskaper:

1 Det existerar ett element kallat $0 \in R$ så att för alla $a \in R$ så är $a + 0 = a = 0 + a$

2 För alla $a \in R$ så existerar det ett element $b \in R$ så att $a + b = 0 = b + a$, detta element betecknas med $-a$ (att det bara finns ett sådant visas senare)

3 För alla $a, b, c \in R$ så gäller att $(a + b) + c = a + (b + c)$ (Additiv associativitet)

4 För alla $a, b \in R$ så gäller att $a + b = b + a$ (Additiv kommutativitet)

5 För alla $a, b, c \in R$ så gäller att $a(b+c) = ab+ac$ samt att $(b+c)a = ba+ca$ (Distributivitet)

6 För alla $a, b, c \in R$ så gäller att $a(bc) = (ab)c$ (Multiplikativ associativitet)

7 Det existerar ett element kallat $1 \in R$ så att $1a = a = a1$ för alla $a \in R$ (Multiplikativ enhet)

När man diskuterar ringar så brukar man mycket sällan tala explicit om funktionerna f och g, man säger helt enkelt operationerna $\cdot$ och $+$.

I det första ringaxiomet pratar vi om ett speciellt element 0 som har egenskapen att $0 + a = a$ för alla $a \in R$. Det finns bara ett element med denna egenskap, ty antag att $0' + a = a$ för alla a. Då måste det speciellt gälla att $0' + 0 = 0$. Men vi vet ju att $0' + 0 = 0 + 0' = 0'$ enligt 4 och 1. Alltså är det unikt. Det kallas ibland för det neutrala elementet. Liknande bevis fungerar för att visa att det bara finns ett element med ettans egenskaper.

I det andra ringaxiomet så kräver vi att varje element a ska ha en additiv invers, det vill säga ett element b så att $a + b = 0$. Eftersom vi vill kalla det elementet för $-a$ så måste vi visa att det är unikt, det ger oss följande annulleringslag.

Proposition 2.2. *Om $a + b = a + c$ så är $b = c$.*

Bevis. Antag $a + b = a + c$. Låt d vara en additiv invers till a. Då är $d + a + b = d + a + c$. Då $d + a = 0$ så följer att $0 + b = 0 + c$, då 0 är neutralt additivt enligt 1 så följer att $b = c$. □

Det följer trivialt att additiva inversen är unik för varje givet element vilka tillåter oss att ge den ett speciellt namn.

Exempel 2.3. Det är ganska lätt att se att heltalen, $\mathbb{Z}$ tillsammans med den vanliga additionen som addition och den vanliga multiplikationen som multiplikationen är en ring.

Exempel 2.4. Även $n \times n$ matriser med matrisaddition och -multiplikation bildar en ring, denna ring har den intressanta egenskapen att den ej är kommutativ om $n \geq 2$, dvs det existerar x, y så att $xy \neq yx$.

Exempel 2.5. Linjära operatorer (dvs funktioner $T : V \to V$ där V är ett vektorrum så att $T(x + y) = T(x) + T(y)$ och $T(\lambda v) = \lambda T(v)$) bildar en ring med addition och multiplikation definierade punktvis på följande sätt: $(T \cdot S)(x) = T(S(x))$ samt $(T + S)(x) = T(x) + S(x)$.

Bevis. Vi måste kontrollera ringaxiomen ett i taget. Vår mängd R består av alla linjära operatorer på V så det första vi ska kontrollera är att mängden är sluten under addition och multiplikation (dvs att vi inte hamnar utanför mängden när vi adderar eller multiplicerar saker). Låt T och S vara godtyckliga linjära operatorer. Vi kontrollerar först att $T + S$ är en linjär operator.

$$(T + S)(\lambda x + y) = T(\lambda x + y) + S(\lambda x + y) = T(\lambda x) + T(y) + S(\lambda x) + S(y) =$$

$$\lambda T(x) + T(y) + \lambda S(x) + S(y) = \lambda(T + S)(x) + (T + S)(y).$$

Motsvarande bevis för multiplikation lämnas till läsaren. Vårt element 0 är operatorn N som definieras av att $N(x) = 0 \, \forall x \in V$, det inses lätt att den uppfyller kravet. Det är lätt att se att vår additiva invers $(-T)(x)$ är just $-(T(x))$. Då addition är en kommutativ och associativ operation i vektorrum så är den det även för våra operatorer. Att de övriga ringaxiomen är tillfredställda lämnas som en övning. Notera att $n \times n$-matriser är ett specialfall av linjära operatorer för ändligtdimensionella vektorrum med något val av bas. □

Exempel 2.6. Kontinuerliga funktioner från $\mathbb{R} \to \mathbb{R}$ med följande operationer: $(f + g)(t) = f(t) + g(t)$, $(f \cdot g)(t) = f(t) \cdot g(t)$ bildar en ring. Samma sak gäller om vi istället tittar på kontinuerliga funktioner från $\mathbb{C} \to \mathbb{C}$.

Exempel 2.7. De positiva heltalen med standardaddition/multiplikation bildar **inte** en ring då det inte finns något positivt heltal n så att $1 + n = 1$, alltså saknar den additiv enhet.

Exempel 2.8. Polynom i en variabel med heltalskoefficienter bildar en ring kallad $\mathbb{Z}[X]$.

Exempel 2.9. Mängden $\{0, 1, 2, 3\}$ med additionen $0+0 = 0, 0+1 = 1, 0+2 = 2, 0+3 = 3, 1+1 = 2, 1+2 = 3, 1+3 = 0, 2+2 = 0, 2+3 = 1, 3+3 = 2$ och multiplikationen $0 \cdot a = 0, 1 \cdot a = a, 2 \cdot a = a + a, 3 \cdot a = a + a + a$ för alla $a \in \{0, 1, 2, 3\}$. Känner du igen den ringen?

Exempel 2.10. Givet en ring R så är $R \times R$ en ring med operationerna $(a, b) + (c, d) = (a + c, b + d)$ samt $(a, b) \cdot (c, d) = (ac, bd)$.

2.2 Lite notation och grundläggande räkneregler

För att underlätta så inför vi en del notation och visar att vissa grundläggande räkneregler fungerar. Det kan vara nyttigt att studera bevisen och försöka se var de olika ringaxiomen används.

Proposition 2.11. $0 \cdot a = 0$

Bevis.

$$a = a \longrightarrow 1 \cdot a = a \longrightarrow (1 + 0) \cdot a = a \longrightarrow$$

$$1 \cdot a + 0 \cdot a = a \longrightarrow a + 0a = a \longrightarrow a + (-a) + 0a = a + (-a) \longrightarrow 0 + 0a = 0 \longrightarrow 0a = 0$$

□

Proposition 2.12. $(-1)a = -a$

Bevis. Låt $(-1)a = b$ och studera $a + b$. Då $a = 1 \cdot a$ enligt 7 så har vi att $a + b = 1 \cdot a + (-1) \cdot a = (1 + (-1))a = 0 \cdot a = 0$. Alltså så är b en additiv invers till a, denna är unik så vi vet att $b = -a$. □

Proposition 2.13. *Om* $|R| > 1$ *så är* $1 \neq 0$.

Bevis. Bra övning. □

För att underlätta notation så inför vi följande två definitioner:

Definition 2.14. Låt n vara ett positivt heltal. Vi definierar då n sett som ett element i en ring R på följande vis: $n = 1 + 1 + ... + 1$ (n st ettor).

Det inses lätt tack vare associativitet att alla vanliga additioner gäller, t.ex så är $1 + 3 = 2 + 2$ då $(1) + (1 + 1 + 1) = (1 + 1) + (1 + 1)$. Dock så är det inte självklart att varje heltal ger upphov till ett unikt element i ringen. I ringen $\mathbb{Z}_3$ så är exempelvis $2 = 5$.

Definition 2.15. Låt n vara ett negativt heltal. Vi definierar då n sett som ett element i en ring R på följande vis: $n = -(-n)$.

Anmärkning 2.16. Detta ger oss möjlighet att skriva saker på ett betydligt smidigare sätt, exempelvis så är $a + a + a + a + a = 5a$ och $na - ma = (n - m)a$, bevisen är uppenbara.

2.3 Kommutativa ringar och delringar

Definition 2.17. Vi säger att en ring R är *kommutativ* om det för alla $a, b \in R$ gäller att $ab = ba$.

Vi ska i denna kursen huvudsakligen studera kommutativa ringar och kommer därmed att anta att våra ringar är kommutativa i fortsättningen om inget annat nämns. Vi kommer även att anta att våra ringar har egenskapen att $1 \neq 0$. Även ickekommutativa ringar som ringar av matriser eller linjära operatorer är dock mycket intressanta att studera, men vi kommer tyvärr inte att hinna med ytterligare studier av dessa i denna kurs.

Definition 2.18. Givet en ring R så definierar vi *polynomringen* $R[x]$ som mängden av polynom i x, dvs formella summor $\sum_{i=0}^{n} a_i x^i$ där $a_i \in R$ och $n < \infty$. Addition och multiplikation av polynom definieras på samma sätt som vanligt.

Anmärkning 2.19. Notera att vi betraktar polynomen inte som funktioner utan som objekt i sig, om vi exempelvis har polynomet $x^2 + x + 1$ sett som ett polynom med koefficienter i $\mathbb{Z}_2$ så kommer den att vara konstant 1 sett som funktion från $\mathbb{Z}_2$ till $\mathbb{Z}_2$. Sett som polynom så är den dock inte samma polynom som polynomet 1.

I linjär algebra så studerar man ganska ofta delrum till vektorrum, en mindre del av det ursprungliga rummet som fortfarande uppfyller kravet att vara ett vektorrum. Denna idé återfinns även hos ringar.

Definition 2.20. Om R är en ring och $H \subset R$ så är H en *delring* till R om H tillsammans med multiplikationen och additionen från R är en ring i sig själv samt $1 \in H$.

För att kontrollera att en delmängd H av en ring R är en delring så får man mycket gratis då många axiom redan är uppfyllda genom ringstrukturen på R. Det enda som behöver kontrolleras är att H är sluten under multiplikation och addition samt att additiv invers, nolla och etta existerar.

Exempel 2.21. $\mathbb{Z}$ är en delring till $\mathbb{C}$.

Exempel 2.22. För alla ringar R så är R en delring till R.

Exempel 2.23. $\mathbb{Q}$ är en delring till $\mathbb{R}$.

Exempel 2.24. Polynom med enbart jämna potenser av x är en delring till polynomringen med variabeln x, dvs polynom $p(x)$ så att $p(x) = g(x^2)$ för något polynom $g(x)$.

2.4 Nolldelare, integritetsområden och inverterbara element

Då vi ursprungligen motiverades till konstruktionen av ringar utifrån $\mathbb{Z}$ så kan man fråga sig om det finns fler egenskaper från $\mathbb{Z}$ eller $\mathbb{Q}$ som man kan hitta hos ringar. En trevlig egenskap hos $\mathbb{Z}$ är att man kan förenkla ekvationer med hjälp av såkallade annuleringslagar. Vi har visat att den första annuleringslagen, $a + b = a + c \longrightarrow b = c$ gäller för ringar i allmänhet, så en naturlig fråga är om motsvarande sak gäller för multiplikation. Antag att $ab = ac$ och att $a \neq 0$, kan vi då dra slutsatsen att $b = c$? Svaret är tyvärr nej i allmänhet, i exempelvis ringen $\mathbb{Z}_4$ så är ju $2 \cdot 2 = 2 \cdot 0$ men $2 \neq 0$. Tydligen så är 2 i detta fall ett element som faktiskt är en delare till noll, vi formaliserar detta som följer:

Definition 2.25. Vi säger att ett element $a \in R$ är en *nolldelare* om $a \neq 0$ och det existerar ett $b \in R$ så att $b \neq 0$ och $ab = 0$.

Exempel 2.26. Alla element som inte är relativt prima med n är nolldelare i $\mathbb{Z}_n$.

Exempel 2.27. Elementet 1 är aldrig en nolldelare, ty om $a \cdot 1 = 0$ med $a \neq 0$ så gäller $a = a \cdot 1 = 0$ vilket ger motsägelse.

Definition 2.28. Ett *integritetsområde* är en ring som saknar nolldelare.

Några exempel på integritetsområden är $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{Z}[x]$ samt $\mathbb{Z}_p$. Om vi har en ring som är ett integritetsområde så kan vi formulera den andra annulleringsregeln:

Proposition 2.29 (Andra annulleringsregeln)**.** *Om R är ett integritetsområde och $a \neq 0$ så medför $ax = ay$ att $x = y$.*

Bevis. Antag att R är ett integritetsområde, att $a \neq 0$ och att $ax = ay$. Då har vi $ax - ay = 0$, varav det följer att $a(x - y) = 0$. Eftesrom R saknar nolldelare, så gäller speciellt att a inte är en nolldelare, så $x - y$ kan inte vara nollskilt. Alltså är $x = y$. □

Proposition 2.30. *Om R' är en delring till R och R är ett integritetsområde så är även R' det.*

Bevis. Inses lätt. □

Så vi kan inte alltid hitta en unik lösning till linjära ekvationer $ax = b$ även om $a \neq 0$. En naturlig nästa fråga är när de överhuvudtaget är lösbara? Vi tar igen $a \neq 0$ och undersöker $ax = 1$. Det man instinktivt kanske vill göra är att dividera med a och få $x = \frac{1}{a}$. Det är dock inte alls säkert att det existerar en sådan multiplikativ invers till alla nollskilda tal. Även för heltalen så saknar ju exempelvis $2x = 1$ lösningar.

