# SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

## Kvantifikatorelimination

av

**Gustav Yilbar Kjellström**

2020 - No K39

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

# Kvantifikatorelimination

Gustav Yilbar Kjellström

---

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Torbjörn Tambour

2020

# Abstract

Syftet med detta arbete är att undersöka hur man eliminerar kvantifikatorer med hjälp av en generalisering av Sturms sats, samt undersöka hur Mathematica klarar av att eliminera kvantifikatorer. Detta arbete inkluderar ett bevis av Sturms sats. Denna sats kan man använda för att bestämma antalet reella nollstället till ett polynom, men fokuset ligger på en generalisering av Sturms sats som säger att man kan skriva om ett uttryck till ett ekvivalent kvantifikatorfritt uttryck och därigenom eliminera alla kvantifikatorer. Slutligen undersöks Mathematicas begränsningar samt hur programmet svarar när man använder dess inbygga funktion *resolve* för att eliminera kvantifikatorer. Detta undersöks på några geometriska former som tex skärning mellan två linjer.

# Tack

Jag vill tacka min handledare Torbjörn Tambour som med tålamod stöttat och hjälpt mig - även vid kort varsel. Tack också för att du tog över handledarskapet när min tidigare handledare Erik Palmgren gick bort. Mina tankar går till Erik och hans anhöriga.

# Kvantifikatorelimination

Gustav Yilbar Kjellström

## Innehåll

# 1 Inledning

Logik har inte alltid varit en del av det matematiska språket. Leibniz började strukturera fram logiken redan på 1600-talet. Han levde i en tid med moderna matematiska notationer, främst inom algebra och analys. Leibniz strävade efter att formulera om matematikens regler så att axiom, satser och definitioner skulle kunna uttryckas med hjälp av matematiska symboler. Med detta ville han göra det enklare att klargöra matematiska bevis och resonemang. Leibniz var före sin tid med detta, men hans arbete lade grunden för den moderna logiken som växte fram under 1900-talet.

# 2 Bakgrund

Tarskis geometri är baserad på elementär geometri. Elementär geometri formulerades av matematikern Euklides ca 300 år f.Kr. Detta publicerade han bland annat i sitt stora verk Elementa. Elementär geometri är uppbyggd av första ordningens logik. Detta betyder att alla variabler $x, y, z, ...$ betraktas som punkter som kan anta alla värden i ett intervall samt vissa logiska symboler, dessa symboler är en del av satslogiken.

## 2.1 Satslogik

Med satslogik kan man med hjälp av olika påståenden dra korrekta slutsatser. Tex om jag tittar på min klocka och ser att klockan är 7 så är klockan antingen 7 eller så går min klocka fel. Detta kan skrivas med hjälp av logiska symboler. Vi börjar med att definiera dem.

**Definition 1.** Logiska symboler:
$\wedge$ motsvarar *och*, påståendet $A \wedge B$ är sant om både A och B är sanna. I annat fall är påståendet falskt.
$\vee$ motsvarar *eller*, påståendet $A \vee B$ är sant ifall A eller B är sant och är endast falskt ifall både A och B är falska.
$\supset$ är ett tecken för *medför* vilket är samma sak som $\rightarrow$ exempelvis $x = 3 \supset x^2 = 9$
$=$ vanliga identitets symbolen
$\neq$ beskriver *inte lika*.
$\equiv$ är ett tecken för *ekvivalens* tex $x \equiv y$ innebär att $x$ är ekvivalent med $y$.
$\neg$ innebär negation vilket motsvarar ordet *inte*. Är $P$ sant så är $\neg$P falskt.

Vi kan skriva det tidigare uttrycket med bokstäver.
$A$ : Jag ser att klockan är 7
$B$ : Min klocka går fel
Med dessa definitioner kan man dra logiska slutsatser, tex $\neg B \rightarrow A$. Om min klocka *inte* går fel, alltså om klockan går rätt så är klockan 7. $A \vee B$ säger att antingen är klockan 7 eller att min klocka går fel.

## 2.2 Första ordningens logik

Predikatlogik bygger på Satslogiken men man har lagt till kvantifikatorer, *för alla* och *det existerar*

**Definition 2.** Kvantifikatorer:
$\forall x$ (för alla) innebär att uttrycket ska gälla för alla $x$.
$\exists x$ (existerar) innebär att uttrycket ska gälla för minst ett $x$.

Om $A$ säger att talet är större än 10. Då säger $\forall x A(x)$ att alla $x$ har egenskapen $A$, alltså att alla $x$ är större än 10. $\exists x A(x)$ innebär att minst ett $x$ är större än 10. Dessa logiska symboler bygger tillsammans med punkter upp första ordningens logik och med denna definition är bara punkter klassade som första ordningens variabler.

## 2.3 Tarskis geometri

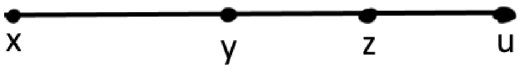
Utifrån dessa punkter kan man skapa olika geometriska figurer tex linjer, trianglar, cirklar, kvadrater osv. Varje geometrisk figur definieras av ett fixt antal punkter. Detta ger oss möjligheten att definiera $\beta$ som betyder *mellan.* Att $z$ ligger på en linje mellan $x$ och $y$ skriver vi $\beta(x, z, y)$ Vi betecknar avståndet med $\gamma$. $\gamma(x, y, u, v)$ betecknar avståndet mellan $x$ och $y$ samt avståndet mellan $u$ och $v$. Den elementära geometrin byggs upp av 13 axiom. Jag kommer att skriva ner de första tre, vill man läsa om de andra tio axiomen kan man göra det i *What is elementary geometry? av Alfred Tarski.*

**Axiom 1.** $\forall xy[\beta(x, y, x) \rightarrow x = y]$

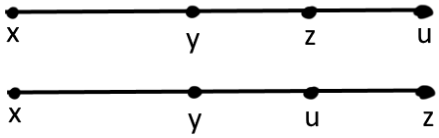
**Axiom 2.** $\forall xyzu[\beta(x, y, u) \wedge \beta(y, z, u) \rightarrow \beta(x, y, z)]$

