Växjö University

Department of Mathematics, Statistics and Computer Science

Reports from MASDA - Rapporter från MASDA

Symmetriska polynom

Omar Hijazi

Maj
1999

MASDA
Växjö University
S-351 95 VÄXJÖ

Report 9933
ISSN 1400-1942
ISRN VXU/MASDA/MA/E/--9933--SE

Abstract

Omar Hijazi, Växjö Universitet
Symmetriska polynom
Växjö 1999

Denna uppsats behandlar symmetriska polynom. Syftet är att klarlägga beviset så att det kan förstås av matematikstuderande på c-nivå (41-60 poäng). Bortsett från beviset för huvudsatsen bevisas även två andra satser som är viktiga, en sats om entydigheten hos polynomen och en sats om att π är transcendent. Huvudsatsen är främst hämtad ur Nagell (1957) *Lärobok i algebra*.

Innehåll

1. Inledning

1.1 Motiv för valet och syfte med uppsatsen

Motivet för att skriva om *symmetriska polynom* är mitt stora intresse för algebra och ekvationer. Då denna del är viktig och ger ökad förståelse för matematiken har jag valt att skriva om detta. Jag skrev om *symmetriska polynom* dels för att öka mina egna kunskaper och för att ge andra studenter med intresse en chans att förstå satsen och dess bevis.
Syftet med denna uppsats är att utifrån redan kända bevis för satsen om symmetriska polynom, göra en sammanställning, som kan förstås av matematikstuderande på C-nivå.

1.2 Historisk bakgrund

Inom teorin för algebraiska ekvationer är det naturligt att studera symmetriska polynom. I en algebraisk ekvation i en obekant *x* av grad *n* är koefficienterna för potenserna x^k speciella symmetriska polynom i rötterna $x_1, \ldots, x_n$ de s.k. symmetriska grundfunktionerna. Dessa har varit kända sedan 1500-talet.
Fransmannen August Louis Cauchy (1789-1857) visade att varje symmetriskt polynom kan skrivas som ett polynom i de symmetriska grundfunktionerna. Isaac Newton (1642-1727) hade tidigare visat att bl.a. potenssumman $x_1^k + \ldots + x_n^k$ kan skrivas som polynom i grundfunktionerna.
Symmetriska polynom spelar en viktig roll i teorin för s.k. grupprepresentationer och där är de s.k. Schurpolynom betydelsefulla. Polynomen har uppkallats efter Issai Schur (1875-1941) som givit betydelsefulla bidrag till teorin för grupper.
Huvudsatsen för symmetriska polynom är ett viktigt hjälpmedel för att utveckla teorin för algebraiska ekvationer. Bland annat spelar den en avgörande roll i det bevis för transcendensen av π som vi ger i avsnitt 4.

2. Symmetriska polynom

2.1 Några begrepp och definitioner

- Homogena polynom

Om ett polynoms termer alla är av samma grad, sägs polynomet vara homogent.

- Heltals polynom

Är polynomets koefficienter hela rationella tal, är det ett heltalspolynom.

- Lineära kvadratiska, kubiska och bikvadratiska polynom

Ett polynom sägs vara lineärt, kvadratiskt, kubiskt eller bikvadratiskt beroende på dess grad är 1, 2, 3 eller 4.

2.2 Definition och exempel på symmetriska och icke symmetriska polynom

En rationell funktion $F(x_1, x_2, \ldots, x_n)$ är symmetrisk om den inte ändras när variablerna permuteras på alla de *n!* möjliga sätten. T.ex. är funktionerna

$$\frac{x_1}{x_2} + \frac{x_2}{x_3} + \frac{x_2}{x_1} + \frac{x_3}{x_1} + \frac{x_1}{x_3} + \frac{x_3}{x_2},$$

$$(x_1 + x_2) + (x_1 + x_3) + (x_2 + x_3),$$

och

$$x_1^4 + x_2^4 + x_3^4 + 5x_1x_2x_3$$

symmetriska funktioner av de tre variablerna x_1, x_2 och x_3. Summan, differensen eller produkten av två symmetriska funktioner är en symmetrisk funktion. Detsamma gäller även för kvoten mellan två symmetriska funktioner.
Exempel på icke symmetriska funktioner är

$$x_1^4 - x_2^4$$

och

$$x_1^3 - x_2^3.$$

2.3 De symmetriska grundfunktionerna

Produkten

$$P(z) = (z - x_1)(z - x_2)\ldots(z - x_n)$$

är tydligen ett symmetriskt polynom av variablerna $x_1, x_2, \ldots x_n$. Utför man multiplikation, får man ett *n*:*te* grads polynom i *z*,

$$P(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_n,$$

där koefficienterna $a_1, a_2, \ldots, a_n$ är symmetriska polynom i variablerna x_i och bestämda av formlerna

$$\begin{aligned}
a_1 &= -\sum x_i, \\
a_2 &= +\sum x_i x_j, \\
a_3 &= -\sum x_i x_j x_k, \\
&\vdots \\
a_v &= (-1)^v \sum x_{i_1} x_{i_2} \ldots x_{i_v}, \\
&\vdots \\
a_n &= (-1)^n x_1, x_2 \ldots x_n.
\end{aligned}$$

Summan i formeln för a_v sträcker sig över alla de i_v som uppfyller villkoren $1 \le i_1 < i_2 \ldots \le n$. Storheten a_v är således ett homogent symmetriskt polynom av variablerna x_i av graden *v* och vi sätter $a_v = (-1)^v \sigma_v$ och kallar $\sigma_1, \sigma_2, \ldots, \sigma_n$ för de *symmetriska grundfunktionerna* av $x_1, x_2, \ldots x_n$.