Definition 2.31. Om R är en ring och $a \in R$ så säger vi att a är ett *inverterbart element* om det existerar $b \in R$ så att $ab = 1$. Vi kallar då b för a:s invers och betecknar $b = a^{-1}$ (inversen är unik, ty om $ab = 1$ och $b'a = 1$ så är $b' = b'1 = b'ab = 1b = b$).

Exempel 2.32. 1 är alltid ett inverterbart element i alla ringar ty $1 \cdot 1 = 1$ per definition.

Exempel 2.33. 0 är aldrig ett inverterbart element så länge $|R| > 1$. Detta gäller eftersom $x \cdot 0 = 0 \neq 1$ för alla $x \in R$ vilket innebär att inget ringelement ger resultatet 1 efter multiplikation.

Exempel 2.34. En nolldelare är aldrig inverterbar.

Exempel 2.35. För kommutativa ringar så gör det detsamma från vilket håll man multiplicerar inversen, ty $xy = yx$. Däremot så är inverser inte lika lätta att handskas med för ickekommutativa ringar då det kan hända att det finns en invers från höger men ingen från vänster, dvs att $xy = 1$ men $yx \neq 1$. Om det finns både höger och vänsterinvers för ett element x (dvs ett element y så att $yx = 1$ och ett element z så att $xz = 1$) så är de lika. Bevis: $yx = 1 \longrightarrow zyx = z \cdot 1 = z \longrightarrow (zy) \cdot x = 1 \cdot x = x = z$. Dock så kan det hända att det bara existerar högerinvers men inte vänsterinvers och tvärsom.

2.5 Kroppar

Ekvationen $ax = b, a \neq 0$ har alltså en lösning åtminstone så länge a är ett inverterbart element (den kan dock fortfarande ha en lösning även om a ej är inverterbart, exempelvis ekvationen $3x = 6$ i ringen $\mathbb{Z}$). Ringar med egenskapen att alla element utom noll är inverterbara är mycket speciella, de kallas kroppar

Definition 2.36. En mängd K med två binära operationer $+, \cdot$ kallas en *kropp* om K är en kommutativ ring med egenskapen att det för alla $a \in K \backslash \{0\}$ så gäller att det existerar $b \in K$ så att $ab = 1$.

Proposition 2.37. *En kropp är ett integritetsområde.*

Bevis. Följer från att elementen i en kropp antingen är inverterbara (och alltså inte nolldelare enligt tidigare exempel) eller elementet 0 (som inte är nolldelare) per definition. □

Exempel 2.38. Det finns många exempel på kroppar, $\mathbb{R}, \mathbb{Q}$ och $\mathbb{C}$ är ju några vi känner till sedan länge, men även $\mathbb{Z}_2$ är en kropp (ofta en mycket användbar sådan). Allmänt så gäller att $\mathbb{Z}_p$ är en kropp för varje givet primtal p (vilket vi redan visat då det finns inverser tack vare Euklides algoritm). Det finns även fler kroppar, t.ex rationella funktioner (dvs formella kvoter av polynom) med koefficienter från någon kropp.

Exempel 2.39. Ett intressant exempel på en kropp är de konstruerbara talen. Dessa består av alla tal i $\mathbb{C}$ som man kan skapa utgående från 0 och 1 med hjälp av en passare och ograderad linjal, dvs alla punkter som man kan få fram som skärningar av olika linjer (som går genom två kända punkter) och cirklar (med centrum i en känd punkt och minst en känd punkt på cirkeln för att kunna bestämma passarens radie). Exempelvis så är 2 konstruerbart då det ligger i skärningen av linjen från 0 till 1 och cirkeln med centrum i 1 och 0 som en punkt på cirklen (dvs med radie 1). Det krävs flera intressanta övningar för att se att detta är en kropp.

Exempel 2.40. Ännu ett exempel på en kropp är de så kallade algebraiska talen, vilka definieras som alla komplexa tal som är rötter till ett polynom med heltalskoefficienter (som inte är identiskt noll). Igen så krävs det en del ansträngning för att verkligen visa att detta är en kropp.

En grundläggande egenskap för kroppar är dess karaktäristik som definieras som följer:

Definition 2.41. *Karaktäristiken* hos en kropp K är det minsta positiva heltal n sådant att $na = 0$ för alla $a \in K$. Om inget sådant tal existerar så definieras karaktäristiken som 0.

Ett exempel på en kropp med karaktäristik skild från noll är $\mathbb{Z}_5$ medans ett exempel på en kropp med karaktäristik noll är $\mathbb{R}$. En intressant fråga är vilken karaktäristik en kropp kan ha. Vi vet att kroppar kan ha noll som karaktäristik (realiserat av $\mathbb{R}$) samt att den kan ha primtalskaraktäristik (realiserat av $\mathbb{Z}_p$). Följande sats ger svar på frågan om det kan finnas fler möjligheter.

Proposition 2.42. *Låt n vara karaktäristiken hos en kropp K. Antingen så är $n = 0$ eller så är n ett primtal.*

Bevis. Nyttig övning. □

Kroppar (fields på engelska) är mycket användbara i bland annat linjär algebra. I kurserna ni har läst så har ni troligen antagit att skalärerna är reella tal eller möjligen komplexa tal. Teorin kan dock utan större problem utvecklas till att innefatta vektorrum över godtyckliga skalärkroppar.

Man kan ta ett liknande steg till och ersätta skalärerna med element i en ring, då faller dock betydligt fler saker samman. Exempelvis så spänner $2, 3$ upp $\mathbb{Z}$, dock så är de linjärt beroende. Trots det så kan man inte ta bort någon av dem då de inte ensamma spänner upp $\mathbb{Z}$. Teorin för vad som händer om man ersätter skalärkropparna med ringar är trots det mycket intressant, sådana objekt kallas för moduler istället för vektorrum och kan studeras i kursen "Moduler och homologisk algebra". Dessa objekt ger upphov till mycket kraftfulla verktyg i flera matematiska ämnen, bland annat algebraisk topologi.

En algebraisk fråga man kan ställa sig är hur (och när) man kan konstruera en kropp utifrån en ring, tydligen så går det ju att göra $\mathbb{Q}$ från $\mathbb{Z}$ så kan

man kopiera den konstruktionsidén? Svaret är ja, givet att ringen i fråga är ett integritetsområde.

Definition 2.43. Givet en ring R sådan att R är ett integritetsområde så definierar vi kroppen $Q(R)$, ibland kallad fraktionskroppen på följande sätt: Vi börjar med R och konstruerar mängden $R \times R \backslash \{0\}$ (dvs ordnade par av element ur R sådana att det andra elementet ej är noll). På denna mängd så konstruerar vi en ekvivalensrelation sådan att $(a, b) \sim (c, d)$ om och endast om $ad - bc = 0$. Vi låter $Q(R)$ bestå av ekvivalensklasser under denna relation med följande addition samt multiplikation: $(a, b) + (c, d) = (ad + cb, bd)$, $(a, b) \cdot (c, d) = (ac, bd)$ där multiplikation och addition definieras via representanter.

Proposition 2.44. *$Q(R)$ är en kropp.*

Bevis. Mycket nyttig övning, visa att ekvivalensrelationen är en ekvivalensrelation, att addition samt multiplikation beter sig väl under ekvivalensrelationen (dvs om $x \sim y, z \sim v$ så är $x + z \sim y + v$ samt $xz \sim yv$) samt att det verkligen uppfyller kraven på en kropp. □

Vi konstruerade ju också $\mathbb{Z}_p$ utifrån $\mathbb{Z}$, även denna konstruktion kan generaliseras vilket vi ser i avsnittet "Ideal och kvotringar".

2.6 Galoisteori*

Ett annat håll man kan gå åt är att fråga sig om polynoms lösbarhet. För exempelvis $\mathbb{Q}$ så finns det ju ickekonstanta polynom med koefficienter i $\mathbb{Q}$ som ej har rötter i $\mathbb{Q}$. Detta gäller dock ej ickekonstanta polynom med koefficienter i $\mathbb{C}$ då de alltid har (minst) en rot i $\mathbb{C}$. Då säger man att $\mathbb{C}$ har egenskapen att den är algebraiskt sluten. En fråga man då kan fundera på är hur man kan utvidga $\mathbb{Q}$ (eller för den delen andra ej algebraiskt slutna kroppar) för att kunna lösa fler och fler polynomekvationer och samtidigt få större och större kroppar som ligger emellan $\mathbb{Q}$ och $\mathbb{C}$.

Ifall man exempelvis lägger till $\sqrt{2}$ så kan man ju få en rot till polynomet $X^2 - 2$. Vad händer om man lägger till alla kvadraträtter av rationella tal (och sedan deras summor, kvoter etc)? Man kan då lösa alla andragradspolynom med koefficienter i $\mathbb{Q}$ eftersom vi har en formel som uttrycker rötterna för en andragradare med hjälp av koefficienterna, kroppsoperationerna ock kvadraträtter. Det visar sig att om man på motsvarande sätt tillåter kubik och fjärderötter så får man med rötter till alla fjärdegradspolynom (detta visas inte helt lätt genom att man hittar formler för att lösa dem). Däremot så får man **inte** med rötterna till alla polynom av grad 5 bara genom att ta 2-, 3- ,4- och 5-rötter. Detta medför att det **inte** finns en motsvarande lösningsformel för femtegradare.

Man använder också liknande typer av resonemang för att lösa en del rent geometriska problem, så kallade konstruktionsproblem då man funderar

på vilka punkter som kan konstruera på ett plan med hjälp av enbart passare och ograderad linjal. Detta kan man läsa mer om i kursen "Algebraiska strukturer" och senare "Galoisteori".

2.7 Faktoriella ringar och primelement

När vi studerade $\mathbb{Z}$ i talteoridelen av kursen bevisade vi bland annat att varje tal har en unik primtalsfaktorisering. Några uppenbara frågor är vad som skulle kunna menas med primtal och unik faktorisering om man pratar om ringar istället. Det är inte helt självklart hur man skall utvidga primtal till ringar i allmänhet.

Ett sätt är att fokusera på egenskapen hos primtal att de inte går att faktorisera i andra tal än inverterbara tal och $\pm$ primtalet självt.

Definition 2.45. Vi säger att ett ickeinverterbart element $a \in R$ där R är ett integritetsområde är *irreducibelt* om $a = bc$ medför att endera b eller c är inverterbart.

Detta säger alltså att irreducibla element är sådana som inte kan "plockas isär" i mindre delar. Några exempel är att primtalen i $\mathbb{Z}$ är irreducibla. De är dock inte de enda irreducibla elementen i $\mathbb{Z}$ då exempelvis -2 också är irreducibel.

Definition 2.46. Låt $a, b \in R$ vara två element i en ring. Vi säger att a och b är *associerade* med varandra om det finns ett inverterbart element $c \in R$ sådant att $a = bc$, $b = ac^{-1}$.

Två element är alltså associerade om de skiljer sig åt endast med en inverterbar faktor. Bland heltalen är de enda inverterbara elementen $\{\pm 1\}$, och därför är två heltal a och b associerade till varandra om och endast om $a = b$ eller om $a = (-1)b = -b$.

Om vi istället betraktar ringen $\mathbb{R}[x]$ bestående av polynom i en variabel x med reella koefficienter, så är exempelvis polynomen $x + 1$ och $\sqrt{2} + \sqrt{2}x$ associerade med varandra, eftersom de bara skiljer sig åt med den inverterbara faktorn $\sqrt{2}$ (dess invers är förstås $1/\sqrt{2} \in \mathbb{R}$). Mer generellt är två polynom p och q i $F[x]$, där F är en kropp, associerade till varandra om och endast om $p = \lambda q$, där $\lambda \in F \setminus \{0\}$.

Vi lämnar som en övning att visa att $\sim$ definierad på en ring R av att $a \sim b$ om och endast om a och b är associerade, definierar en ekvivalensrelation. Det som ska visas är alltså att givet tre element $a, b, c \in R$, så gäller

- $a \sim a$,
- $a \sim b \Leftrightarrow b \sim a$,

- $a \sim b$ och $b \sim c \longrightarrow a \sim c$.

Vi lämnar även som övning att visa att om $a \in R$ är irreducibelt så är varje element som är associerat till a också irreducibelt.

En annan möjlighet är att fokusera på primtalens delbarhetsegenskaper för att definiera primelement.

Definition 2.47. Låt R vara ett integritetsområde. Vi kallar ett icke-inverterbart element $a \in R$ för *primelement* om

$$a|bc \longrightarrow a|b \text{ eller } a|c,$$

för element $b, c \in R$.

Ett element $a \in R$ är med andra ord ett primelement om och endast om det har följande egenskap. Om a delar en produkt av två element, så måste det dela minst en av de två faktorerna.