**Axiom 3.** $\forall xyzu[\beta(x, y, z) \wedge \beta(x, y, u) \wedge x \neq y \rightarrow \beta(x, z, u) \vee \beta(x, u, z)]$

I axiom 1 har vi en linje som börjar i $x$ och slutar i $x$. Då måste $y = x$. I axiom 2 har vi en linje där y ligger mellan $x$ och $u$ samt att på samma linje ligger $z$ mellan $y$ och $u$. Då måste $y$ ligga mellan $x$ och $z$ (se fig. 1) Axiom 3 säger att vi har en linje med punkterna $x, y, z$ och $x, y, u$. Vi i kan inte avgöra vilken av punkterna $u$ och $z$ som kommer först (ligger närmast $x$) därför får vi två möjligheter (se fig. 2 )




Figur 1:




Figur 2:

# 3 Eliminationsprocessen

Kvantifikator elimininationen innebär att för en formel $A$ så ska vi hitta en kvantifikator fri formel $B$ som är bevisbar.

## 3.1 Exempel på elimination

**Exempel 3.1.** $\forall x(ax + b > 0)$
Här har vi ett uttryck som säger att för alla $x$ ska $ax + b > 0$. Vi vill eliminera kvantifikatorn $\forall$ (för alla). I detta exempel har vi två fria variabler $a$ och $b$. Detta betyder att de kan anta alla värden och vi behöver sätta villkor på dem för att eliminera kvantifikatorn.

I vissa fall finns det inga fria variabler (variabler som kan anta vilka värden som helst). I dessa fall blir resultatet antingen sant eller falskt.

**Exempel 3.2.** $\exists x(4x^2 + 5 > 10)$
I detta exempel blir resultatet av en kvantifikatorelimination sant, eftersom det finns minst ett $x$ som uppfyller villkoret och det finns inga fria variabler.

Kommande sats och beviset till satsen är inspirerat från *Foundations of Mathematics* av *Erwin Engeler*

**Sats 3.1** (kvantifikatorelimination)**.** *Givet en första ordningens logik Antagande:För varje formel A på formen* $\exists x(A_1(X) \wedge ... \wedge A_n(x))$ *existerar en kvantifikatorfri formel B sådan att A. Genom att genomföra följande steg kan man eliminera kvantifikatorerna.*

1. *Ändra de innersta kvantifikatorerna till existenskvantifikatorer om de inte redan är det.*

2. *Se till att dessa uttryck är på disjunktiv normalform*

3. *Eliminera dessa genom att använda antagande*

4. *Ifall det resulterade uttrycket är kvantifikatorfritt, evaluera det som sant eller falskt, annars gör om processen från steg 1.*

Formel betyder att vi har en funktion som kan innehålla logiska symboler. Satsen säger att vi kan skriva om formeln $A$ som innehåller kvantifikatorer till ett ekvivalent uttryck $B$ som inte innehåller några kvantifikatorer.
Steg 1 är att kontrollera så att den innersta kvantifikatorn är $\exists$. Om det är $\forall$ så måste man skriva om uttrycket genom att skriva om för alla kvantifikatorn på följande sätt.

$$\forall x P(x) = \neg\exists\neg P(x)$$

Detta betyder att det inte existerar något $x$ sådant att $P(x)$ inte uppfylls, vilket är ekvivalent med att för alla $x$ så uppfylls $P(x)$.

Ett uttryck som är på disjunktiv normalform är skriven på formen $A_1 \vee ... \vee A_l$ där varje är $A$ är på formen $S_1 \wedge ... \wedge S_i$ och varje $S$ är en atom eller en negerad atom. Man kan sammanfatta det förenklat som att man inte får ha nästlade $\neg$ eller $\vee$. Exempel på en formel som är skriven på disjunktiv normalform är $(A \wedge B) \vee C$. Ett exempel på en formler som inte är skriven på disjunktiv normalform är $(A \wedge (B \vee C))$ eller $\neg(A \vee B)$. Jag kommer nu att gå djupare in på steg 3 där man eliminerar kvantifikatorn. Det finns flera sätt att eliminera kvantifikatorer. Jag kommer gå igenom ett sätt som bygger på en generalisering av Sturms sats samt elementär teori för reellt slutna kroppar. Vi börjar med Sturms sats för att sedan gå vidare till generaliseringen av satsen. Beviset för Sturms sats är inspirerat ifrån *lärobok i algebra* av *Trygve Nagell*

## 3.2 Sturms sats

Med Sturms sats kan man bestämma antalet reella nollställen med hjälp av Euklides algoritm som är en metod att bestämma största gemensamma delare av två polynom. Vi kan skriva om $f(x)$ och dess derivata till en produkt av två polynom samt en rest.

$$\begin{aligned} f(x) &= f'(x)g_1(x) - f_2(x) \\ f'(x) &= f_2(x)g_2(x) - f_3(x) \\ f_2(x) &= f_3(x)g_3(x) - f_4(x) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_{m-2}(x) &= f_{m-1}(x)g_{m-1} - f_m(x) \\ f_{m-1}(x) &= f_m(x)g_m(x) \end{aligned} \tag{1}$$

För att använda Sturms sats för att hitta polynomets nollställen så studerar man resttermerna och om man sätter $f'(x) = f_1(x)$ kan man skriva det som en kedja.

$$f(x), f_1(x), f_2(x), f_3(x), ...., f_m(x).$$

Denna kedja kallas den sturmska kedjan. Genom att studera teckenändringar i denna kedja kan man studera antalet nollställen.

Om $b$ och $c$ är reella tal, $b < c$ och $f(b) \neq 0$ samt $f(c) \neq 0$ är differensen mellan antalet teckenväxlingar i följderna

$$f(b), f_2(b), ..., f_{m-1}(b), f_m(b)$$

$$f(c), f_2(c), ..., f_{m-1}(c), f_m(c)$$

lika med antalet reella nollställen i intervallet $[b, c]$.