Ett klassiskt resultat är att varje symmetriskt polynom i $x_1, x_2, \ldots x_n$ kan skrivas som ett polynom i $\sigma_1, \sigma_2, \ldots, \sigma_n$. Vi inleder med två exempel

Exempel

1.

Vi ska skriva om polynomet,

$$x_1^2 + x_2^2 + x_3^2,$$

i σ_1, σ_2 och σ_3 där

$$\sigma_1 = x_1 + x_2 + x_3$$
$$\sigma_2 = x_1 x_2 + x_2 x_3 + x_1 x_3$$
$$\sigma_3 = x_1 x_2 x_3.$$

Vi kvadrerar σ_1 och får

$$\sigma_1^2 = (x_1 + x_2 + x_3)^2 =$$

$$= x_1^2 + x_2^2 + x_3^2 + 2x_1 x_2 + 2x_1 x_3 + 2x_2 x_3 \; .$$

Vi har nu fått det utryck vi söker samt en icke önskvärd restterm $2x_1 x_2 + 2x_1 x_3 + 2x_2 x_3$. Men den är $2\sigma_2$ och allstå är

$$x_1^2 + x_2^2 + x_3^2 = \sigma_1^2 - 2\sigma_2.$$

2.

I detta exempel skall vi skriva om,

$$x_1^2 x_2 x_3 + x_1 x_2^2 x_3 + x_1 x_2 x_3^2 \, ,$$

i σ_1, σ_2 och σ_3. Vi kan bryta ut $x_1 x_2 x_3$

$$(x_1 + x_2 + x_3)(x_1 x_2 x_3) =$$
$$= x_1^2 x_1 x_1 + x_1 x_2^2 x_3 + x_1 x_2 x_3^2 \; .$$

Alltså kan det ursprungliga polynomet skrivas

$$\sigma_1 \cdot \sigma_3.$$

Innan vi bevisar huvudsatsen visar vi att $\sigma_1, \sigma_2, \ldots, \sigma_n$ i viss mening är oberoende av varandra.

Sats 1.

Om man i ett icke identiskt försvinnande polynom $F(z_1, z_2, \ldots z_n)$ istället för variablerna $z_1, z_2, \ldots z_n$ sätter in de symmetriska grundfunktionerna av variablerna $x_1, x_2, \ldots x_n$ där $z_1 = \sigma_1, z_2 = \sigma_2, \ldots, z_n = \sigma_n$ får man ett icke identiskt försvinnande polynom $G(x_1, x_2, \ldots, x_n)$ i variablerna $x_1, x_2, \ldots, x_n$

Bevis:

Den allmänna termen i $F(z_1, z_2, \ldots, z_n)$ är

$$T_v = C_v z_1^{v_1} z_2^{v_2} \ldots z_n^{v_n}$$

där $v = [v_1, v_2, \ldots, v_n]$. Koefficienten C_v är skild från noll. Bland termerna i F väljer vi nu ut en bestämd term T enligt följande metod:

Vi betraktar enbart sådana termer i F i vilka summan $v_1 + v_2 + \ldots + v_n$ har sitt största värde. Låt detta vara k_1. Ifall man endast finner en term T_v med denna egenskap, sätter vi $T = T_v$. Finner man flera sådana termer, betraktar vi bland dessa enbart de i vilka summan $v_2 + v_3 + \ldots + v_n$ har sitt största värde. Låt denna vara k_2. Exponenten v_1 är då entydigt bestämd av relationen $v_1 = k_1 - k_2$. Vi upprepar denna procedur r antal gånger tills vi erhåller en term

$$T = T_v$$

där $v_1, v_2, \ldots, v_n$ är entydigt bestämda av

$$v_s = k_s - k_{s+1}$$

för $s = 1,\ 2, \ldots, r$. Även för $s = r+1,\ \ r+2,\ \ \ldots,\ n$ definierar vi talen k_s enligt formeln

$$k_s = v_s + v_{s+1} + \ldots + v_n.$$

Genom insättning av de symmetriska grundfunktionerna övergår termen T i uttrycket

$$C_v (x_1 + x_2 + \ldots + x_n)^{v_1} (x_1 x_2 + \ldots)^{v_2} \ldots (x_1 x_2 \ldots x_n)^{v_n}\ .$$

I detta polynom $G(x_1, x_2, \dots, x_n)$ förekommer därför potensprodukten

$$x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$$

med koefficienter C_v. På grund av det sätt som vi definierat T kan ingen annan term i $F(z_1, z_2, \dots z_n)$ ge upphov till en potensprodukt av variablerna $x_1, x_2, \dots, x_n$ med exponenterna $k_1, k_2, \dots, k_n$.

Satsen är därmed bevisad.

En konsekvens är följande sats.

Följdsats

Om ett symmetriskt polynom av variablerna $x_1, x_2, \dots, x_n$ kan framställas som ett polynom av grundfunktionerna av samma variabler, alltså på formen

$$P(\sigma_1, \sigma_2, \dots, \sigma_n) = \sum \gamma_v \sigma_1^{v_1} \sigma_2^{v_2} \dots \sigma_n^{v_n}$$

så är denna framställning entydig.

Bevis

Antag att

$$P_1(\sigma_1, \sigma_2, \dots, \sigma_n) = \sum \beta_v \ \sigma_1^{v_1} \sigma_2^{v_2} \dots \sigma_n^{v_n}$$

vore en annan sådan framställning, då måste differensen

$$P(a_1, a_2, \dots, a_n) - P_1(a_1, a_2, \dots, a_n)$$

försvinna identiskt, om vi sätter in $\sigma_1 = x_1 + \dots + x_n$, $\sigma_2 = x_1 x_2 + \dots$, vilket enligt föregående sats är omöjligt.