Exempel 2.48. Låt R vara ett integritetsområde. Visa att alla primelement i R är irreducibla.

Omvändningen till detta påstående är emellertid inte sann. Ett elemet kan vara irreducibelt utan att vara ett primelement. Till exempel gäller detta för $2 \in \mathbb{Z}[i\sqrt{5}]$, vi tittar närmare på den saken.

Per definition har vi att

$$\mathbb{Z}[i\sqrt{5}] = \{a + bi\sqrt{5} \mid a, b \in \mathbb{Z}\}$$

består av de komplexa tal som kan skrivas på formen $a + bi\sqrt{5}$ där a och b är heltal. Eftersom de komplexa talen har strukturen av en ring får vi naturligt inducerade additions- och multiplikationsregler på $\mathbb{Z}[i\sqrt{5}]$. För att visa att $\mathbb{Z}[i\sqrt{5}]$ verkligen blir en ring med dessa operationer räcker det att visa att den är additivt- respektive multiplikativt sluten och att $0, 1$ ligger i den. Exempelvis är

$$(a + bi\sqrt{5})(c + di\sqrt{5}) = ac - 5bd + i\sqrt{5}(ad + bd),$$

så om $a + bi\sqrt{5}$ och $c + di\sqrt{5}$ ligger i $\mathbb{Z}[i\sqrt{5}]$, så gör även deras produkt det.

Exempel 2.49. Elementet $2 \in \mathbb{Z}[i\sqrt{5}]$ är irreducibelt. Antag nämligen att

$$2 = (a + bi\sqrt{5})(c + di\sqrt{5}).$$

Genom att jämföra kvadraten på absolutbeloppen av dessa komplexa tal, får vi att $4 = (a^2 + 5b^2)(c^2 + 5d^2)$. Vi drar slutsatsen att $(a^2 + 5b^2) \leq 4$, eftersom $(c^2 + 5d^2)$ inte kan vara mindre än 1. Men då följer det att $b = 0$. Med precis samma resonemang följer det att $d = 0$, och vi får likheten $2 = ac$. Men eftersom a och c är heltal, så får vi att det ena av dem, säg a, måste vara endera 1 eller -1, och då följer det att $(a + bi\sqrt{5})$ ligger i $\{\pm 1\}$, och är således inverterbart i $\mathbb{Z}[i\sqrt{5}]$. Då drar vi slutsatsen att 2 är irreducibelt, eftersom det inte går att skriva som en produkt av icke-inverterbara faktorer.

Exempel 2.50. Elementet $2 \in \mathbb{Z}[i\sqrt{5}]$ är inte ett primelement. Vi har nämligen att

$$2|(1+i\sqrt{5})(1-i\sqrt{5}),$$

utan att för den sakens skull 2 delar någon av de två faktorerna (lägg märke till att $(1+i\sqrt{5})(1-i\sqrt{5})$ bara är ett annorlunda sätt att skriva 6 på). Vi visar att 2 inte delar den första faktorn. Om $2(a+bi\sqrt{5}) = 1+i\sqrt{5}$ för något element $a+bi\sqrt{5} \in \mathbb{Z}[i\sqrt{5}]$, så får vi genom att jämföra real- respektive imaginärdel i de båda leden att $2a = 1$ respektive $2b = 1$. Detta är förstås omöjligt om $a, b \in \mathbb{Z}$, och på samma sätt visas att 2 inte heller delar den andra faktorn.

Exempel 2.51. Visa att om a är ett primelement och b är associerat till a så är b ett primelement.

Detta ger oss någon form av uppfattning om vad primtal skulle kunna generaliseras till (om man dessutom ser $-2, -3, -5$ osv som någon slags primtal). Det ger oss också möjligheten att prata om unik faktorisering, som vi vet så har varje heltal en unik faktorisering i primtal (upp till tecken och ordning). Motsvarande egenskap för ringar kallas att vara en faktoriell ring.

Definition 2.52. Vi definierar en ring R som *faktoriell* om R är ett integritetsområde sådant att varje nollskilt ickeinverterbart element $a \in R$ är en produkt av irreducibla element. Samt att om två produkter $p_1p_2\cdots p_m = q_1q_2\cdots q_n$ av irreducibla element är lika så är $m = n$ och termerna kan omindexeras så att p_i är associerat till q_i.

Ringarna $\mathbb{Z}$, $\mathbb{C}[x]$ samt $\mathbb{Q}$ är alla exempel på faktoriella ringar, medan $\mathbb{Z}[i\sqrt{5}]$ är ett exempel på en icke-faktoriell ring. Bevisen lämnas som en övning.

Proposition 2.53. *Alla kroppar är faktoriella.*

Bevis. Alla kroppar är integritetsområden, och då alla element i en kropp är antingen noll eller inverterbara så finns det inga ickeinverterbara nollskilda element och alltså inga irreducibla element. □

3 Homomorfier och isomorfier

Ett ganska allmänt problem inom matematiken är att ta mängder med vissa sorts strukturer och sedan undersöka funktioner mellan sådana mängder som bevarar strukturerna (eller ännu mer generellt funktorer mellan klasser, men det är det mest logiker/kategoriteoretiker som tycker om).

I exempelvis envariabelanalys så är strukturerna gränsvärden på $\mathbb{R}$ och funktioner som bevarar gränsvärden blir då kontinuerliga funktioner ty om $\lim_{n\to\infty} a_n = a$ så är f kontinuerlig om det för alla sådana gränsvärden gäller att $\lim_{n\to\infty} f(a_n) = f(a)$. I linjär algebra så är mängderna med strukturer vektorrum och funktionerna som bevarar dessa är så kallade linjära funktioner där $f(av + bw) = af(v) + bf(w)$.

Kontinuerliga funktioner i envariabelanalys är alltså sådana där det gör detsamma om man, givet en konvergerande följd $\{a_n\}_{n\in\mathbb{N}}$, först beräknar dess gränsvärde och sedan applicerar funktionen eller om man först applicerar funktionen på varje element och sedan beräknar gränsvärdet av den resulterande följden av funktionsvärden. Egenskapen som karaktäriserar *linjära* funktioner inom linjär algebra har en motsvarande egenskap; där gör det detsamma om man först adderar två vektorer och sedan applicerar funktionen på summan, eller om man först applicerar funktionen på var och en av de två vektorerna och sedan adderar resultaten.

Även för ringar finns det en motsvarande slags funktioner, som spelar en lika viktig roll som kontinuerliga funktioner gör i envariabelanalys, och linjära avbildningar gör i linjär algebra. De kallas för ringhomomorfismer.

3.1 Ringhomomorfismer

Definition 3.1. En funktion $f\colon R \to R'$ där R samt R' är ringar kallas en *ringhomomorfism* om det för alla $a, b \in R$ gäller att $f(a + b) = f(a) + f(b)$ samt $f(ab) = f(a)f(b)$ och $f(1) = 1'$ där 1 är multiplikativa enheten i R och $1'$ är multiplikativa enheten i R'.

Då vi endast tittar på ringhomomorfismer i denna kurs så kommer vi att använda begreppen homomorfism och ringhomomorfism för samma sak. I allmänhet så kan man också definiera homomorfismer också mellan andra objekt, t.ex grupper och då prata om grupphomomorfismer (vilket då definieras på liknande sätt men så att de respekterar gruppens struktur).

Anmärkning 3.2. En ringhomomorfism avbildar alltid noll på noll ty $f(a) = f(0 + a) = f(0) + f(a)$, använd första annulleringslagen och vi får $0 = f(0)$.

3.1.1 Några exempel på ringhomomorfismer

Exempel 3.3. En (ganska tråkig) ringhomorfism är identitetsfunktionen från R till R, dvs en funktion f sådan att $f(x) = x$. Den uppfyller uppenbarligen kraven.

Exempel 3.4. Låt R vara en godtycklig ring och låt $a \in R$, studera homomorfismen $v_a : R[x] \to R$ definierad så att $v_a(p) = p(a)$ för $p \in R[x]$. Detta är en ringhomomorfism eftersom givet $p, h \in R[x]$, har vi

$$\begin{aligned} v_a(p+h) &= (p+h)(a) = p(a) + h(a) = \\ &= v_a(p) + v_a(h). \\ v_a(ph) &= (ph)(a) = p(a)h(a) = \\ &= v_a(p)v_a(h). \\ v_a(1) &= 1. \end{aligned}$$

Denna ringhomomorfism motsvarar just att undersöka vad ett polynom har för värde i en given punkt $x = a$.

Exempel 3.5. Det finns också alltid en naturlig ringhomomorfism $\mathbb{Z} \to R$ för alla R, nämligen den som skickar n på $n \cdot 1$. Detta eftersom $f(n+m) = (n+m) \cdot 1 = n \cdot 1 + m \cdot 1 = f(n) + f(m)$ samt motsvarande för multiplikation.

Exempel 3.6. En funktion från $\mathbb{Z}$ till $\mathbb{Z}_n$ som skickar ett heltal på heltalets restklass (modulo n) är en homomorfism.

Exempel 3.7. Det existerar inga ringhomomorfismer från $\mathbb{Z}_2$ till $\mathbb{Z}_3$. Bevis: Antag att det finns en sådan ringhomomorfism f. Då vet vi att $f(0) = 0$ enligt ovan och $f(1) = 1$ enligt kraven på ringhomomorfism. Då måste det gälla att $0 = f(0) = f(1+1) = f(1) + f(1) = 1 + 1 = 2 \neq 0$ så vi får en motsägelse.

Exempel 3.8. Om R är en delring till R' så har vi en inklusionshomomorfism $i : R \to R'$ så att $i(x) = x$. Ett specialfall av detta är att se R som delring till $R[x]$, dvs att i helt enkelt till elementet $r \in R$ tillordnar det konstanta polynomet $r \in R[x]$.

3.2 Isomorfier

Det händer ganska ofta att man studerar objekt som är på något vis är lika matematiskt utan att objekten beskrivs på samma sätt. Ett exempel kan vara vektorer i ett plan och par av reella tal. Matematiskt sett så beskriver de samma sak även om man har två olika uttryckssätt. Ett annat exempel kan vara komplexa tal och par av reella tal (a, b) med operationerna $(a, b) + (c, d) = (a + c, b + d)$ samt $(a, b)(c, d) = (ac - bd, bc + ad)$.

Det är helt klart att det är samma objekt man diskuterar, men hur skall man uttrycka detta matematiskt? För att säga att de algebraiskt är samma objekt så använder man begreppet isomorfi som definieras som följer:

Definition 3.9. Två ringar R och R' är *isomorfa* om det existerar en ringhomomorfism $h : R \to R'$ så att h är bijektiv och att dess invers h^{-1} också är en ringhomomorfism. Denna ringhomomorfism kallas då för en isomorfi.

Anmärkning 3.10. Detta ger upphov till en ekvivalensrelation $R \sim R'$ om och endast om R är isomorf till R'. Beviset att det verkligen är en ekvivalensrelation är en nyttig övning.

Några andra nyttiga övningar kan vara att visa att om R är isomorf till R' och R har någon av egenskaperna vi diskuterat tidigare (faktoriell, är en kropp etc) så har också R' denna egenskap.

3.2.1 Några exempel på isomorfier

Exempel 3.11. Varje ring R är isomorf till sig själv genom identitetsfunktionen.

Exempel 3.12. Identitetsfunktionen måste inte vara den enda isomorfin från en ring till sig själv, exempelvis så är polynomringen $R[x, y]$ i två variabler isomorfi till sig själv genom att man byter plats på x och y.

Exempel 3.13. Alla kroppar med två element är isomorfa till $\mathbb{Z}_2$ (skicka nollan på noll och ettan på ett).

Exempel 3.14. Polynomringarna $R[x]$ och $R[y]$ är isomorfa, låt $f(p(x)) = p(y)$.

Proposition 3.15. *Om vi har en injektiv homomorfism $\varphi : R \to R'$ så är bilden $\Im(R)$ isomorf med R.*

Bevis. Funktionen φ är en bijektion från $R \to \Im(R)$. Det återstår bara att visa att $\Im(R)$ är en delring eftersom en isomorfi måste gå mellan två ringar, men den är uppenbart sluten under addition, ty om $a = \varphi(a')$ samt $b = \varphi(b')$ ligger i $\Im(R)$ så gäller att $a + b = \varphi(a') + \varphi(b') = \varphi(a' + b')$ ligger i $\Im(R)$. Resten av kraven fås på motsvarande sätt. □

Anmärkning 3.16. Om vi har en homomorfi $h : R \to R'$ som är inverterbar så är dess invers automatiskt en homomorfi.

Bevis. Antag att h är en inverterbar homomorfism med invers h'. Vi ska visa att $h'(a + b) = h'(a) + h'(b)$ samt $h'(ab) = h'(a)h'(b)$. Undersök $c = h'(a + b) - h'(a) - h'(b)$. Vi tittar på $h(c) = h(h'(a + b) - h'(a) - h'(b)) = h(h'(a + b)) - h(h'(a)) - h(h'(b)) = a + b - a - b = 0$. Då h avbildar c på noll och vi vet (då h är en homomorfi) att h avbildar 0 på 0 och h är inverterbar så måste $c = 0$ vilket ger att $h'(a + b) = h'(a) + h'(b)$. Motsvarande bevis för multiplikation lämnas till den intresserade läsaren. □

Orsaken till att vi tar med kravet att h:s invers skall vara en homomorfi i definitionen trots att det egentligen inte krävs enligt kommentaren då det alltid gäller är för att belysa likheterna med denna definition på att vara "samma" objekt med definitionerna i andra ämnen.