Vill man ha antalet nollställen i hela intervallet så jämför man antalet teckenförändringar när man sätter in $+\infty$ samt $-\infty$ i alla funktioner och studerar antalet förändringar i varje kedja. Skillnaden i antalet förändringar ger antalet nollställen.

### 3.2.1 Bevis för Sturms sats

För att bevisa Sturms sats behöver vi gå igenom några fakta som används i beviset. Vi antar att vi har ett polynom $f(x)$ med enbart enkla nollställen alltså inga multipla rötter. Att funktionen enbart har enkla nollställen innebär att $f(x)$ och $f'(x)$ inte har några gemensamma nollställen heller. Detta kan vi se om vi antar att $f(c) = 0$ alltså att funktionen har ett nollställe i $x = c$. Då kan vi skriva om $f(x)$ med hjälp av faktorsatsen till $f(x) = (x - c)g(x)$ för något polynom $g(x)$ och eftersom $f(x)$ inte har några multipla rötter så är $g(c) \neq 0$. Om vi deriverar $f(x)$ får vi följande:

$$f'(x) = g(x) + (x - c)g'(x)$$

Nu kan vi se att $f'(c) = g(c) \neq 0$, därav kan inte $f(x)$ och $f'(x)$ ha några gemensamma nollställen.

Det andra vi behöver etablera är att om $f(x)$ har ett enkelt nollställe i $c$ så växlar $f(x)$ tecken vid $c$. Vi såg nyss att vi kan skriva $f(x) = (x - c)g(x)$ där $g(c) \neq 0$. Om vi antar att $g(c) > 0$ då måste även $g(x) > 0$ i någon omgivning $[c - h, c + h]$. För $c - h < x < c$ så är $f(x) = (x - c)g(x) < 0$ och för $c < x < c + h$ så måste $f(x) > 0$. Alltså byter $f(x)$ tecken vid $c$. Detta gäller endast funktioner med enkla nollställen, tex $f(x) = x^2$ byter inte tecken i $x = 0$. att denna funktion inte byter tecken beror på att den har en dubbelrot.

Med detta i åtanke kan vi nu börja med beviset av Sturms sats. Vi börjar med att skriva Euklides algoritm för $f(x)$ och $f'(x)$.

$$\begin{aligned} f(x) &= f'(x)q_1(x) - f_2(x) \\ f_1(x) &= f_2(x)q_2(x) - f_3(x) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_{m-1}(x) &= f_m(x)q_m - f_{m+1}(x) \end{aligned} \tag{2}$$

Här är $f'(x) = f_1$ och $f_{m+1}$ konstanta polynom. Två konsekutiva polynom $f_i(x), f_{i+1}$ kan inte ha samma nollställe $c$ för då skulle de tidigare polynomen försvinna för det värdet och $f(c) = f'(c)$. Vilket är en motsägelse.
Vi antar nu att ekvationen $f(x) = 0$ har en rot $c$. Har vi ett litet intervall omkring $x = c, [c-h, c+h]$ har $f(x)$ och $f_1(x)$ samma tecken före men motsatt tecken efter. Alltså kan endast en teckenväxling ske i det fall då ett av polynomen $f_i$ passerar genom noll. Om $f_i(c) = 0$ blir $f_{i-1}(c) = -f_{i+1}(c)$ och dessa är $\neq 0$. Detta utifrån de fakta vi etablerade innan beviset. I ett litet intervall kommer $f_{i-1}(x)$ och $f_{i+1}(x)$ att behålla sina motsatta tecken och detta medför att exakt en teckenväxling sker i intervallet. Eftersom $f_{i-1}(x)$ och $f_{i+1}(x)$ behåller sina motsatta tecken så ändras inte antalet teckenväxlingar för $f_i(x), 1 \leq i < m$. Då $f_m$ är konstant kan en förändring i antalet teckenväxlingar endast sker när $x$ passerar ett nollställe till $f(x)$. Vi kan göra en tabell för att se detta tydligare. Vi antar att $f_{i+1} < 0$ i någon omgivning $[c-k, c+k]$ och då får vi teckenschemat

| | | $c$ | |
|---|---|---|---|
| $f_{i-1}(x)$ | + | + | + |
| $f_i(x)$ | + | 0 | - |
| $f_{i+1}(x)$ | - | - | - |
| Antal teckenväxlingar:$V$ | 1 | | 1 |

Vi ser att antalet teckenväxlingar inte ändras då $x$ passerar $c$ utan att det endast är teckenväxling före och efter. Om vi istället studerar vad som händer när $x$ passerar nollstället $c$ till $f(x)$ och om vi antar att $f(x)$ växlar tecken från positivt till negativt vid $c$ samt att i en omgivning av $c$ så är $f'(x) < 0$. Då får vi tecken schemat.

| | | c | |
|---|---|---|---|
| $f(x)$ | + | 0 | - |
| $f'(x)$ | - | - | - |
| Antal teckenväxlingar:V | 1 | | 0 |

Här ser vi att antalet teckenväxlingar ändras med 1 och vi ser att antalet nollställen till $f(x)$ kan skrivas som $V(a) - v(b)$ i intervallet $[a, b]$

**Exempel 3.3.** Bestäm antalet nollställen till funktionen $f(x) = x^4 + x^3 - 7x^2 - x + 6$ genom att använda Sturms sats.
Vi löser detta genom att skriva upp Euklides algoritm genom att använda polynomdivision.

$$f(x) = x^4 + x^3 - 7x^2 - x + 6 = (4x^3 + 3x^2 - 14x - 1)(\frac{x}{4} + \frac{1}{16}) - (\frac{59}{16}x^2 - \frac{x}{8} - \frac{97}{16})$$
$$f'(x) = 4x^3 + 3x^2 - 14x - 1 = (\frac{59}{16}x^2 - \frac{x}{8} - \frac{97}{16}) * (\frac{64}{59}x + \frac{2960}{3481}) - (\frac{25472}{3481}x - \frac{14464}{3481})$$
$$f_2 = (-\frac{59}{16}x^2 - \frac{x}{8} - \frac{97}{16}) = (\frac{25472}{3481}x - \frac{14464}{3481})(-\frac{205379}{407552}x - \frac{21822389}{81102848}) - (\frac{783225}{158404}) \quad (3)$$