2.4 Potenssummor uttryckta i de elementära grundfunktionerna

Vi betraktar först polynom i tre variabler x_1, x_2, x_3 och sätter

$$S_n = x_1^n + x_2^n + x_3^n.$$

Naturligtvis gäller att $S_0 = x_1^0 + x_2^0 + x_3^0 = 3$, $S_1 = x_1 + x_2 + x_3 = \sigma_1$ och vi har visat att

$$S_2 = \sigma_1^2 - 2\sigma_2.$$

Vi vet att x_1, x_2, x_3 är rötter till $x^3 - \sigma_1 x^2 + \sigma_2 x - \sigma_3 = 0$ vilket ger

$$\begin{aligned} x_1^3 - \sigma_1 x_1^2 + \sigma_2 x_1 - \sigma_3 &= 0 \\ x_2^3 - \sigma_1 x_2^2 + \sigma_2 x_2 - \sigma_3 &= 0 \\ x_3^3 - \sigma_1 x_3^2 + \sigma_2 x_3 - \sigma_3 &= 0. \end{aligned}$$

Efter addition får vi

$$x_1^3 + x_2^3 + x_3^3 - \sigma_1 S_2 + \sigma_2 S_1 - 3\sigma_3 = 0$$

dvs

$$S_3 = \sigma_1 S_2 - \sigma_2 S_1 + 3\sigma_3.$$

Eftersom $S_1 = \sigma_1$ och $S_2 = \sigma_1^2 - 2\sigma_2$ så är

$$S_3 = \sigma_1^2 - 3\sigma_2\sigma_1 + 3\sigma_3.$$

Därmed har vi lyckats uttrycka S_3 i de symmetriska grundfunktionerna. För att på samma sätt bestämma S_n konstaterar vi att x_1, x_2, x_3 också är rötter till ekvationen

$$x^n - \sigma_1 x^{n-1} + \sigma_2 x^{n-2} - \sigma_3 x^{n-3} = 0$$

om $n \geq 3$. Vänster ledet har vi ju fått genom att multiplicera

$$x^3 - \sigma_1 x^2 + \sigma_2 x - \sigma_3$$

med x^{n-3}.

Alltså har vi

$$\begin{aligned}
x_1^n - \sigma_1 x_1^{n-1} + \sigma_2 x_1^{n-2} - \sigma_3 x_1^{n-3} &= 0 \\
x_2^n - \sigma_1 x_2^{n-1} + \sigma_2 x_2^{n-2} - \sigma_3 x_2^{n-3} &= 0 \\
x_3^n - \sigma_1 x_3^{n-1} + \sigma_2 x_3^{n-2} - \sigma_3 x_3^{n-3} &= 0.
\end{aligned}$$

Som i föregående fall, får vi efter addition

$$S_n - \sigma_1 S_{n-1} + \sigma_2 S_{n-2} - S_{n-3}\sigma_3 = 0$$

dvs

$$S_n = \sigma_1 S_{n-1} - \sigma_2 S_{n-2} + S_{n-3}\sigma_3.$$

Vi får en rekursions formel som vi kan använda för att uttrycka $S_n, n = 1,2,\ldots$ i $\sigma_1, \sigma_2, \sigma_3$ eftersom vi vet att $S_0 = 3, S_1 = \sigma_1, S_2 = \sigma_1^2 - 2\sigma_2$.

Antag att vi på samma sätt vill uttrycka

$$S_n = x_1^n + x_2^n + \ldots + x_k^n$$

i de elementära symmetriska funktionerna $\sigma_1, \ldots, \sigma_k$. Vi måste då på något sätt först uttrycka $S_1, S_2, \ldots, S_{k-1}$ i $\sigma_1, \ldots, \sigma_k$. Därefter kan vi beräkna S_n för $n \geq k$ genom rekursions formel

$$(1) \quad S_n = -\sigma_1 S_{n-1} + \sigma_2 S_{n-2} - \ldots - (-1)^k \sigma_k S_{n-k}.$$

Talen $x_1, \ldots, x_k$ är nämligen rötter till ekvation

$$x^k - \sigma_1 x^{k-1} + \sigma_2 x^{k-2} + \ldots + (-1)^k \sigma_k = 0$$

varför

$$x_j^k = \sigma_1 x_j^{k-1} - \sigma_2 x_j^{k-2} - \ldots - (-1)^k \sigma_k$$

för $\gamma = 1,2,\ldots,k$. Vi får då (1) genom att multiplicera dessa likheter med x_j^{n-k} och addera dem för $j = 1,\ldots,k$. Observera att $S_0 = x_1^0 + \ldots + x_k^o = k$.

2.5 Huvudsatsen för symmetriska polynom

Sats 2.

Varje symmetriskt polynom

$$(1) \qquad S(x_1, x_2, \ldots, x_n) = \sum \beta_{k_1, k_2, \ldots, k_n} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$$

kan entydigt framställas som ett polynom av de symmetriska grundfunktionerna $\sigma_1, \sigma_2, \ldots \sigma_n$ av variablerna $x_1, x_2 \ldots, x_n$ under formen

$$(2) \qquad P(\sigma_1, \sigma_2, \ldots, \sigma_n) = \sum \gamma_{v_1, v_2, \ldots, v_n} \sigma_1^{v_1} \sigma_2^{v_2}, \ldots, \sigma_n^{v_n},$$

där koefficienterna γ är homogena lineära heltalspolynom av koefficienterna β.