I exempelvis topologi så säger man att två topologiska rum X och Y är **homeomorfa** om det existerar en kontinuerlig inverterbar funktion $f : X \to Y$ sådan att dess invers också är kontinuerlig (här följer det inte att inversen är kontinuerlig, eller hur?). Samma definition dyker också upp i differentialtopologi där man istället för kontinuerlig ställer det hårdare kravet att funktionen skall vara slät (dvs oändligt många ggr deriverbar med rätt definition på deriverbar).

Man kan även prata om isomorfier mellan grupper, mängder som bara uppfyller de första tre ringaxiomen (dvs vi struntar i multiplikationen och kräver inte att addition är kommutativ) med nästan precis samma definition av isomorfi.

Styrkan med dessa definitioner är att eftersom man ser till att de bevarar de viktiga egenskaperna för ämnet i fråga (öppna mängder för topologiska rum, gruppoperationen för grupper etc) så bevarar den egenskaper som uttrycks i dessa termer. I vårt fall så är det exempelvis egenskapen att vara en kropp eller att vara faktoriell, medans i exempelvis topologin så kan det vara egenskaper som att sitta ihop eller att lokalt se ut på ett visst sätt.

4 Ideal och kvotringar

4.1 Ideal

Precis i början av kursen när vi studerade talteori tittade vi på alla tal som gav en given rest vid division med talet n och satte dem i samma ekvivalensklass. Ifall vi speciellt studerar de talen som ger resten noll (dvs de som är delbara med n) så har de flera intressanta egenskaper. De är slutna under multiplikation och addition och även om man multiplicerar dem med något tal som inte är delbart med n så får man fortfarande ett tal delbart med n (de bildar dock ingen delring i allmänhet, eller hur?).

Vi kan också observera samma sak med exempelvis kontinuerliga funktioner, om vi tittar på funktionerna $f(x)$ så att $f(r) = 0$ för något givet $r \in \mathbb{R}$ så har de också dessa egenskaper, om vi adderar två sådana funktioner så får vi en ny sådan funktion, och om vi multipicerar dem med en godtycklig kontinuerlig funktion så får vi en ny funktion som är noll i r.

Delmängder av ringar med dessa egenskaper kallas ideal och definieras som följer.

Definition 4.1. Låt I vara en icke-tom delmängd av en ring R. Vi säger att I är ett *ideal* i R ifall $a, b \in I, r \in R$ medför att $a + b, ra \in I$.

För att en icke-tom delmängd av en ring ska få kallas ett ideal, måste den alltså uppfylla två villkor. För det första måste summan av två element från I ligga i I. För det andra måste ra vara ett element i I för varje val av $r \in R$ och $a \in I$. Man ska alltså kunna addera två element från idealet, samt kunna multiplicera element från idealet med godtyckliga ringelement, utan att lämna idealet.

Anmärkning 4.2. Om I är ett ideal så ligger 0 i I. Detta är uppenbart ty då I är icketom så har vi något $x \in I$, men då har vi även $0x = 0 \in I$.

Anmärkning 4.3. R och $\{0\}$ är alltid ideal (uppenbart).

Då dessa ideal alltid finns i alla ringar så är de samtidigt inte så intressanta som andra ideal.

Definition 4.4. Idealen R och $\{0\}$ kallas för *triviala ideal*.

Proposition 4.5. *Om φ är en homomorfism från $R \to R'$ så är $Ker(\varphi)$ ett ideal, där $ker(\varphi)$ består av alla element x sådana att $\varphi(x) = 0$.*

Bevis. Elementet 0 ligger uppenbarligen i $ker(\varphi)$. Om $f(a) = 0 = f(b)$ så gäller att $f(a + b) = f(a) + f(b) = 0 + 0 = 0$. Om $r \in R$ samt $a \in ker(\varphi)$ så är $f(ra) = f(r)f(a) = f(r) \cdot 0 = 0$. □

Anmärkning 4.6. Om $I \subset R$ är ett ideal sådant att $1 \in I$ så är $I = R$, ty $1 \in I \longrightarrow r \cdot 1 = r \in I$ för alla $r \in R$.

Anmärkning 4.7. De enda idealen som ligger i kroppar är de triviala idealen.

Bevis. Om $I = \{0\}$ så är vi klara, annars så kan vi ta $x \in I \backslash \{0\}$, då K är en kropp så kan vi hitta ett tal $y \in K$ sådant att $yx = 1$. Då följer det att $1 \in I$ och alltså är $I = R$ enligt ovan. □

Proposition 4.8. *Låt I samt I' vara ideal i ringen R. Då är deras snitt, $I \cap I'$ ett ideal.*

Bevis. Låt I och I' vara ideal som i satsen. Låt $\hat{I} = I \cap I'$. Då 0 ligger i alla ideal så ligger 0 i både I och I' och alltså även i $\hat{I}$ vilket gör att $\hat{I}$ klarar det första kravet, icketomhet. Antag att a, b ligger i $\hat{I}$. Då måste a, b ligga i I och i I'. Alltså så måste $a + b$ och ca ligga i både I och I' då de är ideal. Men då måste $a + b$ och ca ligga i deras snitt $\hat{I}$. □

Definition 4.9. För element $a_1, a_2, a_3, \ldots, a_n \in R$ så definierar vi $\langle a_1, a_2, a_3, \ldots, a_n \rangle$ som det minsta idealet i R som innehåller $a_1, a_2, a_3, \ldots, a_n$. Ett annat (ekvivalent) sätt att definiera detta ideal är följande

$$\langle a_1, a_2, a_3, \ldots, a_n \rangle = Ra_1 + Ra_2 + \ldots + Ra_n,$$

där

$$Ra_1 + Ra_2 + \ldots + Ra_n = \{r_1a_1 + r_2a_2 + \ldots + r_na_n \mid r_1, \ldots, r_n \in R\}.$$

Vi säger att ett ideal $I \subset R$ är *ändligt genererat* om $I = \langle a_1, \ldots, a_n \rangle$ för några väl valda element $a_1, \ldots, a_n \in R$.

Unionen av två ideal behöver inte nödvändigtvis vara ett ideal. Ett enkelt motexempel är att ta de heltal som är delbara med 2 samt de heltal som är delbara med 3 sedda som ideal i $\mathbb{Z}$. Deras union innehåller precis de tal som är delbara med två eller tre. Då ligger 3 och -2 i unionen, om det vore ett ideal så skulle $3 + (-2) = 1$ ligga i unionen, men 1 är varken delbart med två eller tre.

Definition 4.10. Vi säger att I är ett *äkta ideal* till R om I är ett ideal och $I \neq R$.

Ifall vi igen betraktar heltalen och tar som (äkta) ideal alla tal delbara med 4 så är det ganska klart att vi kan hitta ett större äkta ideal, nämligen alla tal delbara med 2. Därefter så tar det dock stopp, ty om vi skulle försöka göra ett större ideal så måste vi ta med ett udda tal. Då kommer alltså både $2k + 1$ och $-2k$ att ligga i vårt ideal, alltså så kommer deras summa att göra det och vi får med ett i idealet. Detta ger nu enligt kommentar 4.6 att vi har hela ringen som vårt ideal och då är det ej längre äkta. Idealet som består av alla jämna tal är alltså ett maximalt äkta ideal vilket formaliseras som följer.

Definition 4.11. Vi säger att $I \subset R$ är ett *maximalt ideal* om I är ett äkta ideal och om det inte ligger i något strikt större äkta ideal.

Om I är ett maximalt ideal och $I \subset J$ där J är ett ideal, så följer det alltså att J måste vara lika med endera R eller I.

Notera att maximala ideal ej behöver vara unika, i ringen av heltal så är exempelvis både $\langle 2 \rangle$ och $\langle 3 \rangle$ maximala ideal.

Vi noterar snabbt att idealet bestående av alla heltal delbara med n är just idealet $\langle n \rangle$, bestående just av de talen som kan skrivas som an för $a \in \mathbb{Z}$. En naturlig fråga är då ifall alla ideal i alla ringar kan genereras på detta sätt av ett enda element. Även om resultatet stämmer i ringen $\mathbb{Z}$, så är svaret tyvärr nej i allmänhet. Ideal som kan genereras av ett enda element är tillräckligt intressanta för att förtjäna ett eget namn, de kallas för huvudideal.

Definition 4.12. Vi definierar mängden Ri där $i \in R$ som alla element i R som kan skrivas på formen ai där $a \in R$ och kallar detta för huvudidealet genererat av i.

Definition 4.13. Vi säger att ett ideal $I \subset R$ är ett *huvudideal* om $I = Ri$ för något $i \in I$.

Några exempel på huvudideal är $R = R1$, $0 = R0$, polynom med en rot i 1 (dvs $R[X](X-1)$) och jämna tal $\mathbb{Z}(-2)$. Observera att det inte är entydigt bestämt vad generatorn är då $\mathbb{Z}2 = \mathbb{Z}(-2)$.

Alla ideal är inte huvudideal, ett exempel på ett ideal som inte är ett huvudideal är alla kontinuerliga funktioner $f : \mathbb{R} \to \mathbb{R}$ så att $f(0) = 0$, beviset lämnas som en intressant övning.

Just ringen $\mathbb{Z}$ har dock egenskapen att alla ideal är huvudideal. Även polynomringen $\mathbb{C}[X]$ har denna egenskap. Sådana ringar kallas huvudidealringar vilket vi formaliserar som följer.

Definition 4.14. Ett integritetsområde R är en *huvudidealring* om alla ideal $I \subset R$ är huvudideal.

Några exempel på huvudidealringar är kroppar, $\mathbb{Z}$ samt polynom med koefficienter i kroppar (visas i avsnittet om polynom).

Förra kapitlet så diskuterades faktoriella ringar. Dessa hör ihop med huvudidealringar på följande vis.

Låt R vara en ring. Vi kallar en familj av ideal $\{I_j\}_{j=1}^{\infty}$ i R för en växande kedja av ideal om $I_j \subset I_{j+1}$ för $j = 1, 2, \ldots$. En växande kedja

$$I_1 \subset I_2 \subset I_3 \subset \ldots$$

sägs stabilisera sig vid $j = n$ om det finns ett tal $n \in \mathbb{N}$ sådant att

$$I_n = I_{n+1} = I_{n+2} = \ldots,$$

eller med andra ord om $I_n = I_{n+k}$ för alla $k \in \mathbb{N}$.

Lemma 4.15. *Varje växande kedja i en huvudidealring stabiliserar sig förr eller senare.*

Bevis. Låt

$$I_1 \subset I_2 \subset I_3 \subset \dots$$

vara en växande kedja av ideal. Vi påstår att unionen

$$I = \bigcup_{j \in \mathbb{N}} I_j$$

bildar ett ideal i R. För att visa det behöver vi visa att $a, b \in I \longrightarrow a + b \in I$ och $a \in I, r \in R \longrightarrow ra \in I$. Om $a, b \in I$, så finns det index $k, l \in \mathbb{N}$ sådana att $a \in I_k$ och $b \in I_l$; låt $m = \max\{k, l\}$. Då har vi, eftersom kedjan av ideal är växande, att a och b båda ligger i I_m. Men I_m är ett ideal, så $a + b \in I_m$, och därmed ligger $a + b$ även i I eftersom $I_m \subset I$. Beviset för att $a \in I, r \in R \longrightarrow ra \in I$ görs med ett liknande resonemang och lämnas som övning. På grund av antagandet i lemmat, måste I vara ett huvudideal, så vi har $I = Rg$ för något $g \in R$. Naturligtvis har vi $g \in I$, och därför måste g ligga i I_n för något $n \in \mathbb{N}$. Men det medför att kedjan stabiliseras vid $j = n$. Antag nämligen att $a \in I_{n+k} \subset I$ för något $k \in \mathbb{N}$. Eftersom $I = Rg$, så har vi $a = rg$ för något $r \in R$, men då ligger ju a i I_n, så $I_n = I_{n+k}$. □

Lemma 4.16. *I en huvudidealring R är varje irreducibelt element ett primelement.*

Bevis. Antag att $p \in R$ är ett irreducibelt element, och att $p|bc$ där $b, c \in R$. Vi behöver visa att p i så fall delar endera b eller c. Vi bildar idealet $I = \langle p, b \rangle$, och eftersom R är ett huvudidealområde har vi $I = \langle d \rangle$ för något väl valt $d \in R$. Det följer att $p = rd$ för något $r \in R$ eftersom $p \in I$, men då p är irreducibelt måste endera r eller d vara inverterbar. Vi undersöker de två möjliga fallen var för sig.