Resttermerna (sista termerna) skapar den sturmska kedjan

$$\begin{aligned} f(x) &= x^4 + x^3 - 7x^2 - x + 6 \\ f'(x) &= 4x^3 + 3x^2 - 14x - 1 \\ f_2 &= \frac{59}{16}x^2 - \frac{x}{8} - \frac{97}{16} \\ f_3 &= \frac{25472}{3481}x - \frac{14464}{3481} \\ f_4 &= \frac{783225}{158404} \end{aligned} \tag{4}$$

För att ta reda på antalet nollställen behöver vi studera teckenväxlingarna i den sturmska kedjan i ändpunkterna på intervervallet. I vårt fall vill vi studera hela intervallet, $-\infty, +\infty$. Men det kan även vara intressant att ta reda på hur många av rötterna som är positiva samt negativa, detta kan göras genom att studera antal teckenväxlingar i punkten 0. Nedan är en teckentabell av den sturmska kedjan.

| | $+\infty$ | 0 | - $\infty$ |
|---|---|---|---|
| $f(x)$ | + | + | + |
| $f'(x)$ | - | - | + |
| $f_2(x)$ | + | - | + |
| $f_3(x)$ | - | - | + |
| $f_4(x)$ | + | + | + |
| Antal teckenväxlingar:V | 4 | 2 | 0 |

Antalet teckenväxlingar i hela intervallet blir $V(+\infty) - V(-\infty) = 4 - 0 = 4$ Alltså har vi 4 nollställen. VI kan även studera antalet positiva samt negativa nollställen genom att jämföra ändpunkterna med 0. $V(+\infty) - V(0) = 4 - 2 = 2$ samt $V(0) - V(-\infty) = 2 - 0 = 2$ Vi har således 2 positiva samt 2 negativa nollställen.

Detta är ett sätt att formulera Sturms sats på. Man kan formulera Sturms sats på flera sätt. Nedan är ett annat sätt att formulera Sturms sats på som är mera generell.

**Sats 3.2** (Sturms sats)**.** *För varje polynom $p(x, x_1, ..., x_n)$ med heltals koefficienter, så existerar en kvantifikatorfri formel $B(x, x_1, ..., x_n, a, b)$ sådan att:*

$$a < b \supset .B(x_1, ..., x_n, x_n) \equiv \exists x(a \leq x \leq b \wedge p(x, x_1, x, ..., x_n) = 0)$$

**Sats 3.3.** *Generalisering av Sturms sats*
*För varje kvantifikatorfri formel $A(x, x_1, ..., x_n)$ finns det en kvantifikatorfri $B(x, x_1, ..., x_n, a, b)$ sådan att*

$$a < b \supset .B(x_1, ..., x_n, a, b) \equiv \exists x(a < x < b \wedge A(x, x_1, ..., x_n))$$

Nu gäller inte Sturms sats för polynom utan även varje kvantifikatorfri formel $A(x_1, x_2, ..., x_n)$. Jag kommer att gå igenom beviset för generaliseringen av Sturms sats. Vi kommer att se att vi kan skriva om högerledet som innehåller kvantifikatorn $\exists$, till ett kvantifikatorfritt uttryck $B$. $B$ är ett villkor på koefficienterna $a$ och $b$ som garanterar att varje $p$ har minst ett nollställe mellan $a$ och $b$. Varje $p$ här betraktas som en funktion av $x_1, ..., x_n$. Satsen är bevisbar i elementär teori för reella tal.

**Sats 3.4.** *Elementär teori för reellt slutna kroppar*
*Språk och logik: Första ordningens predikat kalkyl med likhet; individuella variabler:* $x, y, z, ...$; *individuella konstanter:* $0, 1$; *funktionssymboler;*$+, \cdot, -, ^{-1}$; *standard predikat* $\leq$.
*Axiom*

*(i)* *Axiom för kroppar*

*(ii)* *order axiom*

*(iii)* $\forall x \exists y(x = y^2 \vee -x = y^2)$

*(iv)* *För varje naturligt tal n:*
$\forall x_0 \forall x_1 ... \forall x_{2n} \exists y(x_0 + x_1 y + x_2 y^2 + ... x_{2n} y^(2n) + y^{2n+1} = 0$

## 3.3 Bevis för generalisering av Sturms sats

**Lemma 3.1.** *Låt* $p_1, ..., p_k, q_1, ..., q_l$ *vara polynom i* $x, x_1, ..., x_n$ *med heltalskoefficienter. Då är* $p_1 = 0 \wedge p_2 = 0 \wedge ... \wedge p_k = 0 \wedge q_1 > 0 \wedge ... \wedge q_l > 0$ *ekvivalent med en kvantifikatorfri formel vars grad i* $x$ *är mindre än graden av* $p_i$ *med avseende på* $x$ *för varje polynom* $p_i$.