Bevis

Satsen bevisas först för $n = 2$ med variablerna x_1 och x_2. Vi har då $\sigma_1 = -(x_1 + x_2) = -x_1 - x_2$ och $\sigma_2 = x_1 x_2$. Om $S(x_1, x_2)$ är ett godtyckligt symmetriskt polynom av x_1 och x_2 är

$$S(x_1, x_2) = S(x_1, -\sigma_1 - x_1) = A_0 x_1^m + A_1 x_1^{m-1} + \ldots + A_m .$$

Koefficienterna A_i är polynom i σ_1 med koefficienter som är homogena lineära heltalspolynom av koefficienterna i S. Detta inses med hjälp av binomialteoremet. Sätter man nu

$$\phi(z) = A_0 z^m + A_1 z^{m-1} + \ldots + A_m$$

och dividerar med

$$f(z) = (z - x_1)(z - x_2) = z^2 - \sigma_1 z + \sigma_2$$

så får man

$$\phi(z) = Q(z) f(z) + A + Bz.$$

Eftersom koefficienten för z^2 i $f(z)$ är 1 så är A och B polynom i σ_1 och σ_2 med koefficienter som är homogena lineära heltalspolynom av koefficienterna i S. För $z = x_1$ får man alltså

$$S(x_1, x_2) = A + Bx_1$$

Men eftersom $S(x_1, x_2)$ är symmetriskt måste man även ha

$$S(x_1, x_2) = A + Bx_2.$$

Då nu x_1 och x_2 är av varandra oberoende variabler leder det till att $B = 0$ och alltså till

$$S(x_1, x_2) = A.$$

Satsen är därmed bevisad för $n = 2$ eftersom A är ett polynom i σ_1, σ_2.

Vi skall nu visa satsen med hjälp av induktion över antalet variabler n. Antag att påståendet är sant för alla symmetriska polynom i $n-1$ variabler. En symmetrisk funktionen $S(x_1, x_2, \ldots, x_n)$ i n variabler ordnas efter fallande potenser av x_1, alltså

$$(3) \qquad S = S_0 x_1^u + S_1 x_1^{u-1} + \ldots + S_{\mu-1} x_1 + S_\mu .$$

Koefficienterna $S_0, S_1, \ldots, S_\mu$ är tydligen symmetriska polynom av de $n-1$ variablerna $x_2, x_3, \ldots, x_n$ med koefficienter som är homogena lineära heltalspolynom av koefficienterna i S. De symmetriska grundfunktionerna av dessa $n-1$ variabler betecknas med

$$\sigma_1', \sigma_2', \ldots, \sigma_{n-1}'$$

och som förut sätter vi $a_j' = (-1)^j \sigma_j'$ och $a_j = (-1)^j \sigma_j$. Om nu

$$f(z) = \prod_{i=1}^{n} (z - x_i) = z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n$$

är

$$\frac{f(z)}{z - x_1} = z^{n-1} + a_1' z^{n-2} + \ldots + a_{n-2}' z + a_{n-1}'.$$

där

$$\begin{aligned}
a_1' &= x_1 + a_1 \\
a_2' &= x_1^2 + a_1 x_1 + a_2 \\
a_3' &= x_1^3 + a_1 x_1^2 + a_2 x_1 + a_3 \\
&\vdots \\
a_{n-1}' &= x_1^{n-1} + a_1 x_1^{n-2} + \ldots + a_{n-1}
\end{aligned}$$

Detta inses på följande sätt:

Om vi i likheten

$$f(z) = (z - x_1)\left(z^{n-1} + a_1^{'} z^{n-2} + \ldots + a_{n-2}^{'} z + a_{n-1}\right) =$$
$$= z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n$$

utför multiplikation och identifierar koefficienterna får vi

$$a_1 = a_1^{'} - x_1, a_2 = a_2^{'} - a_1^{'} x_1, \ldots, a_{n-1} = a_{n-1}^{'} - a_{n-2}^{'} x_1.$$

Detta ger att

$$a_v^{'} = a_v + a_{v-1}^{'} x_1 = a_v + \left(a_{v-1} + a_{v-2}^{'} x_1\right) x_1 = a_v + a_{v-1} x_1 + a_{v-2}^{'} x_1^2 = \ldots =$$
$$= a_v + a_{v-1} x_1 + a_{v-2} x_1^2 + \ldots + a_1 x^{v-1} + x^v.$$

Eftersom satsen enligt förutsättningen gäller för $n-1$ variabler, kan koefficienterna $S_0, S_1, \ldots, S_\mu$ i (3) uttryckas som polynom i $a_1, a_2, \ldots, a_n$ och x_1 med koefficienter som är homogena lineära heltalspolynom av koefficienterna i *S*.

Genom att i (3) införa ovanstående uttryck för $\sigma_1^{'}, \sigma_2^{'}, \ldots, \sigma_{n-1}^{'}$ och sedan ordna efter fallande potenser av x_1 får man

$$S = A_0 x_1^m + A_1 x_1^{m-1} + \ldots + A_{m-1} x_1 + A_m.$$

Koefficienterna $A_0, A_1, \ldots, A_m$ är här polynom i $\sigma_1, \ldots, \sigma_n$ med koefficienter som är homogena lineära heltalspolynom av koefficienterna i *S*.

Sätter man nu, precis som innan för $n = 2$

$$\phi(z) = A_0 z^m + A_1 z^{m-1} + \ldots + A_{m-1} z + A_m$$

och dividerar med *f(z)*, får man

$$(4) \qquad \phi(z) = Q(z) f(z) + \psi(z)$$

med

$$\psi(z) = C_0 z^{n-1} + C_1 z^{n-2} + \ldots + C_{n-2} z + C_{n-1}.$$

Koefficienterna $C_0, C_1 \ldots, C_{n-2}, C_{n-1}$ är här polynom i $\sigma_1, \sigma_2, \ldots, \sigma_n$ med koefficienterna som är homogena lineära heltalspolynom av koefficienterna i *S* eftersom högstagradskoefficienten i $f(x)$ är 1.