Antag först att d är inverterbar. Vi påstår att detta är ekvivalent med att $\langle d \rangle = R$: Den ena inklusionen, $\langle d \rangle \subset R$ är klar, och den andra inklusionen, $R \subset \langle d \rangle$, följer av att $r = erd \in \langle d \rangle$ för ett godtyckligt element $r \in R$, där e är inversen till d. Alltså har vi $\langle b, p \rangle = R$, så det finns element $s, t \in R$ sådana att $ps + bt = 1$. Vi multiplicerar denna relation med c för att få

$$c = psc + bct.$$

Eftersom p delar varje term i högerledet, måste p även dela c.

Antag då istället att det är r som är inverterbar, med invers r'. Vi multiplicerar relationen $p = rd$ med r' och erhåller $pr' = d$. Eftersom $b \in \langle b, p \rangle = \langle d \rangle$ så har vi $b = xd$ för något $x \in R$, och det medför att $b = xpr'$. Så p delar b.

Sammanfattningsvis har vi visat att p delar endera b eller c, dvs att p är ett primelement. □

Lemma 4.17. *Låt R vara en huvudidealring. Varje nollskilt element $a \in R$ som inte är inverterbart kan skrivas som en produkt av ändligt många irreducibla element.*

Bevis. Vi börjar med att visa att a har åtminstone en irreducibel faktor. Om a själv är irreducibel så stämmer förstås detta, och i annat fall kan vi skriva $a = a_1 b_1$ där varken a_1 eller b_1 är inverterbar. Då har vi $a \in \langle a_1 \rangle$, och

$$\langle a \rangle \subsetneq \langle a_1 \rangle.$$

Inklusionen är strikt eftersom $\langle a \rangle = \langle a_1 \rangle$ skulle innebära att $a_1 = ac$ och $a = acb_1$ för något $c \in R$. Men det skulle i sin tur innebära, eftersom R är ett integritetsområde, att b_1 är inverterbar. Om a_1 inte är irreducibel kan vi skriva $a_1 = a_2 b_2$, där varken a_2 eller b_2 är inverterbara, och då erhåller vi på samma sätt

$$\langle a \rangle \subsetneq \langle a_1 \rangle \subsetneq \langle a_2 \rangle.$$

Om nu inte a_2 skulle vara irreducibel skriver vi $a_2 = a_3 b_3$ med icke-inverterbara element a_3, b_3 och erhåller en ännu längre kedja

$$\langle a \rangle \subsetneq \langle a_1 \rangle \subsetneq \langle a_2 \rangle \subsetneq \langle a_3 \rangle.$$

Vi inser att om inget av elementen a_i, $i = 1, 2, 3, \ldots$ konstruerade på detta vis är irreducibelt så kan kedjan av växande ideal förlängas obegränsat. Men eftersom R är en huvudidealring kan vi tillämpa lemma 4.15 och dra slutsatsen att detta är omöjligt. Alltså måste kedjan stanna vid ett ideal $\langle a_r \rangle$ som är genererat av en irreducibel faktor till a. Det visar att a har en irreducibel faktor.

Låt $a \in R$ vara ett nollskilt icke-inverterbart element. Om a inte är irreducibel kan vi enligt bevisets första del skriva $a = p_1 c_1$ där p_1 är irreducibel och c_1 inte är inverterbar. Precis som i bevisets första del får vi då $\langle a \rangle \subsetneq \langle c_1 \rangle$. Om nu inte c_1 är irreducibel så kan vi skriva $c_1 = p_2 c_2$ där p_2 är irreducibel och c_2 inte är inverterbar. Genom att fortsätta den här proceduren får vi, om inget av de erhållna elementen c_r är irreducibelt, en strikt växande kedja av ideal

$$\langle a \rangle \subsetneq \langle c_1 \rangle \subsetneq \langle c_2 \rangle \subsetneq \langle c_3 \rangle \ldots,$$

men enligt lemma 4.15 så är detta inte möjligt. Alltså kommer vi förr eller senare att erhålla ett irreducibelt element c_r, och det innebär att

$$p = p_1 p_2 p_3 \ldots p_r c_r$$

är en produkt av ändligt många irreducibla element. □

Sats 4.18. *Varje huvudidealring R är faktoriell.*

Bevis. Vi behöver visa att varje nollskilt och icke-inverterbart element $a \in R$ kan skrivas som en ändlig produkt av irreducibla element, och att denna faktorisering är unik så när som på associering och faktorernas ordning. Enligt

lemma 4.17 kan a skrivas som en produkt av ändligt många irreducibla element:

$$a = p_1 p_2 \dots p_k,$$

och det enda som återstår att visa är att entydighetsdelen av satsen. Antag därför att vi även har $a = q_1 q_2 \dots q_l$, så att

$$p_1 p_2 \dots p_k = q_1 q_2 \dots q_l.$$

Vi kan anta att $k \leq l$. Enligt lemma 4.16 så är p_1 ett primelement, och måste därför dela en av faktorerna i högerledet. Genom att omindexera dem kan vi anta att $p_1 | q_1$. Det medför, eftersom q_1 är irreducibelt, att $p_1 = eq_1$ för något inverterbart element e. Men då kan vi förkorta med p_1 och erhålla

$$p_2 \dots p_k = (e^{-1} q_2) q_3 \dots q_l.$$

Resultatet följer nu genom induktion över k. Eftersom $e^{-1} q_2$ är irreducibelt om och endast om q_2 är det, så är vi nämligen tillbaka i den ursprungliga situationen, men med en irreducibel faktor mindre på varje sida. □

Då $\mathbb{Z}$ är ett huvudidealring så är den faktoriell.

Per definition är ett primelement p ett element sådant att $p|ab$ medför att p delar endera a eller b. Det finns en motsvarande egenskap för ideal, nämligen att vara ett primideal.

Definition 4.19. Ett äkta ideal I kallas för *primideal* om $ab \in I$ medför att a ligger i I eller b ligger i I.

Primideal hör ihop med primelement på följande sätt:

Proposition 4.20. *Ett huvudideal Rp i ett integritetsområde R är ett primideal om och endast om p är ett primelement.*

Bevis. Vi visar först att p primt medför att Rp är ett primideal. Antag att $ab \in Rp$. Det innebär att $ab = rp$ för något $r \in R$. Då $p|ab$ så medför det att p delar antingen a eller b. Antag att $p|a$, då är $a = pc$ och alltså ligger a i Rp och Rp är ett primideal. Vi visar andra hållet, antag att Rp är ett primideal och antag att $p|ab$. Då ligger ab i Rp och alltså ligger antingen a eller b i Rp. Vi antar att a ligger i Rp. Detta ger att $a = cp$ för något c, alltså $p|a$, vilket ger att p är ett primelement. □

Proposition 4.21. *I $\mathbb{Z}$ så är alla primideal maximalideal.*

I allmänhet så gäller inte detta, ett motexempel kan vara ringen $Q[x, y]$, dvs polynom i två variabler. Där är idealet genererat av x ett primideal men ej ett maximalideal. Det andra hållet gäller däremot.

Proposition 4.22. *Om $I \subset R$ är ett maximalideal så är I ett primideal.*

Bevis. Övning. □

4.2 Kvotringar

Då vi tidigare i talteorin räknade modulo n så sade vi att $a = b$ om n delade $a - b$. Uttryckt i vårt nya idealspråk så skulle vi kanske säga att $a = b$ om $a - b$ låg i idealet som genererades av n. Denna idé ger upphov till kvotringar, kursens kanske svåraste och mest intressanta begrepp.

Definition 4.23. Om $I \subset R$ är ett äkta ideal så definierar vi en ring R/I som följer:

- R/I består av ekvivalensklasser $\bar{a}$ av element från R med $\bar{a} = \bar{b}$ om och endast om $(a - b) \in I$ (elementen $\bar{a}$ kan även skrivas som $[a]$).
- $\overline{a} + \overline{b} = \overline{a + b}$
- $\bar{a} \cdot \bar{b} = \overline{a \cdot b}$

För att övertyga sig om att detta är en ring, behöver man visa att de två operationerna är väldefinierade, dvs att de inte beror på val av representanter för elementen i fråga, samt att alla ringaxiomen är uppfyllda. Vi lämnar detta som övning, och kommer att ta upp det på någon av lektionerna.

Ett annat vanligt skrivsätt för elementet $\overline{a} \in R/I$ är $a + I$.

Sats 4.24 (Noethers (första) isomorfisats)**.** *Låt φ vara en homomorfism från R till R'. Då är $Im(\varphi)$ isomorf till $R/Ker(\varphi)$.*

Bevis. Vi skapar isomorfin som följer, låt $\bar{x}$ vara ett element i $R/Ker(\varphi)$. Vi definierar $f(\bar{x}) = \varphi(x)$. Detta är väldefinierat ty om $\bar{x} = \bar{y}$ så är $y - x \in Ker(\varphi)$. Då gäller att

$$\begin{aligned} f(\bar{x}) &= \varphi(x) = \varphi(x) + 0 = \varphi(x) + \varphi(y - x) = \varphi(x + y - x) = \varphi(y) = \\ &= f(\bar{y}) \end{aligned}$$

Alltså är f väldefinierad. Då funktionen φ är en homomorfism så är f en homomorfism. Den är injektiv ty

$$f(\bar{x}) = f(\bar{y}) \longrightarrow f(\bar{x}) - f(\bar{y}) = 0 \longrightarrow 0 = f(\bar{x} - \bar{y}) = \varphi(x - y) \longrightarrow \bar{x} = \bar{y}.$$

Den är uppenbart surjektiv ty om $x \in Im(\varphi)$ så existerar $y \in R$ så att $\varphi(y) = x$. Då gäller att $f(\bar{y}) = x$. Då vi har en bijektiv homomorfism så har vi en isomorfism. □

Några exempel på kvotringar är som följer:

Exempel 4.25. $\mathbb{Z}/\mathbb{Z}p = \mathbb{Z}_p$.

Exempel 4.26. $R/\{0\}$ är isomorf till R.

Exempel 4.27. $\mathbb{R}[x]/\langle x^2 + 1\rangle$ är isomorf till de komplexa talen $\mathbb{C}$.

Lemma 4.28. *Låt R vara en ring och $\mathfrak{a} \subset R$ ett äkta ideal. Då finns det en 1-1 motsvarighet mellan mängden av ideal i R som innehåller $\mathfrak{a}$ och mängden av alla ideal i $R/\mathfrak{a}$: Vi associerar till ett ideal $\mathfrak{b} \supset \mathfrak{a}$ i R idealet $\mathfrak{b}/\mathfrak{a} \subset R/\mathfrak{a}$, och till idealet $\mathfrak{c} \subset R/\mathfrak{a}$ associerar vi idealet $\{r \in R \mid r + \mathfrak{a} \subset \mathfrak{c}\}$ i R. Här betyder $\mathfrak{b}/\mathfrak{a}$ idealet i $R/\mathfrak{a}$ som består av element på formen $b + \mathfrak{a}$, med $b \in \mathfrak{b}$.*

Bevis. Beviset kommer att ges som övningsuppgift, och gås igenom på en lektion. Det kräver inte någon speciellt trick, eller särskild idé, utan är ganska rätt fram. □

Ett ekvivalent sätt att säga detta på är som följer. Låt R vara en ring och $\mathfrak{a} \subset R$ ett ideal. Det finns en naturlig homomorfism $\pi : R \longrightarrow R/\mathfrak{a}$ som till $r \in R$ associerar $r + \mathfrak{a} \in R/\mathfrak{a}$; den kallas vanligen för den kanoniska kvotprojektionen, och är förstås surjektiv. Nu kan 1-1-korrespondensen beskrivas med hjälp av π på följande sätt.

$$\begin{array}{ccc} \{\text{Ideal i } R \text{ som innehåller } \mathfrak{a}\} & \longrightarrow & \{\text{ideal i } R/\mathfrak{a}\} \\ \mathfrak{b} & \mapsto & \pi(\mathfrak{b}) \end{array}$$

med invers

$$\begin{array}{ccc} \{\text{ideal i } R/\mathfrak{a}\} & \longrightarrow & \{\text{Ideal i } R \text{ som innehåller } \mathfrak{a}\} \\ \mathfrak{c} & \mapsto & \pi^{-1}(\mathfrak{c}). \end{array}$$

En naturlig fråga är nu om man kan lista ut några egenskaper av kvotringen R/I utifrån egenskaper hos R och I. Även här kan vi dra vissa paralleller till talteorin. Om man räknar modulo n så är alla (nollskilda) element inverterbara om och endast om n är ett primtal, vilket är ekvivalent med att $\mathbb{Z}n$ är ett maximalideal. Detta innebär att Z_p är en kropp. Detta följer av följande teorem om kvotringar.

Proposition 4.29. *Ett ideal $M \subset R$ är maximalt om och endast om R/M är en kropp.*

Bevis. Låt M vara ett maximalt ideal; vi gör ett motsägelsebevis för att R/M är en kropp. Antag därför att R/M inte är en kropp; då finns det ett nollskilt element $\overline{a} \in R/M$ som inte är inverterbart. Studera då idealet $M + Ra = \{m + ra \mid r \in R, m \in M\}$ i R. Det innehåller M strikt eftersom M inte innehåller elementet a (att $\overline{a}$ är nollskilt i R/M betyder precis att $a \notin M$). Men M är maximalt, så vi får $M + Ra = R$. Speciellt så gäller det då att $1 = m + ra$ för något $r \in R$ och $m \in M$. Men då är ju $\overline{ra} = \overline{1}$ vilket innebär att $\overline{a}$ är inverterbart. Från denna motsägelsen drar vi slutsatsen att R/M är en kropp.