*Bevis.* Beviset bygger på att reducera summan $h$ som är summan av graden av $p_i$ och $q_j$. h kan därav skrivas så här: $h = deg\, p_1 + ... + deg\, p_k + deg\, q_1 + ... + deg\, q_l$ Beviset använder induktion över $h$. Ifall $h = 0$ har vi inget att visa då graden är noll. Detta är vår induktionsbas. Induktionsantagandet är att påståendet är sant då den totala graden är $< h$. När $h > 0$ får vi olika fall beroende på vad k är. När $k = 0$ har vi inget att visa, då vi inte har några $p$ termer. Om $k = 1$ så har vi endast en $p$ term, men vi har fortfarande $l$ st $q$ termer. Vi kan skriva det på denna form:

$$p = 0 \wedge q_1 > 0 \wedge q_2 > 0 \wedge ... \wedge q_l > 0 \tag{5}$$

Däremot blir det endast intressant ifall graden av $q_i \geq$ graden av $p$. För att minska graden på $q_i$ gör vi ett variabelbyte där $p = ax^m + ...$, och $q = bx^n + ...$ där $a, b \neq 0$ och $m \leq n$. Vi sätter $Q = a^2 q_1 - abx^{n-m} p(x)$ Då kan man skriva om ekvation (5) till:

$$p = 0 \wedge Q > 0 \wedge q_2 > 0 \wedge ... \wedge q_l > 0 \tag{6}$$

Det viktiga här är att $Graden(Q) < graden(q_1)$ Därmed har vi även reducerat graden av h. När vi studerar $k \geq 2$ sätter vi $p_1 = a_1 x^{m_1} + ..., p_2 = a_2 x^{m_2} + ...$

där $a_1, a_2 \neq 0$ och $m_1 \geq m_2$. Nu låter vi $P = a_2p_1 - a_1x^{m_1-m_2}p_2$ och får nu på exakt samma sätt som i fallet $k = 1$ att:

$$P = 0 \wedge p_2 = 0 \wedge ... \wedge p_k = 0 \wedge q_1 > 0 \wedge ... \wedge q_l > 0 \tag{7}$$

Nu har vi lyckats minska graden av $h$. □

**Lemma 3.2.** *Låt $A(x, x_1, ..., x_n)$ vara en kvantifikatorfri formel av grad $h$ i $x$. Välj $a$ och $b$ så att de är parametrar och att de är distinkta från $x, x_1, ...x_n$. Då existerar en kvantifikatorfri formel $B(x_1, ..., x_n, a, b)$ som uppfyller*

$$a < b. \supset B(x_1, ..., x_n, a, b) \equiv \exists x(a < x < b \wedge A(x, x_1, ..., x_n)) \tag{8}$$

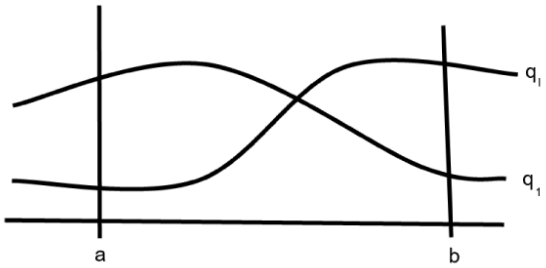
*Graden av $B$ i $a, b$ är begränsad av $h + 1$.*

*Bevis.* Vi behöver återigen dela upp det hela i olika fall. Om $h = 0$ behöver vi inte göra något då variabeln $x$ inte förekommer i $A$ och vi kan välja $B$ till $A$. Om $h > 0$ kan vi skriva om $A$ på formen:

$$p_1 = 0 \wedge p_2 = 0 \wedge ... \wedge p_k = 0 \wedge q_1 > 0 \wedge ... \wedge q_l > 0 \tag{9}$$

Nu argumenterar vi beroende på vad $k$ är. Om $k = 0$ har $A$ formen:$q_1 > 0 \wedge ... \wedge q_l > 0$ Ett påstående av denna form $\exists x(a < x < b \wedge q_1 > 0 \wedge ... \wedge q_l > 0)$ medför att någonstans mellan $a$ och $b$ måste alla polynom $q_1, ..., q_l$ vara strikt positiva. Detta kan ske på tre olika sätt. Första sättet är ifall alla polynom är strikt större än 0 för alla $x$ i intervallet $a, b$, se figur 3, vilket går att skriva på detta sätt:

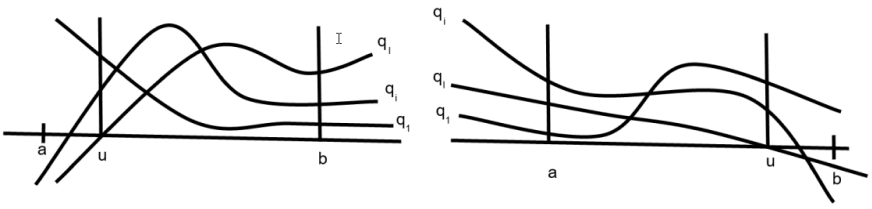
$$G_0(a, b) \equiv \forall x(a < x < b \supset .q_1 > 0 \wedge ... \wedge q_l > 0) \tag{10}$$




Figur 3:

I det andra fallet är inte alla polynom strikt större än 0. Men det finns ett i,$1 \leq i \leq l$ sådan att ett polynom skär $x$-axeln i punkten $v$, medan alla andra polynom är positiva i intervallet $a, v$ eller $v, b$ se figur 4. Då kan vi teckna ett uttryck $G_i$ såhär

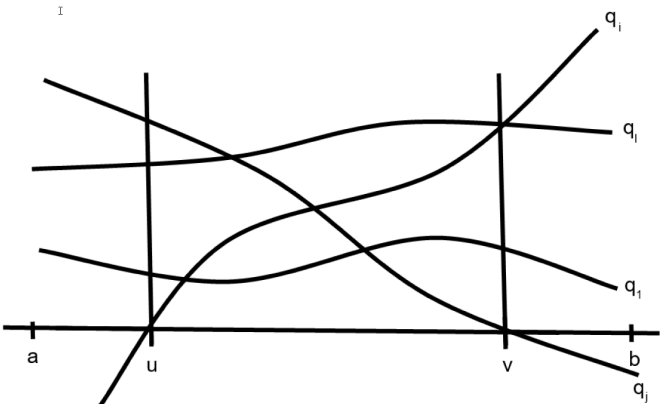
$$G_i(a, b) \equiv \exists v(a < v < b \wedge q_i(v) = 0 \wedge G_0(a, v)) \vee \exists v(a < v < b \wedge q_i(v) = 0 \wedge G_0(v, b)) \tag{11}$$




Figur 4:

I det tredje fallet har vi ett intervall mellan $u$ och $v$ där $(u, v) \subseteq (a, b)$. I det här delintervallet är alla polynom strikt positiva utom polynomen $q_i$ och $q_j$ som skär x-axeln i punkten $u$ respektive $v$ se figur 5. I detta fall definierar vi $i, j$ såhär: $1 \leq i \leq l$ och $1 \leq j \leq l$ och vi kan teckna ett uttryck $H_{ij}$ som uppfyller detta

$$H_{ij}(a, b) \equiv \exists u \exists v (a < u < v < b \wedge q_i(u) = 0 \wedge q_j(v) = 0 \wedge G_0(u, v)) \tag{12}$$




Figur 5:

Nu kan vi skriva om uttrycket till en kombination utav $G_0, G_i$ och $H_{ij}$

$$\begin{aligned} \exists x(a < x < b \wedge q_1 > 0 \wedge ... \wedge q_l > 0) \equiv G_0(a, b) \vee G_1(a, b) \vee ... \\ \vee G_l(a, b) \vee H_{11}(a, b) \vee ... \vee H_{1l}(a, b) \vee ... \vee H_{ll}(a, b) \end{aligned} \tag{13}$$

Vi börjar med att reducera $G_0(a,b)$

$$\begin{aligned} G_0(a,b) &\equiv \forall a < x < b \supset .q_1 > 0 \wedge ... \wedge q_l > 0 \\ &\equiv \forall x(a < x < b) \supset q_1 > 0 \\ &\wedge \forall x(a < x < b) \supset q_2 > 0 \\ &\vdots \\ &\wedge \forall x(a < x < b) \supset q_l > 0 \end{aligned} \tag{14}$$

Eftersom varje polynom är strikt större än 0 i det öppna intervallet $]a,b[$ måste antingen $q$ eller dess första nollskillda derivata vara positiv i punkten $a$, så länge $q$ inte har några nollställen i intervallet. Detta går att skriva såhär:

$$\begin{aligned} \forall x(a < x < b \supset q_i > 0) \equiv \neg\exists x(a < x < b \wedge q_i = 0 \wedge q_i(a) > 0) \\ \vee(q_i(a) = 0 \wedge q_i\prime(a) > 0) \\ \vee q_i(a) = 0 \wedge q_i\prime(a) = 0 \wedge q_i\prime\prime(a) > 0 \\ \vdots \\ \vee q_i(a) = 0 \wedge q_i\prime(a) = 0) \wedge ... \wedge q_i^{h-2}(a) = 0 \wedge q_i^{h-1}(a) = 0 \end{aligned} \tag{15}$$

Nu kan vi se att för varje $i$, $1 < i < l$ så är graden x i alla dessa formler strikt mindre än $h$. Genom vårt induktionsantagande kan dessa formler bli reducerade och därför kan vi även reducera $G_0(a,b)$ och vi får det på en form som ser ut såhär:

$$\bigwedge_{i=1}^{l}(\neg B_i(x_1,...,x_n,a,b) \wedge K_i(a)) \tag{16}$$

Denna är på formen:

$$K(a) \wedge L(b) \tag{17}$$

Enligt induktionsantagandet är $a$ och $b$ begränsade av $h$.

Nu kommer vi till reduktion av $G_i(a,b), 1 < i < l$. Vi kan skriva om $G_0(a,v)$ och $G_0(v,b)$ på den nya formen som vi har i ekv (14):

$$\begin{aligned} G_0(a,v) &\equiv K(a) \wedge L(v) \\ G_0(v,b) &\equiv K(v) \wedge L(b) \end{aligned} \tag{18}$$

Nu kan vi skriva om $G_i(a,b)$:

$$\begin{aligned} G_i(a,b) \equiv K(a) \wedge \exists v(a < v < b \wedge q_i(v) = 0 \wedge L(v)) \\ \vee\exists v(a < v < b \wedge q_i(v) = 0 \wedge K(v) \wedge L(b)) \end{aligned} \tag{19}$$

Graden av $G_i(a,b)$ är ≤ h. Nu kan vi använda lemma 1 för att minska graden till $h-1$ och då kan vi använda oss av induktionsantagandet och därför kan vi reducera uttrycket.

Nu har vi bara kvar att reducera $H_{ij}(a,b), 1 < i,j < l$

$$H_{ij}(a,b) \equiv \exists u\exists v(a < u < v < b \wedge q_i(u) = 0 \wedge q_j(v) = 0 \wedge G_0(u,v)) \tag{20}$$

Vi kan återigen skriva om $G_0(u,v)$ och då får vi det på den här formen:

$$\begin{aligned} H_{ij}(a,b) \equiv \exists u(a < u < b \wedge q_i(u) = 0 \wedge K(u) \\ \wedge \exists v(u < v < b \wedge q_j(v) = 0 \wedge L(v))) \end{aligned} \tag{21}$$

Vi kan nu använda oss av lemma 1 två gånger. Först kan vi ta bort den innersta kvantifikator eftersom v är bundet av graden h och graden av $(q_j) < h-1$ där av kan vi använda oss av lemma 1. Därefter kan vi använda oss av lemma 1 igen eftersom u också är bunden av h och då kan vi reducera detta till graden av $q-1$ och återigen kan induktions antagandet användas. Alltså kan $h_{ij}$ bli reducerat.

Nu har vi kvar att visa att detta även gäller för $k = 1$ och $k = 2$. Jag kommer inte gå igenom detta, men för $k = 1$ använder man en liknande metod som för $k = 0$, fast man måste ha med derivatorna. Man kommer då att se att man kan reducera A. I det sista fallet $k = 2$ reducerar man uttrycket till fallet k=1.

□

Nu har vi visat att alla uttryck med kvantifikatorer går att skriva om till ett kvantifikatorfritt uttryck. Denna metod fungerar även för att eliminera ett system av ändligt många kvantifikatorer.

$$\exists x_1 \exists x_2, ..., \exists x_n(p_1 = 0 \wedge ... \wedge p_k = 0 \wedge q_1 > 0 \wedge ... \wedge q_l > 0 \tag{22}$$

Genom att genomföra eliminations processen $n$ gånger skapar vi ett ekvivalent kvantifikatorfritt system av polynom, ekvationer och olikheter.