Sätter man in $z = x_1$ i (4) får man $f(x_1) = 0$ och

$$S = \phi(x_1) = \psi(x_1).$$

Eftersom S är symmetriskt i $x_1, x_2, \ldots, x_n$ måste denna likhet fortfarande gälla om x_1 ersätts med vilken som helst av de övriga variablerna x_i.

Men av detta följer att polynomet

$$C_0 z^{n-1} + C_1 z^{n-2} + \ldots + C_{n-2} z \; + C_{n-1} - S$$

försvinner för n olika värden $z = x_1, z = x_2, \ldots, z = x_n$. Enligt algebrans fundamentalsats måste polynomet då försvinna identiskt, eftersom dess grad är högst $n-1$. Alltså måste man ha $S = C_{n-1}$, dvs. S är framställd som ett polynom i $\sigma_1, \sigma_2, \ldots, \sigma_n$ vars koefficienter är homogena lineära heltalspolynom av koefficienterna i S.

Därmed är satsen fullständigt bevisad.

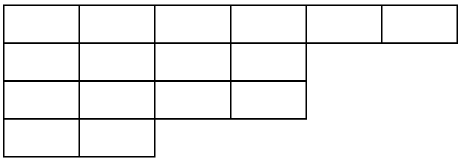
3. Youngs tabell och symmetriska polynom

3.1 Youngdigram

Ett Young-diagram är en samling lådor eller celler. Dessa är ordnade i rader med minskande antal celler för varje rad, med det lägsta antalet celler längst ned. Om n är totala antalet celler så svarar varje diagram mot en partition av n dvs.

$$n = n_1 + n_2 + \ldots + n_k \, , \; n_1 \geq n_2 \ldots \geq n_k$$

Partitionen av 16 i $6+4+4+2$ motsvarar alltså följande diagram:



λ

Vi betecknar partitioner med den grekiska bokstaven λ. Formuleringen $\lambda \leftarrow n$ används för att förklara att λ är en partition av n. $|\lambda|$ betyder antalet partitioner av λ. Genom att använda sig av ett diagram istället för partitioner kan vi sätta in element i cellerna. En insättning av positiva heltal i varje cell kallas en *numrering* av diagrammet om följande gäller

1. I varje rad är elementen växande.
2. I varje kolumn är elementen strängt växande.

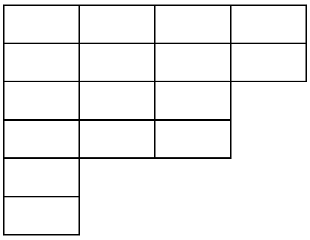
En numrering av ovanstående partition av 16 är t.ex.

1	2	2	3	3	5
2	3	5	5		
4	4	6	6		
5	6				

En numrering av λ

En numrering av ett diagram kallas också en tabell. En *standardtabell* är en tabell där varje element endast uppträder en gång. Ifall vi vänder diagrammet längs dess diagonal (från övre vänster hörn till det högra hörnet längst ned) får vi *diagrammets konjugat*. Konjugatet av λ skrivs som $\overline{\lambda}$. Konjugatet av ovannämnda diagram svarar mot partitionen $16 = 4+4+3+3+1+1$. Observera att i $\overline{\lambda}$ är radernas längder lika med längderna av kolumnerna i λ.

Vidare svarar λ och $\bar{\lambda}$ mot partitioner av samma tal.



$\bar{\lambda}$

3.2 Algoritmer

En viktig algoritm på Youngs tabeller är Schensteds ”utstötnings” algoritm eller *radinsättning*. Med hjälp av den konstruerar man utifrån en tabell T och ett positivt heltal x en ny tabell kallad $T \leftarrow x$. Algoritmen han formuleras på följande sätt.

Om x minst är lika stort som alla element i första raden i T, lägg då x i en ny cell i slutet av första raden. Om inte, finn elementet längst till vänster som är strikt större än x i första raden. Sätt in x istället för det större elementet och ”stöt ut” elementet. Ta nu elementet som blev utstött från första raden och repetera ovan process fast nu på rad 2. Detta fortskrider tills alla element är på plats, även de som blev utstötta, enligt algoritmen. Istället för att ett element stöts ut helt från tabellen, så bildas en ny cell längst ned på diagrammet med detta element.

Vi ger ett exempel och sätter in 2 i nedanstående tabell.

1	2	2	3
2	3	5	5
4	4	6	
5	6		

2: an stöter bort 3: an från första raden.

1	2	2	**2**
2	3	5	5
4	6	6	
5	6		

3: an stöter i sin tur bort den första 5: an på andra raden.

1	2	2	2
2	3	**3**	5
4	4	6	
5	6		

5: an stöter vidare bort 6: an från 3:e raden.

1	2	2	2
2	3	3	5
4	4	**5**	
5	6		

6:an kan slutligen sättas in i slutet av sista raden.

1	2	2	2
2	3	3	5
4	4	5	
5	6	**6**	

Det finns ett lemma som bygger på ”utstötnings” algoritmen som ”avslöjar” resultatet av två insättningar efter varandra. Detta gör att man kan relatera storleken på elementen som sätts in på de platser i de nya cellerna. En rad-insättning bestämmer den uppsättning celler, *R*, vilka element stöts ut tillsammans med cellen där det sista utstötta elementet hamnar. Låt oss kalla detta för *utstötningsrutten*. I exemplet ovan är utstötningsrutten de markerade siffrorna på bilden nedan.