Antag då istället att R/M är en kropp. En ring är en kropp om och endast om $\{0\}$ är det enda äkta idealet, så enligt antagandet finns det inte några nollskilda äkta ideal i R/M. Men enligt 1-1 korrespondensen mellan ideal i R som innehåller M och idealen i R/M, betyder detta precis att det inte finns några äkta ideal som innehåller M, så M är maximalt. □

Man kan bygga nya kroppar genom att hitta maximalideal och kvota med dem, några tidigare exempel var maximalidealen $\mathbb{Z}p$ i $\mathbb{Z}$ som vi kvotade med för att få tag i idealet $\mathbb{Z}_p$. Ett annat trevligt sätt att hitta maximala ideal är genom att ta irreducibla element i ett huvudidealring och sedan skapa huvudidealet genererat av detta irreducibla element. Detta kommer då att bli ett maximalideal (inte så svårt att bevisa). Ett exempel kan vara att kvota med det irreducibla polynomet $x^3 - 2$ i polynomringen $\mathbb{Q}[x]$ för att få en kropp.

Alla primideal i $\mathbb{Z}$ är maximalideal och tvärsom så man kan också fråga sig vad som händer i allmänhet om man kvotar med ett primideal. Då vi vet att att egenskapen att vara primideal är något svagare än att vara maximalideal (då alla maximalideal är primideal men ej tvärsom) så kan man misstänka att vi borde få en kvotring med en egenskap som kroppar har, men som är svagare än att vara en kropp.

Proposition 4.30. *Ett ideal $M \subset R$ är ett primideal om och endast om R/M är ett integritetsområde.*

Bevis. Låt M vara ett primideal som i satsen och antag att $(a + M)(b + M) = 0$ för något par $a + M, b + M \in R/M$. Då har vi alltså att $(a + M)(b + M) = ab + M = 0 \in R/M$. Det är detsamma som att säga att $ab \in M$, och eftersom M är ett primideal kan vi dra slutsatsen att endera a eller b ligger i M. Men det är ekvivalent med att säga att endera $a+M$ eller $b+M$ är lika med 0 i R/M. Alltså saknar R/M nolldelare, vilket är detsamma som att vara ett integritetsområde.

Antag då istället att R/M är ett integritetsområde, och låt $a, b \in R$ vara element vars produkt ligger i M. Vi behöver visa att endera a eller b ligger i M. Vi ser att

$$(a + M)(b + M) = ab + M = 0 + M \in R/M,$$

där vi i sista steget använder att $ab \in M$. Det betyder att produkten $(a+M)(b+M) = 0$ i R/M. Men R/M är ett integritetsområde, så endera $a + M$ eller $b + M$ måste vara 0 där, och det betyder precis att endera a eller b måste ligga i M. □

5 Euklidiska ringar.

Vi har tidigare visat att $\mathbb{Z}$ är en faktoriell ring, dvs att varje heltal kan skrivas som en produkt av primelement och att en sådan faktorisering är unik så när som på faktorernas ordning och associering. Samma sak visade sig gälla även för polynomringar $K[X]$ över en kropp K [1]. I båda fallen hade divisionsalgoritmen en avgörande roll under bevisets gång. Med hjälp av den kunde vi nämligen visa att varje ideal i såväl $\mathbb{Z}$ som $K[X]$ kan genereras av ett enda element, dvs att de är huvudidealringar, och sedan tillämpade vi sats 4.18. Vi tar

[1] dvs om $P = p_1 \dots p_r = q_1 \dots q_s$ är två faktoriseringar av P i irreducibla polynom, så är $r = s$, och det finns en bijektion $\sigma : \{1, \dots, r\} \longrightarrow \{1, \dots, r\}$ sådan att p_i och $q_{\sigma(i)}$ är associerade för $i = 1, \dots, r$

vara på den observationen och inför ett namn på de ringar där vi har tillgång till en divisionsalgoritm, och på samma sätt kan dra slutsatsen att ringen i fråga är en huvudidealring.

Definition 5.1. Ett integritetsområde R kallas för *euklidisk* om det finns en funktion $N : R \setminus \{0\} \longrightarrow \mathbb{N}$ sådan att

1. $N(a) \leq N(ab)$ för alla $a, b \in R \setminus \{0\}$.
2. För varje val av $a, b \in R \setminus \{0\}$ finns det element $q, r \in R$ sådana att $a = bq + r$, och antingen $r = 0$, eller $N(r) < N(b)$.

En funktion med dessa egenskaper kallas för en euklidisk norm på R.

En euklidisk norm antar värden i $\mathbb{N}$ som inte är en ring, så i synnerhet är en euklidisk norm inte en ringhomomorfism.

Exempel 5.2. Heltalen bildar en euklidisk ring, med $N : \mathbb{Z} \setminus \{0\} \longrightarrow \mathbb{N}, N(a) = |a|$. Villkor (1) är uppfyllt eftersom $N(ab) = |ab| = |a||b| \geq |a| = N(a)$ för alla $a, b \in \mathbb{Z} \setminus \{0\}$, och villkor (2) är uppfyllt på grund av divisionsalgoritmen för heltal.

Exempel 5.3. Polynomringar $K[X]$ över en kropp K är euklidiska ringar eftersom man kan definiera en euklidisk norm $N : K[X] \setminus \{0\} \longrightarrow \mathbb{N}$ genom

$$N(P(X)) = (P : \text{s grad}).$$

Vi lämnar som övning att kontrollera att detta verkligen är en euklidisk norm.

I euklidiska ringar finns det ett enkelt villkor för att avgöra om element är inverterbara eller inte.

Lemma 5.4. *Låt R vara en euklidisk ring, med tillhörande norm N. Ett nollskilt element $a \in R$ är inverterbart om och endast om $N(a) = N(1)$.*

Bevis. Antag först att a är inverbart. Det betyder att det finns ett element $b \in R$ sådant att $ab = 1$. Eftersom N är en euklidisk norm vet vi å ena sidan att $N(1) \leq N(1 \cdot a) = N(a)$, men å andra sidan så är $N(a) \leq N(ab) = N(1)$, så vi drar slutsatsen att $N(a) = N(1)$.

För att visa den omvända implikationen, antar vi istället att $N(a) = N(1)$. Eftersom N är en euklidisk norm, finns det element $q, r \in R$ sådana att $1 = aq + r$, och antingen $r = 0$ eller $N(r) < N(1)$. Vi kan utesluta $N(r) < N(1)$ eftersom det inte kan gälla samtidigt som $N(1) \leq N(1 \cdot r) = N(r)$. Då återstår bara möjligheten att $r = 0$, vilket betyder precis att a är inverterbart, eftersom $1 = aq$. □

Sats 5.5. *Varje euklidisk ring är en huvudidealring.*

Under bevisets gång ska vi använda att om $N : R \setminus \{0\} \longrightarrow \mathbb{N}$ är en euklidisk norm, och $I \subset R$ ett äkta nollskilt ideal, så finns det ett element $a \in I$ (inte nödvändigtvis unikt) med *minimal* norm bland de nollskilda elementen i I, dvs ett element $a \in I$ sådant att $N(r) \geq N(a)$ för varje nollskilt $r \in I$. För att övertyga sig om existensen av ett sådant element kan man konstatera att varje icke-tom delmängd av $\mathbb{N}$ har ett minsta element, så speciellt har bilden av $I \setminus \{0\}$ under N, alltså $N(I \setminus \{0\}) = \{m \in \mathbb{N} \mid \exists r \in I \setminus \{0\} : N(r) = m\}$ ett minsta element.

Bevis. Låt R vara en euklidisk ring, och låt $I \subset R$ vara ett ideal. Det är klart att nollidealet är ett huvudideal, så vi kan anta att $I \neq \{0\}$. Låt $b \in I \setminus \{0\}$ vara ett element med minimal norm bland de nollskilda elementen i I, och låt $a \in I$. Eftersom R är euklidisk, kan vi skriva $a = bq + r$, där endera $r = 0$ eller $N(r) < N(b)$. Då a och b ligger i I, och I är ett ideal, ser vi att även $r = a - bq$ tillhör I. Men då kan vi utesluta $N(r) < N(b)$ eftersom b har minimal norm bland elementen i I, och då återstår bara alternativet $r = 0$. Det innebär att det godtyckliga elementet $a \in I$ kan skrivas som $a = bq$, så $I = Rb$ är ett huvudideal. □

Sats 5.5 kan formuleras lite mer precist: Låt I vara ett äkta ideal i en euklidisk ring R, och låt $a \in I$ vara ett element sådant att $N(a)$ är minimal bland de nollskilda elementen i I. Då har vi $I = Ra$.

5.1 Egenskaper hos ringen $\mathbb{Z}[i]$ av Gaussiska heltal.

Det är enkelt att övertyga sig om att delmängden $\{a + bi \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$ av $\mathbb{C}$ har strukturen av en ring; det räcker ju att kontrollera att den är additivt och multiplikativt sluten. Den betecknas vanligen $\mathbb{Z}[i]$, och kallas för ringen av *Gaussiska heltal*. Vi ska visa att $\mathbb{Z}[i]$ är euklidisk, och sedan använda sats 5.5 samt sats 4.18 för att dra slutsatsen att den är faktoriell. Sedan undersöker vi vilka dess primelement är, och ger några exempel på hur man kan primfaktorisera ett givet gaussiskt heltal.

Eftersom $\mathbb{C}$ är ett integritetsområde måste även $\mathbb{Z}[i]$, som är en delring av $\mathbb{C}$, vara ett integritetsområde (villkoren är ju att $1 \neq 0 \in \mathbb{Z}[i]$, samt att $\mathbb{Z}[i]$ ska vara kommutativ och sakna nolldelare; var och en av dessa följer från motsvarande egenskap i $\mathbb{C}$). Följande sats visar att det finns en euklidisk norm på $\mathbb{Z}[i]$, och att ringen av Gaussiska heltal således är euklidisk.

Proposition 5.6. *Funktionen $N : \mathbb{Z}[i] \setminus \{0\} \longrightarrow \mathbb{N}$, $N(a + bi) = |a + bi|^2$ är en euklidisk norm på de Gaussiska heltalen.*

Lägg märke till att N är multiplikativ: $N(zw) = N(z)N(w)$ för alla $z, w \in \mathbb{Z}[i]$. Detta är dock inte en allmän egenskap för euklidiska normer; till exempel har vi ju sett att

$$N(P) = (P : \text{s grad})$$

definierar en euklidisk norm på polynomringen $K[X]$ där K är en kropp, men den är inte multiplikativ eftersom $N(PQ) \neq N(P)N(Q)$ i allmänhet.

Bevis. Krav (1) för en euklidisk norm är uppfyllt av N, eftersom vi har

$$N(zw) = |zw|^2 = |z|^2|w|^2 \geq |z|^2 = N(z)$$

för varje val av element $z, w \in \mathbb{Z}[i] \setminus \{0\}$.

Vi undersöker krav (2): Låt $z, t \in \mathbb{Z}[i] \setminus \{0\}$; vi behöver hitta element $q, r \in \mathbb{Z}[i]$ sådana att $z = qt + r$, där endera $N(r) < N(t)$ eller $r = 0$. Låt x, y vara real-, respektive imaginärdelen av z/t, så att $z/t = x + iy$; låt vidare $q = a+ib \in \mathbb{Z}[i]$, där a, b är heltal sådana att $|a-x| \leq 1/2$, och $|b-y| \leq 1/2$. Med detta val av a och b har vi att $|\frac{z}{t} - q| \leq \frac{\sqrt{2}}{2} < 1$. (Vi har helt enkelt approximerat $z/t \in \mathbb{C}$ med ett gaussiskt heltal q på minimalt avstånd från z/t, och konstaterat att detta avstånd inte under några omständigheter överstiger $\sqrt{2}/2$).

Vi låter

$$r = z - qt,$$

och noterar att $r \in \mathbb{Z}[i]$, eftersom var och en av z, q, och t ligger i $\mathbb{Z}[i]$. Då har vi förstås att $z = qt + r$, och det återstår bara att visa att endera $r = 0$ eller $N(r) < N(t)$. Ifall $r = 0$ är vi klara, och ifall $r \neq 0$ skriver vi:

$$r = t(\frac{z}{t} - q).$$

Eftersom $|\frac{z}{t} - q| < 1$ får vi att

$$|r| = |t||\frac{z}{t} - q| < |t|.$$

Det följer att $|r|^2 < |t|^2$, vilket betyder precis att $N(r) < N(t)$, och därmed har vi visat att N uppfyller även krav (2). □

Med hjälp av satserna 5.5 och sats 4.18, kan vi alltså utan ytterligare ansträngning dra slutsatsen att $\mathbb{Z}[i]$ är faktoriell. Då blir vi förstås nyfikna och ställer oss genast följdfrågan: Vilka av de gaussiska heltalen är primelement, och hur går primfaktoriseringen av ett givet gaussiskt heltal till? Sats 5.9 karaktäriserar primelementen i $\mathbb{Z}[i]$, men innan vi kan bevisa den behöver vi ett antal hjälpresultat:

Lemma 5.7. *De inverterbara gaussiska heltalen är precis* $\{\pm 1, \pm i\}$.