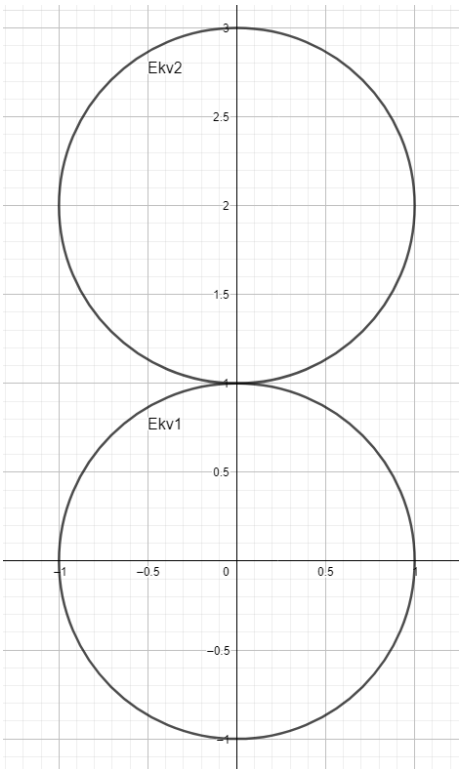
# 4 Tillämpningar

Nedan studeras vad eller om programmet Mathematica klarar av att eliminera kvantifikatorer samt studera hur den uttrycker svaren. Detta görs genom att använda Mathematicas inbyggda funktion *Resolve*. Denna funktion eliminerar kvantifikatorer och skriver om till ett kvantifikator fritt uttryck.

## 4.1 Skärning mellan två cirklar

**Exempel 4.1.** Vi börjar med ett enklare exempel där vi har två cirklar (ekv 1) $x^2 + y^2 = 1$ och (ekv 2) $x^2 + (y-b)^2 = 1$. Vi har alltså enhetscirkeln och en cirkel med radie 1 och mittpunkt i (0,b). Cirklarna skär varandra om $-2 \leq b \leq 2$. Om man ber Mathematica att göra denna elimination svarar Mathematica att $b = 0 || b^2 \leq 4$ vilket är helt korrekt. Dock har Mathematica separerat $b = 0$, vilket inte behövs då det ingår i $b^2 \leq 4$. Den kanske gör det för att i fallet $b = 0$ så får cirklarna identiska ekvationer.

**Exempel 4.2.** I nästa exempel testar vi två variabler. Först testar med cirklarna $x^2 + y^2 = 1$ och $(x-a)^2 + (y-b)^2 = 1$. I detta fall har vi enhetscirkeln samt en cirkel med radie 1 och mittpunkt i $(a,b)$ När man testar att skriva in följande i Mathematica:




Figur 6:
Här ser vi exempel på när cirklarna skär varandra, i detta fall är $b = 2$

Resolve[Exists[$x, y, x^2 + y^2 == 1$ && $(x - a)^2 + (y - b)^2 == 1$], Reals]

Mathematica svarade på följande sätt:
$a^2 - 2b + b^2 == 0||$
$a^2 + 2b + b^2 == 0||$ $(a^2 + b^2 > 0\&\& - 4a^2 + a^4 + a^2b^2 \leq 0\&\&a^4 + 4b^2 - b^4 \geq 0)||$
$(a^2 + b^2 > 0\&\& - 4a^2 + a^4 + a^2b^2 \leq 0\&\&a^4 - 4b^2 + 2a^2b^2 + b^4 == 0)$

Detta är inte lika tydligt som föregående exempel vad Mathematica svarar, några förenklingar krävs för att tolka svaret. Det första Mathematica svarar är $a^2 - 2b + b^2 = 0$ . Detta kan skrivas som $a^2 + (b - 1)^2 = 1$ vilket beskriver en ny cirkel med radie 1 och medelpunkt i $(0, 1)$ och denna cirkel skär enhetscirkeln. Andra raden i svaret som Mathematica ger, $a^2 + 2b + b^2 == 0$ är det samma som första raden bara att det beskriver en cirkel med radie 1 med mittpunkt i $(0, -1)$ istället för $(0, 1)$. Därefter kommer Mathematica med villkoret att $a^2 + b^2 > 0$ . Detta är onödigt för Mathematica att ha med då $a$ och $b$ är reella, men av någon anledning har Mathematica med det uttrycket. Nästa uttryck $-4a^2+a^4+a^2b^2 \leq 0$

ser komplicerat ut. Men om man dividerar bort $a^2$ , blir det lite mera förståligt ( man måste dock hantera $a = 0$ separat). Vi får då: $-4 + a^2 + b^2 \leq 0$, alltså att $a^2 + b^2 \leq 4$. Detta säger oss att avståndet mellan origo och $(a, b)$ är mindre än 2. Om vi har en cirkel som uppfyller detta kriterium så skär den enhetscirkeln. Sista villkoret: $a^4 + 4b^2 - b^4 \leq 0$ kan vi skriva om till $(b^2 - 2)^2$