1	2	2	**2**
2	3	**3**	5
4	5	**5**	
5	6	**6**	

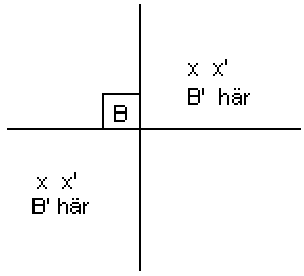
Lemma

Antag att vi gör två insättningar, radvis, efter varandra. Först en insättning x i en tabell T och sedan en insättning x' i tabell resultatet $T \leftarrow x$. Detta ger upphov till två rutter R och R' och två nya celler B och B'.

1. Om $x \leq x'$, så är R strikt till vänster om R' och B är då strikt till vänster om och svagt under[1] B'.

2. Om $x \geq x'$, då är R' svagt till vänster om R och B' är då svagt till vänster om och strikt under B.

[1] ”Svagt under” betyder ”strikt under” eller ”på samma nivå som”.



Bevis

Detta handlar om att hålla reda på vad som händer när elementen ”stöter” genom en viss rad.
Antag att $x \leq x'$ och x stöter ut ett element y från första raden. Element y' utstött av x' måste strikt ligga till höger om cellen där x stöttes ur. Detta på grund av att elementet i cellen eller de till vänster inte är större än x.
Notera att rutten för *R* inte kan sluta ovanför *R'* och om *R'* stannar först, så ändras inte rutten för *R* till höger. Alltså cell *B* måste strikt finnas till vänster om och svagt under *B'*. På bilden nedan motsvarar x = **X,** x'= X, B = **B** och B' = B.

					X		***X***	
					X	X		
				X		X		
			X	X				
		X		B				
		X						
	B							

Å andra sidan, om $x > x'$ och x och x' stöter ut element y respektive y', så kommer cellen i första raden där y' stöts ut att vara till vänster om cellen där y stöttes ut. Alltså är $y > y'$ och algoritmen kan upprepas.

Operationen kan användas till att konstruera produkten av en tabell, $T \cdot U$ från två godtyckliga tabeller *T* och *U*. Man sätter in *U:s* element i *T* och gör det radvis med början med den undre raden och elementet längst till vänster.
Vi skriver

$$T \cdot U = \left(\left(\ldots\left((T \leftarrow x_1) \leftarrow x_2\right) \leftarrow \ldots\right) \leftarrow x_{s-1}\right) \leftarrow x_s$$

där $x_1, \ldots, x_s$ är elementen i *U*.

Ibland arbetar man också med sneda diagram. Ett ”snett” diagram är det diagram som erhålls genom att ta bort ett mindre Young diagram från ett större. Om λ och μ är två diagram med partitionerna

$$\lambda = (\lambda_1, \ldots, \lambda_k),\ \mu = (\mu_1, \ldots, \mu_l)$$

så säger vi att $\mu \subset \lambda$ om

$$l \le k \text{ och } \mu_i \le \lambda_i, \; i = 1, \ldots, l.$$

Den ”sneda” tabell får vi genom att ”ta bort” μ från λ betecknar vi med λ / μ . En sned tabell har ett eller flera inre hörn.

Ett inre hörn är en cell i den mindre skuggade diagrammet μ. T.ex.

				2
				6
	2	4	4	
2	3	6		
5	5			

μ

I exemplet ovan finns det 4 celler på andra raden och 1 cell på den tredje raden som är s.k. inre hörn.

Ett yttre hörn är en cell i λ där varken cellen under eller till höger är i λ. Det betyder att den sista cellen på andra, tredje, fjärde och femte raden är yttre hörn. Värt att notera är att ett inre hörn kan även vara ett yttre hörn.

3.3 Schurpolynomen

Låt λ vara ett Young diagram med högst *m* rader. Antag vidare att *T* är en numrering av λ och att vi i numreringen använder talen *1, 2,…, m*. Mängden av sådana numreringar betecknar vi med σ_λ. Till varje $T \in \sigma_\lambda$ bildar vi ett monom

$$x^T = \prod_{i=1}^{m} (x_i)^{n_i}$$

där

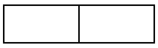
$$n_i = \text{antalet gånger } i \text{ förekommer i } T.$$

Schurpolynomet $S_\lambda(x_1, \ldots, x_m)$ till Young diagrammet λ definieras nu genom

$$S_\lambda(x_1, \ldots, x_m) = \sum_{T \in \sigma_\lambda} x^T .$$

Exempel

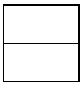
1. Om $\lambda = (2)$ är diagrammet



så är

$$S_\lambda(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$$

2. Om $\lambda = (1^2)$ är

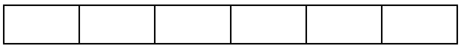


så är

$$S_\lambda(x_1, x_2) = x_1 x_2 .$$

Allmänt gäller

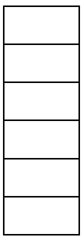
Om $\lambda = (n)$ dvs



n st

så är $S_x(x_1, \ldots, x_m) =$ summan av alla olika homogena monom av graden *n*.

Om $\lambda = (1^n)$ dvs



n st

så måste $1 \leq n \leq m$ och

$$S_\lambda(x_1, \ldots, x_m) = \sigma_n(x_1, \ldots, x_m)$$

där σ_n är de symmetriska grundfunktionerna .
I de exempel vi bekantat oss vid är Schurpolynomen symmetriska. Det gäller allmänt och vi har följande sats.

Sats 3.

Schurpolynomen är symmetriska polynom.

Beviset bygger på två lemma.

Lemma 1

Om λ är ett Youngdiagram så gäller

$$S_\lambda(x_1,\dots,x_m)\,\sigma_p(x_1,\dots,x_m)=\sum_\mu S_\mu(x_1,\dots,x_m)$$

där vi summerar över alla μ som vi får λ genom att lägga till *p* celler så att inte två celler läggs till samma rad.