Bevis. Enligt 5.4 är $z = a + bi \in \mathbb{Z}[i]$ inverterbart om och endast om $N(z) = N(1)$, dvs om och endast om $a^2 + b^2 = |1|^2 = 1$. Men om $a, b \in \mathbb{Z}$ uppfyller $a^2 + b^2 = 1$ så måste antingen a eller b vara noll, och den andra måste vara ± 1; alltså är de inverterbara elementen i $\mathbb{Z}[i]$ precis $\{\pm 1, \pm i\}$. □

Sats 5.8. *Ett udda primtal $p \in \mathbb{N}$ kan skrivas som $p = a^2 + b^2$, med $a, b \in \mathbb{N}$ om och endast om $p \equiv 1 \pmod 4$.*[2]

Bevis. Antag först att $p = a^2+b^2$. Eftersom p är udda så måste a^2 och b^2 ha olika paritet[3], vilket är fallet precis då a och b har olika paritet. Vi kan därför anta att $a = 2s$, och $b = 2r+1$ för några $r, s \in \mathbb{Z}$. Men då får vi $a^2+b^2 = 4(r^2+s^2+s)+1$, så $p \equiv 1 \pmod 4$.

Antag å andra sidan att $p \equiv 1 \pmod 4$. Eftersom $\mathbb{Z}_p^*$ är cyklisk finns ett element n av ordning fyra.[4] Följdaktligen har n^2 multiplikativ ordning två i $\mathbb{Z}_p^*$ och måste således vara kongruent med $-1 \pmod p$, (ekvationen $x^2 - 1 = 0$ har bara två lösningar eftersom polynomet $X^2 - 1 \in \mathbb{Z}_p[X]$ inte kan ha fler än två nollställen; bara en av lösningarna har ordning två). Det betyder att $p|(n^2 + 1)$ i $\mathbb{Z}$, och eftersom $\mathbb{Z}$ är en delring av $\mathbb{Z}[i]$, gäller att p delar $n^2 + 1 = (n + i)(n - i)$ även i $\mathbb{Z}[i]$.

Om p vore ett primelement i $\mathbb{Z}[i]$, så skulle p antingen dela $n + i$ eller $n - i$, vilket är omöjligt: om $p(a + bi) = n + i$ med $a, b \in \mathbb{Z}$, skulle vi nämligen få, genom att jämföra imaginärdelarna i VL och HL, att $pb = 1$. Att p inte heller delar $n - i$ visas med ett liknande argument. Alltså är p inte ett primelement i $\mathbb{Z}[i]$, och eftersom vi vet att $\mathbb{Z}[i]$ är faktoriell vilket säger oss att icke primelement kan reduceras, kan vi ta ett första steg i primfaktoriseringen av p, och skriva $p = (a + bi)(c + di)$ med $a, b, c, d \in \mathbb{Z}$, och varken $a + bi$ eller $c + di$ är inverterbara. Vi tar normen på VL. resp. HL. för att få $p^2 = (a^2 + b^2)(c^2 + d^2)$, och använder 5.7 för att försäkra oss om att både $a^2 + b^2$ och $c^2 + d^2$ är skilda från 1. Men eftersom p är ett primtal i $\mathbb{N}$, måste vi ha $a^2 + b^2 = p$, och det slutför beviset av vår sats. □

Observera att för $p = 2$, dvs det enda jämna primtalet, har vi $p = a^2 + b^2$ om vi väljer lämpliga $a, b \in \mathbb{N}$.

Sats 5.9. *Primelementen i $\mathbb{Z}[i]$ är (så när som på associering):*

1. *primtal $p \in \mathbb{N} \subset \mathbb{Z}[i]$ sådana att $p \equiv 3 \pmod 4$, och*
2. *gaussiska heltal $z = a + bi$ sådana att $N(z) = a^2 + b^2$ är ett primtal.*

Bevis. Alla primtal $p \in \mathbb{N}$ med $p \equiv 3 \pmod 4$ är primelement, ty om

$$p = (a + bi)(c + di) \text{ med } a, b, c, d \in \mathbb{Z}$$

[2]Beviset använder att $\mathbb{Z}_p \setminus \{0\}$ är cyklisk (dvs det existerar ett element a så att $\{1, a, a^2, a^3, .., a^{p-2}\} = \mathbb{Z}_p \setminus \{0\}$). Detta kommer att tas upp på en lektion.

[3]två heltal har *olika paritet* om de ger olika rest vid division med 2, dvs om det ena talet är udda och det andra är jämnt.

[4]Vi definierar ordningen för ett element n som det minsta naturliga talet l sådant att $n^l = 1$. Om vi använder att $\mathbb{Z}_p^* = \{1, a, a^2, \ldots, a^{p-2}\}$ för något väl valt $a \in \mathbb{Z}_p^*$, och antagandet att $p - 1 = 4k$ för något $k \in \mathbb{N}$, ser vi att t.ex. $n = a^k$ har ordning fyra.

vore en faktorisering av p, där varken $a+bi$ eller $c+di$ är inverterbar, så får vi

$$p^2 = N(p) = N(a+bi)N(c+di) = (a^2+b^2)(c^2+d^2),$$

vilket medför att $p = a^2 + b^2$. Detta är dock omöjligt då $p \equiv 3 \pmod 4$, som vi såg i början av beviset för sats 5.8. Gaussiska heltal av typ (2) är också primelement, eftersom om $z = p_1 p_2$ är en faktorisering av z i $\mathbb{Z}[i]$ där varken p_1 eller p_2 är inverterbar, så får vi även en faktorisering av $N(z) = N(p_1)N(p_2)$, med $N(p_1), N(p_2) \neq 1$, men det går inte då $N(z)$ är ett primtal.

Nu återstår att visa att det inte finns några andra primelement i $\mathbb{Z}[i]$. Låt $z \in \mathbb{Z}[i]$ vara ett primelement; vi behöver visa att z antingen är associerat med ett primtal $p \in \mathbb{N}$ med $p \equiv 3 \pmod 4$, eller med ett gaussiskt heltal z sådant att $N(z)$ är ett primtal. Vi visar att z är av typ (1) ifall primfaktoriseringen av det naturliga talet $N(z)$ innehåller en faktor som är kongruent med $3 \pmod 4$, och att z i annat fall är av typ (2).

Antag först att $p|N(z)$, där $p \in \mathbb{N}$ är ett primtal med $p \equiv 3 \pmod 4$. Då är p ett primelement även i $\mathbb{Z}[i]$ enligt bevisets första del, och eftersom $p|N(z) = z\bar{z}$ drar vi slutsatsen att p måste dela antingen z eller $\bar{z}$. Om $p|z$ så är p och z associerade eftersom det inte finns någon icke-trivial faktorisering av primelementet z. Om $p|\bar{z}$ så har vi $px = \bar{z}$ för något $x \in \mathbb{Z}[i]$, men det medför att $p|z$, eftersom $p\bar{x} = \bar{p}\bar{x} = \overline{px} = z$, och då drar vi återigen slutsatsen att p och z är associerade. Vi har alltså visat att om primfaktoriseringen av $N(z)$ (i $\mathbb{N}$) innehåller en primfaktor $p \equiv 3 \pmod 4$ så är $z = p$, så när som på en faktor $\pm 1, \pm i$.

Antag då istället att primfaktoriseringen av $N(z)$ inte innehåller någon primfaktor $p \in \mathbb{N}$ med $p \equiv 3 \pmod 4$. Då måste primfaktoriseringen av $N(z)$ i $\mathbb{N}$ endera innehålla en faktor 2, eller en faktor $p \in \mathbb{N}$ med $p \equiv 1 \pmod 4$, alltså ett tal på formen $a^2 + b^2$ för några väl valda $a, b \in \mathbb{N}$ enligt sats 5.8 (och kommentaren direkt efter beviset). Då är både $t = a + bi \in \mathbb{Z}[i]$ och $\bar{t}$ tal av typ (2), således primelement, och $N(z) = z\bar{z}$ är delbart med $p = t\bar{t}$. Det medför att $t|z\bar{z}$, men då måste antingen $t|z$ eller $t|\bar{z}$. I det första fallet får vi att z och t är associerade, och i det andra fallet får vi att z och $\bar{t}$ är associerade - så i vilket fall är z associerat med ett tal av typ (2), och det slutför beviset. □

Vi beskriver nu hur faktoriseringen av ett givet gaussiskt heltal $z = x+iy$ går till. Om x och y inte är relativt prima, börjar vi med att bryta ut den största gemensamma faktorn till x och y. Varje primfaktor i (x, y) på formen $4k + 3$ är en primfaktor även i $\mathbb{Z}[i]$, så dessa behåller vi som de är. Varje primfaktor i (x, y) som antingen är 2 eller på formen $4k+1$ skriver vi som $p = (a+bi)(a-bi)$; detta är möjligt på grund av sats 5.8, och dessutom är $a + bi$ respektive $a - bi$ primelement i $\mathbb{Z}[i]$ enligt sats 5.9 eftersom de har norm p.

Sedan återstår det att faktorisera gaussiska heltal på formen $z = x + iy$, där x och y är relativt prima. Börja med att faktorisera $N(z) = z\bar{z}$ i naturliga primtal. Varje primfaktor 2 i $N(z)$ (det kan finnas högst en ifall x och y är relativt prima; varför?) ger upphov till en faktor $1 + i$ eller $1 - i$, men eftersom

dessa är associerade med varandra är z automatiskt delbart med $1+i$ om $N(z)$ har en faktor 2. För varje primfaktor på formen $4k+1$ i $N(z)$ får vi en faktor $a+bi$ eller $a-bi$, men dessa är inte associerade med varandra så man behöver pröva från fall till fall vilken av dem det är som delar z. Hur som helst kan z inte vara delbart med både $a+bi$ och $a-bi$ eftersom man i annat fall skulle få en heltalsfaktor a^2+b^2 i z, vilket skulle strida mot att x och y är relativt prima.

Exempel 5.10. Vi faktoriserar $z = 390 + 210i$.

Vi har $(390, 210) = 30 = 2 \cdot 3 \cdot 5$, så det första vi behöver göra är att primfaktorisera 30 i $\mathbb{Z}[i]$. Vi har $2 = (1+i)(1-i)$; 3 är kongruent med 3 modulo 4 och är därför färdigfaktoriserad; 5, slutligen, är kongruent med 1 modulo 4 och kan därför skrivas som summan av två kvadrater $5 = 1^2 + 2^2$. Därför får vi $5 = (2+i)(2-i)$. Av estetiska skäl kan man vilja byta ut $1-i$ mot $-i(1+i)$; då får man $30 = -3i(1+i)^2(2+i)(2-i)$.

Sedan återstår att faktorisera $z/30 = 13+7i$. Vi har $N(13+7i) = 13^2+7^2 = 218 = 2 \cdot 109$. Faktorn 109 är kongruent med 1 modulo 4, så vi kan skriva den som summan av två heltalskvadrater: $109 = 10^2 + 3^2$. Alltså är $13+7i$ delbart med antingen $10+3i$ eller $10-3i$; vi avgör genom prövning.

$$\frac{13+7i}{10+3i} = \frac{(13+7i)(10-3i)}{109} = \frac{151+31i}{109} \notin \mathbb{Z}[i],$$

alltså måste $13+7i$ vara delbart med $10-3i$, och mycket riktigt har vi

$$\frac{13+7i}{10-3i} = 1+i.$$

Eftersom både $10-3i$ och $1+i$ är primelement, blir den slutliga faktoriseringen

$$\begin{aligned} 390+210i &= -3i(1+i)^2(2+i)(2-i)(10-3i)(1+i) \\ &= -3i(1+i)^3(2+i)(2-i)(10-3i). \end{aligned}$$

6 Polynom

Vi kan nu tillämpa våra ringkunskaper på olika polynomringar.

Vi minns vår definition av polynom:

Definition 6.1. Givet en ring R så definierar vi *polynomringen* $R[x]$ som mängden av polynom i x, dvs formella summor $\sum_{j=0}^{n} a_i x^i$ där $a_i \in R$ och $n < \infty$. Addition och multiplikation av polynom definieras på samma sätt som vanligt.

En naturlig fråga är vilka egenskaper hos polynomringar med koefficienter i $\mathbb{C}$ (dvs de polynom vi känner till bra sedan tidigare) som återfinns hos allmänna polynomringar.

En mycket naturlig definition är då följande:

Definition 6.2. Vi säger att $a \in R$ är en *rot* till polynomet $p(x) \in R[x]$ om $p(a) = 0$.

För våra klassiska polynom med koefficienter i $\mathbb{C}$ så är det ekvivalent att a är en rot till $p(x)$ och att $(x-a)|p(x)$. Detta gäller även för allmänna polynomringar.