Beviset som är tekniskt utelämnas. Vi illustrerar lemmat med ett exempel

Exempel

Beräkna $S_\lambda(x_1,x_2)$, $S_\lambda(x_1,x_2)\sigma_1(x_1,x_2)$ och $\sum_\mu S_\mu$ om λ är

Det första vi gör är att skriva upp de möjliga numreringarna av λ och motsvarande monom.

1	1
2	

Motsvarande monom är $(x_1^2x_2)$.

1	2
2	

Motsvarande monom är $(x_1x_2^2)$.

Alltså är

$$S_\lambda(x_1,x_2)=x_1^2x_2+x_1x_2^2$$

och

$$\begin{aligned}&S_\lambda(x_1,x_2)\sigma_1(x_1,x_2)=(x_1^2x_2+x_1x_2^2)(x_1+x_2)=\\&=x_1^3x_2+2x_1^2x_2^2+x_1x_2^3.\end{aligned}$$

Vi beräknar nu $\sum_{\mu} S_{\mu}(x_1, x_2)$ och får två diagram

μ_1

och

μ_2

Vi beräknar numreringarna av μ_1 samt motsvarande monom och får

1	1	1
2		

Monomet är $(x_1^3 x_2)$

1	1	2
2		

Monomet är $(x_1^2 x_2^2)$

1	2	2
2		

Monomet är $(x_1 x_2^3)$.

Alltså är $S_{\mu_1}(x_1, x_2) = x_1^3 x_2 + x_1^2 x_2^2 + x_1 x_2^3$.

Det finns endast en numrering av μ_2 dvs

1	1
2	2

Och motsvarande Schurpolynom är

$$S_{\mu_2}(x_1, x_2) = x_1^2 x_2^2.$$

Alltså är

$$S_{\mu_1}(x_1, x_2) + S_{\mu_2}(x_1, x_2) = x_1^3 x_2 + 2x_1^2 x_2^2 + x_1 x_2^3$$

vilket är samma resultat som vi ovan fick av

$$S_\lambda(x_1, x_2)\,\sigma_1(x_1, x_2).$$

Lemma 2

Om λ är ett Young diagram med partitionen

$$\lambda_1 + \ldots + \lambda_l$$

så är

$$\sigma_{\lambda_1} \cdot \sigma_{\lambda_2} \cdot \ldots \cdot \sigma_{\lambda_l} = \sum K_{\mu\lambda} s_\mu$$

där vi summerar över alla partitioner μ och där $K_{\mu\lambda}$ är antalet följder

$$\mu^{(1)} \subset \mu^{(2)} \subset \ \ldots \subset \mu^{(l)} = \mu$$

sådana att μ är en kolumn med λ_1 celler och $\mu^{(i)} / \mu^{(i-1)}$ är λ_i celler där två inte ligger i samma rad.

Satsen följer nu ur dessa båda lemma på följande sätt:

Om vi ordnar diagrammen i lexikografisk ordning så gäller

$$K_{\mu\lambda} = 0 \text{ om } \lambda > \mu$$

och

$$K_{\mu\lambda} = 1.$$

Detta innebär att matrisen $K_{\mu\lambda}$ är triangulär och vi kan alltså lösa ut s_μ uttryckt i $\sigma_{\mu_1}, \ldots, \sigma_{\mu l}$. Alltså är s_μ symmetriskt. För detaljerna i bevisen, som bl.a. bygger på utstötningsalgoritmen, hänvisas till Fulton [1].

4. Bevis av att π är transcendent

Vi skall nu använda huvudsatsen för symmetriska funktioner till att visa att π är transcendent. Det första beviset gavs 1882 av CL. F Lindemann (1852-1939).

Sats 4.

π är transcendent

Bevis

Antag att motsatsen gäller, att π är noll för ett icke-negativt polynom över Q. Då är $i\pi$, där $i = \sqrt{-1}$ också är ett nollställe till ett polynom i $Q[x]$. Låt $\theta_1(x) \in Q[x]$ vara ett polynom med nollställen $\alpha_1 = i\pi, \alpha_2, \ldots, \alpha_n$ Ett välkänt teorem av Euler ger oss

$$e^{i\pi} + 1 = 0$$

så att

$$(e^{\alpha_1} + 1)(e^{\alpha_2} + 1)\ldots(e^{\alpha_n} + 1) = 0 \qquad (1)$$

Vi konstruerar nu ett polynom med heltalskoefficienter där nollställen är de exponenter $\alpha_{i_1} + \ldots + \alpha_{i_r}$ som uppträder i utvecklingen av produkten (1). Till exempel ger termer på formen

$$e^{\alpha_s} \times e^{\alpha_t} \times 1 \times 1 \times 1 \ldots 1$$

exponent $\alpha_s + \alpha_t$. Tas de parvis över s,t får vi $\alpha_1 + \alpha_2, \ldots, \alpha_{n-1} + \alpha_n$. De elementära symmetriska polynom i dessa variabler är symmetriska i $\alpha_1, \ldots, \alpha_n$. Så dessa kan med hjälp av sats 2 uttryckas som polynom i den elementära symmetriska polynomer av $\alpha_1, \ldots, \alpha_n$. Härav följer att paren $\alpha_s + \alpha_t$ satisfierar en polynom ekvation $\theta_2(x) = 0$ där θ_2 har rationella koefficienter. Alltså

$$\theta_1(x)\theta_2(x)\ldots\theta_n(x)$$

är ett polynom över Q vars nollställen är exponenter till *e* i utvecklingen av ekvation (1). Här är θ_j det polynom som har nollställena $\alpha_{v_1} + \ldots + \alpha_{v_j}$. Delar man med en lämplig potens av *x* och multiplicerar med ett lämpligt heltal får vi ett polynom $\theta(x)$ över Z, vars nollställen är icke-negativa exponenter $\beta_1, \ldots, \beta_r$ till *e* i utvecklingen av ekvationen (1). Nu kan ekvation (1) skrivas enligt

$$e^{\beta_1} + \ldots + e^{\beta_r} + e^0 + \ldots + e^0 = 0,$$

det vill säga

$$(2) \qquad e^{\beta_1} + \ldots + e^{\beta_r} + k = 0\,,$$

där $k \in$ Z. Termen $1 \cdot 1 \cdot \ldots \cdot 1$ finns i utvecklingen så att $k > 0$.
Antag alltså att

$$\theta(x) = cx^r + c_1 x^{r-1} + \ldots + c_r$$

där $c, c_1, \ldots, c_r$ är heltal. Nu är $c_r \neq 0$ då 0 inte är ett av $\theta : s$ nollställen.

Definiera

$$f(x) = \frac{c^s x^{p-1}\{\theta(x)\}^p}{(p-1)!},$$

där $s = rp - 1$ och p är något godtyckligt primtal. Observera att $f(x)$ har graden $p - 1 + rp = p + s$ .Sätt också

$$F(x) = f(x) + f'(x) + \ldots + f^{(s+p+r-1)}(x).$$

Notera att $f^{(s+p+r)}(x) = 0$ eftersom $s + p + r \geq$ grad $f(x)$. Vi har

$$\frac{d}{dx}\{e^{-x}F(x)\} = -e^{-x} f(x),$$

ty

$$\frac{d}{dx}\left(e^{-x}F(x)\right) = -e^{-x}F'(x) - e^{-x}F(x) =$$
$$= e^{-x}\left(F'(x) - F(x)\right) = e^{-x}\left(f'(x) + f''(x) + \ldots + f^{(s+p+r)}(x) - f(x) - f'(x) - \ldots - f^{(s+p+r-1)}(x)\right)$$
$$= -f(x)e^{-x}$$

Om vi integrerar båda leden får vi

$$e^{-x} F(x) - F(0) = -\int_0^x e^{-y} f(y)dy$$

Sätt $y = \lambda x$ och multiplicera med e^x då får vi

$$F(x) - e^x F(0) = -x\int_0^1 \exp[(1-\lambda)x]\, f(\lambda x)d\lambda$$

Sätt x lika med $\beta_1, \ldots, \beta_r$, summera och använd ekvation (2). Vi får

$$(3) \qquad \sum_{j=1}^{r} F(\beta_j) + kF(0) = -\sum_{j=1}^{r} \beta_j \int_0^1 \exp[(1-\lambda)\beta_j] f(\lambda\beta_j) d\lambda .$$

Vi påstår att vänster ledet av ekvation (3) är ett icke-negativt heltal för alla tillräckligt stora p. Eftersom β_j är nollställe av multipliciteten p till $f(x)$ så är

$$\sum_{j=1}^{r} f^{(t)}(\beta_j) = 0 ,$$

om $0 < t < p$. Varje derivata $f^{(t)}(\beta_j)$ med $t \geq p$ innehåller en faktor p därför att vi måste derivera $\{\theta(x)\}^p$ minst p antal gånger för att erhålla en term $\neq 0$. För varje $t \geq p$ är

$$\sum_{j=1}^{r} f^{(t)}(\beta_j) ,$$

ett symmetriskt polynom i β_j av grad $\leq s$. Alltså är det ett polynom av grad $\leq s$ i koefficienterna c_i / c enligt sats 2 . Faktorn c^s i definitionen av $f(x)$ gör det till ett heltal. Så för $t \geq p$ gäller

$$\sum_{j=1}^{r} f^{(t)}(\beta_j) = pk_t$$

för lämpligt $k_t \in Z$. Nu tittar vi på $F(0)$. Vi har

$$f^{(t)}(0) = \begin{cases} 0 & (t \leq p-2) \\ c^s c_r^p & (t = p-1) \\ l_t p & (t \geq p) \end{cases}$$

för lämpligt $l_t \in Z$. Följaktligen är vänster led av ekvation (3)

$$Kp + kc^s c_r^p$$

för något $K \in Z$. Nu är $k \neq 0, c \neq 0$ och $c_r \neq 0$. Om vi tar

$$p > \max (k, |c|, |c_r|$$

då är vänster ledet av ekvation (3) ett icke-negativt heltal icke delbart med p.

Sista delen av beviset är att vi beräknar storleken av höger ledet i ekvation (3). Då

$$\left| f(\lambda\beta_j \right| \leq \frac{|c|^s |\beta_j|^{p-1} (m(j))^p}{(p-1)!}$$

om $0 \leq \lambda \leq 1$ och

$$m(j) = \sup_{0\leq\lambda\leq 1} \left|\theta(\lambda\beta_j)\right|$$

är

$$\left| -\sum_{j=1}^{r} \beta_j \int_0^1 \exp[(1-\lambda)\beta_j] f(\lambda\beta_j) d\lambda \right| \leq \sum_{j=1}^{r} \frac{|\beta_j|^p |c^s| |m(j)|^p B}{(p-1)!}$$

där

$$B = \left| \max_j \int_0^1 \exp[(1-\lambda)\beta_j] d\lambda \right| .$$

Uttrycket går mot 0 om p går mot ∞. Alltså får vi en motsägelse, vilket i sin tur betyder att π är transcendent.

5. Litteratur

Dickson, Eugene Leonard, 1992: *Theory of numbers*. New York: Chelsea Publishing Company.

Fulton, William, 1997: *Young tableaux*. New York: Cambridge University Press.

Nagell, Trygve, 1949: *Lärobok i algebra*. Uppsala: Almqvist & Wiksells.

Stewart, Ian, 1989: *Galois theory*. New York: Chapman and Hall.