Sats 6.3. *[Faktorsatsen] Ett polynom $p \in R[x]$ har en rot $a \in R$ om och endast om $p(x) = (x-a) \cdot h(x)$ för något $h(x) \in R[x]$.*

Bevis. Ena hållet är uppenbart. Andra hållet är en bra övning, ett tips är att först visa att det stämmer för $a = 0$ och sedan generalisera det genom att flytta nollstället på lämpligt vis. □

Detta är visserligen ett trevligt resultat, men våra ringar kan fortfarande ha flera faktorer än graden. Tag exempelvis polynomet $2x$ sett som polynom med koefficienter i $\mathbb{Z}_4$. Det polynomet har både 0 och 2 som rötter, alltså så skall x samt $x-2$ båda vara faktorer. Detta medför inte att deras produkt är faktorer då vi inte har unik faktorisering i allmänna polynomringar. Det gäller ju att $2x = 2(x-2)$.För att undvika sådana situationer så kan vi lägga på kravet att R skall vara ett integritetsområde.

Korollarium 6.4. *Låt $p \in R[x]$ där R är ett integritetsområde. Om $a_i, 1 \leq i \leq n$ är parvis olika rötter till p så följer det att $p(x) = h(x) \cdot (x-a_1) \cdot (x-a_2) \cdot \cdots \cdot (x-a_n)$.*

Bevis. Antag att vi har p och a_i som i satsen. Då är a_n en rot till p och vi har enligt Sats 6.3 att $p(x) = h_n(x) \cdot (x-a_n)$ för något polynom $h_n \in R[x]$. Då a_{n-1} är en rot till p så måste $0 = p(a_{n-1}) = h_n(a_{n-1}) \cdot (a_{n-1} - a_n)$. Eftersom $a_{n-1} \neq a_n$ så är $a_{n-1} - a_n \neq 0$. Då vi är på ett integritetsområde så ger andra annulleringslagen att då $0 = 0 \cdot (a_{n-1} - a_n) = h_n(a_{n-1}) \cdot (a_{n-1} - a_n)$ så följer det att $0 = h_n(a_{n-1})$. Alltså så är a_{n-1} en rot till h_n och vi kan igen applicera Sats 6.3. Induktion ger nu beviset. □

Sats 6.5. *En ring R är ett integritetsområde om och endast om $R[x]$ är ett integritetsområde.*

Bevis. Ena hållet är trivialt, ty om R inte är ett integritetsområde så är knappast $R[x]$ det då vi kan hitta nolldelare bland konstanta polynom. Detta visar att om $R[x]$ är ett integritetsområde så måste R också vara det. Andra hållet är lite klurigare. Låt R vara ett integritetsområde. Antag att $p(x) \cdot h(x) = 0$ och undersök sedan termerna av högst grad med nollskild koefficient (om det ej existerar så är vi klara, eller hur?). Då är $p(x) = ax^n +$ "termer av lägre grad" och $h(x) = bx^m +$ "termer av lägre grad" vilket ger $0 = abx^{n+m} +$ "termer av lägre grad". Då nollpolynomet har koefficienten noll framför varje term så följer det att koefficienten framför x^{n+m} skall vara noll. Då är speciellt $ab = 0$ vilket ger motsägelse då a, b var nollskilda och element i ett integritetsområde. Alltså så måste $h \cdot p = 0$ medföra att p eller h är lika med noll, dvs $R[X]$ är ett integritetsområde. □

Polynomringar över ett integritetsområde har flera trevliga egenskaper, bland annat så uppför sig graden som förväntat. För att diskutera detta begrepp så definierar vi graden av ett polynom som följer:

Definition 6.6. Givet ett polynom $p(x)$ så definierar vi *graden* av p, också kallat $\delta(p)$ som det högsta heltal d sådant att koefficienten framför x^d är nollskild. Vi definierar graden för nollpolynomet som $-\infty$.

Denna gradfunktion har trevliga egenskaper (trots att den inte är en homomorfism). Följande två regler kan enkelt härledas (med uppenbara definitioner av vad som händer med $-\infty$).

- $\delta(p + h) \leq max(\delta(p), \delta(h))$
- $\delta(ph) \leq (\delta(p) + \delta(h))$. Likhet gäller om R är ett integritetsområde.

En av de trevligare egenskaperna hos heltalen är unik faktorisering så en naturlig fråga är om samma egenskap går att återfinna hos polynomringar med lämpliga krav?

Sats 6.7. *Om K är en kropp, så är polynomringen $K[x]$ faktoriell.*

Bevis. Gradfunktionen uppfyller kraven som ställs för att $K[x]$ skall vara en euklidisk ring. Då euklidisk medför faktoriell så är vi klara. Vi skall alltså visa att den uppfyller kraven.

Det första kravet, att $\delta(a(x)) \leq \delta(a(x)b(x))$ är uppfyllt, ty kroppar är alltid integritetsområden så $\delta(a(x)b(x)) = \delta(a(x)) + \delta(b(x))$. Då vi enbart studerar nollskilda polynom så är gradfunktionen ickenegativ, alltså är $\delta(a(x)) \leq \delta(a(x)b(x))$. Vi måste också visa att gradfunktionen uppfyller den andra egenskapen, att för varje par av polynom $a(x), b(x)$ så existerar det polynom $q(x), r(x)$

sådana att $a(x) = b(x)q(x) + r(x)$ samt att $\delta(r(x)) < \delta(b(x))$ eller $r(x) = 0$. Vi gör detta med induktion på $a(x)$:s grad. Om $a(x) = a$ är ett konstant polynom så är vi klara ty antingen så är $b(x)$ av högre grad än $a(x)$ och vi kan välja $r(x) = a(x)$ eller så har $b(x)$ samma grad som $a(x)$, men då kan vi ta $q(x) = ab^{-1}$. Detta ger oss vårt basfall. Antag nu att det är visat för polynom $a(x)$ av grad upp till n. Tag ett godtyckligt $a(x)$ av grad $n+1$ och ett godtyckligt $b(x)$ Om $\delta(b(x)) > \delta(a(x))$ så tar vi $r(x) = a(x), q(x) = 0$ och är klara. Annars så kommer vi att ha $a(x) = cx^{n+1} + ...$ samt $b(x) = dx^{\delta(b)} + ...$ med $c, d \neq 0$. Tag då $\hat{a}(x) = a(x) - cd^{-1}x^{n+1-\delta(b(x))}b(x)$. Då har vi att $\delta(\hat{a}(x)) < n + 1$ och alltså enligt induktionsantagandet att $\hat{a}(x) = \hat{q}(x)b(x) + r(x)$ med antingen $r(x) = 0$ eller $\delta(r(x)) = \delta(b(x))$. Välj då $q(x) = \hat{q}(x) + cd^{-1}x^{n+1-\delta(b(x))}$ enligt tidigare så har vi då

$$q(x)b(x) + r(x) = \hat{q}(x)b(x) + cd^{-1}x^{n+1-\delta(b(x))}b(x) + r(x) = a(x)$$

vilket skulle bevisas.

□

7 Algebraisk geometri*

En väg som leder till nya intressanta problem och trevliga lösningar på gamla problem kan man få genom att studera vad som händer med gamla objekt då man tar dem till nya situationer. Ifall vi tittar på cirkeln i $\mathbb{R}^2$ definierad av $x^2 + y^2 = 1$ så kanske det känns som man har ganska bra koll på hur den ser ut och beter sig. Men ifall vi tar exakt samma ekvation och istället tittar på lösningsmängden i $\mathbb{C}^2$ så får man något nytt. Hur ser denna komplexifierade cirkel ut? Vad händer om man tar andra objekt och gör samma sak, exempelvis en tredjegradskurva $y^2 = x^2(x+1)$? Det visar sig att (med ett par modifikationer för hur saker ska bete sig vid oändligheten) att cirkeln ser ut som en sfär, medan tredjegradskurvan faktiskt ser ut som en torus. Genom att använda de komplexa bitarna så kan man sedan få reda på saker om rent reella fenomen eftersom man inte bara kan undersöka strukturen hos de reella bitarna utan även får reda på hur de komplexa bitarna interagerar.

Detta ger upphov till ytterligare frågor, om man exempelvis klipper bort precis de reella bitarna, kommer de komplexa bitarna att sitta ihop eller trillar de isär i olika delar? Kan man dra dessa slutsatser enbart genom information om de reella bitarna? Hur bestämmer graden på objektet dess möjliga topologiska egenskaper?

Huvudsakligen är man intresserad av mer generella (geometriska) objekt som ofta lever i affina rum k^n där k är en algebraiskt sluten kropp. I studiet av dessa har man stor nytta av våra nyvunna algebrakunskaper.

Ämnet algebraisk geometri har som utgångspunkt studiet av så kallade algebraiska mängder. En algebraisk mängd är, per definition, den gemensamma nollställemängden till en ändlig familj av polynom.

Definition 7.1. Om k är en algebraiskt sluten kropp och $n \in \mathbb{N}$, säger vi att $X \subset k^n$ är en *algebraisk mängd* i k^n om det finns polynom $p_1, \ldots, p_m \in k[x_1, \ldots, x_n]$ sådana att

$$X = \{x \in k^n \mid p_i(x) = 0 \text{ för i=1,...,m}\}.$$

Exempel 7.2. Om vi väljer $n = 2$, tar $k = \mathbb{C}$, och låter $p_1(x, y) = x^2 + y^2 - 1$, ser vi att cirkeln

$$C = \{(x, y) \in \mathbb{C} \mid x^2 + y^2 = 1\}$$

är en algebraisk mängd.

Ett annat grundläggande begrepp inom algebraisk geometri är *reguljära funktioner*. Om X är en algebraisk mängd, så säger vi att en funktion $f : X \longrightarrow k$ är en *reguljär funktion på* X om det finns ett polynom $p \in k[x_1, \ldots, x_n]$ sådat att $f(x) = p(x)$ för alla $x \in X$. Observera att p definierar en k-värd funktion på hela k^n, och att f bara är definierad på en delmängd till k^n. Villkoret $f(x) = p(x)$ betyder precis att restriktionen av p till X sammanfaller med funktionen f.

Det kan mycket väl hända att olika polynom definierade på k^n har samma restriktion till X. Till exempel definierar polynomen $x^2 + y^2 + x - 1$ och x i $k[x, y]$ samma reguljära funktion på cirkeln C, eftersom $x^2 + y^2 = 1$ för alla $(x, y) \in C$.

De reguljära funktionerna på en algebraisk mängd bildar en ring. Ett tillräckligt och nödvändigt villkor för att två olika polynom $p_1, p_2 \in k[x, y]$ ska definiera samma reguljära funktion på cirkeln är att deras differens är 0 på C, och det är fallet precis om deras differens är en multipel av $x^2 + y^2 - 1$. Det bör kanske nämnas att detta inte är helt självklart då man inte alltid direkt kan ta polynomet som definierar vår algebraiska mängd (exempelvis så definierar ju x^2 och x precis samma algebraiska mängd). Det finns dock en sats (Hilberts Nullstellensatz) som i detta fall säger att vårt polynom fungerar. Därför är ringen av reguljära funktioner på cirkeln lika med

$$k[x, y]/I(C),$$

där $I(C) = \langle x^2 + y^2 - 1 \rangle$ betecknar huvudidealet i $k[x, y]$ som genereras av $x^2 + y^2 - 1$.

Nu har vi alltså på ett naturligt sätt, utgående från en algebraisk mängd $X \subset k^n$ bildat ett ideal $I(X)$ bestående av de polynom i $k[x_1, \ldots, x_n]$ vars restriktion till X är 0. Det idealet i sin tur används för att beskriva ringen av reguljära funktioner på X, säg A. All information om den algebraiska mängden X finns nu kodad i ringen A, så man förlorar inte någon kunskap om X genom att studera enbart A. Vi säger att två algebraiska mängder är isomorfa om och endast om deras ringar av reguljära funktioner är isomorfa. Detta gör att vi kan frigöra oss från studiet av enskilda algebraiska mängder som ligger i vissa speciella k^n. Exempelvis så ser vi de algebraiska mängderna som beskrivs av x^2+y^2-1 i k^2 och de gemensamma nollställena till de två polynomen x^2+y^2-1 och z i k^3 som isomorfa. Samma idé kan sägas framkomma i geometrin (jämför såkallad ”intrinsic geometry”) och topologin där man vill flytta fokus från det omgivande rummet till själva objektet.

Om man istället utgår från ett ideal J i $k[x_1, \ldots, x_n]$, kan man bilda en algebraisk mängd i k^n, nämligen den som beskrivs genom

$$X = \{x \in X \mid p(x) = 0 \text{ för alla } p \in J\}.$$

På så sätt får man en korrespondens mellan algebraiska mängder i k^n och ideal i polynomringen $k[x_1, \ldots, x_n]$. Om man dessutom lägger ett villkor på idealen, nämligen att $a^m \in I \longrightarrow a \in I$ (exempelvis så får idealet inte innehålla x^2 utan att innehålla x), så blir den beskrivna korrespondensen bijektiv. Till varje ideal hör precis en algebraisk mängd och till varje algebraisk mängd hör precis ett ideal. Alltså finns det en mycket nära koppling mellan ren algebra och de geometriska objekt som kan definieras i termer av nollställen till polynom.