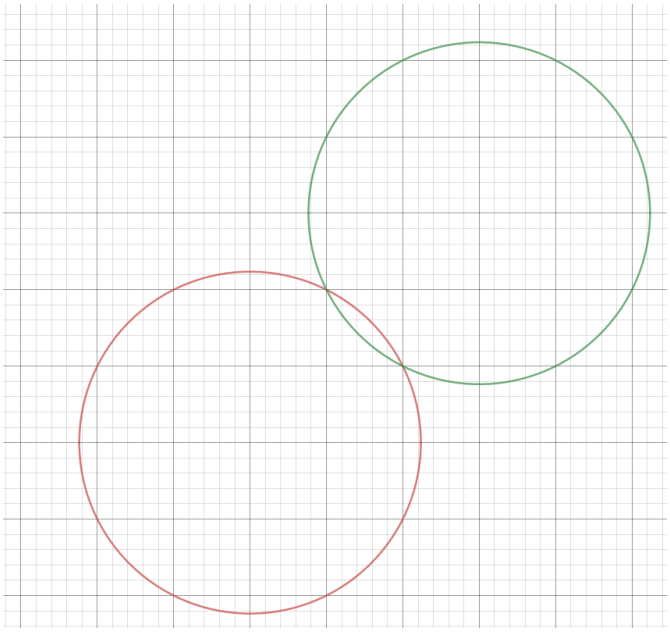
**Exempel 4.3.** Nu har vi testat två variabler. Vad händer om vi låter alla variabler vara fria i dessa två cirklar. Vi börjar med att definiera två cirklar, en med radie $r$ och origo i punkten $(a, b)$. Den kan uttryckas med standard ekvationen för en cirkel:

$$(x - a)^2 + (y - b)^2 = r^2 \tag{23}$$

Den andra cirkeln har radie $s$ och origo i punkten $(c, d)$. På samma sätt kan vi uttrycka en ekvation för denna cirkeln.

$$(x - c)^2 + (y - d)^2 = s^2 \tag{24}$$

Om vi ritar ut dessa cirklar i ett koordinat system kan det se ut som i figur 7.



Figur 7:
Bilden visar de två cirklarna som beskrivs av ekv(23) och ekv(24)

Om vi nu skriver om ekv(23) och ekv(24) till första ordningens formel med fria variabler kan ekv(23) skrivas på formen: $C(a, b, r, x, y)$ och ekv(24) kan på formen: $C(c, d, s, x, y)$ där $a, b, c, d, r, s, x, y$ är fria variabler. Nu vill vi studera

problemet när de båda cirklarna skär varandra i en punkt $(x, y)$. Detta kan vi skriva som ett uttryck på formen:

$$\exists x \exists y C(a, b, r, x, y) \wedge C(c, d, s, x, y) \tag{25}$$

Vi kan skriva ihop ekv (23) och ekv (24) såhär:

$$\exists x \exists y (x - a)^2 + (y - b)^2 = r^2 \wedge (x - c)^2 + (y - d)^2 = s^2 \tag{26}$$

Detta kan omvandlas till en kvantifikatorfri form enligt lemma 1. Då går det att skriva på denna form:

$$D(a, b, c, d, r, s) \tag{27}$$

D är nu ett system av likheter och olikheter av polynom med variablerna $a, b, c, d, r, s$. Nu ska jag undersöka om Mathematica kan lösa detta problem, detta returnerar Mathematica (fig. 8).

```
In[27]:= Resolve[
  Exists[{x, y}, (x - a)^2 + (y - b)^2 == r^2 && (x - c)^2 + (y - d)^2 == s^2], Reals]

Out[27]= (a^2 + b^2 - 2 a c + c^2 - 2 b d + d^2 == 0 &&
    a^2 b + b^3 - 2 a b c + b c^2 + a^2 d - b^2 d - 2 a c d + c^2 d - b d^2 + d^3 - b r^2 + d r^2 + b s^2 - d s^2 != 0 &&
    a^8 - 4 a^6 b^2 - 2 a^4 b^4 + 12 a^2 b^6 + 9 b^8 - 8 a^7 c + 24 a^5 b^2 c + 8 a^3 b^4 c - 24 a b^6 c + 28 a^6 c^2 -
      60 a^4 b^2 c^2 - 12 a^2 b^4 c^2 + 12 b^6 c^2 - 56 a^5 c^3 + 80 a^3 b^2 c^3 + 8 a b^4 c^3 + 70 a^4 c^4 -
      60 a^2 b^2 c^4 - 2 b^4 c^4 - 56 a^3 c^5 + 24 a b^2 c^5 + 28 a^2 c^6 - 4 b^2 c^6 - 8 a c^7 + c^8 - 8 a^6 b d +
      24 a^4 b^3 d + 8 a^2 b^5 d - 24 b^7 d + 48 a^5 b c d - 96 a^3 b^3 c d - 16 a b^5 c d - 120 a^4 b c^2 d +
      144 a^2 b^3 c^2 d + 8 b^5 c^2 d + 160 a^3 b c^3 d - 96 a b^3 c^3 d - 120 a^2 b c^4 d + 24 b^3 c^4 d +
      48 a b c^5 d - 8 b c^6 d + 4 a^6 d^2 + 12 a^4 b^2 d^2 - 52 a^2 b^4 d^2 + 4 b^6 d^2 - 24 a^5 c d^2 -
```

Figur 8:

Nu har vi fått fram ett kvantifikatorfritt uttryck(fig. 8), det är extremt mycket längre än ovan (bifogar resterande se bilaga 1). Det är svårtolkat vad detta betyder. Ett sätt att få fram ett enklare uttryck kanske skulle kunna vara att skapa en linje som skär punkterna där cirklarna skär varandra.

## 4.2 Skärning mellan två linjer

Nu går vi vidare och studerar skärningen mellan två linjer. Vi börjar med att skapa två linjer $y = ax + b, y = cx + d$.

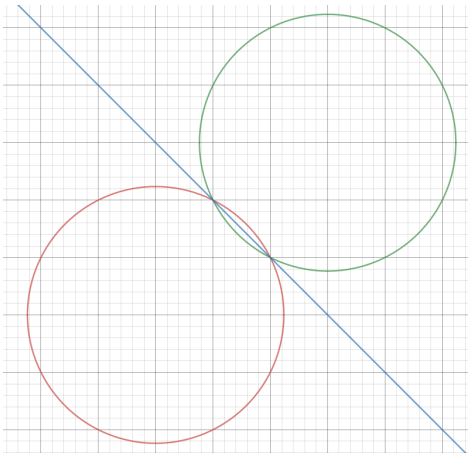
Nu kan vi få fram ett uttryck för att dessa linjär skär varandra. Våra fria variabler i detta fall är $a, b, c, d$

$$\exists x \exists y (y = ax + b) \wedge (y = cx + d) \tag{28}$$

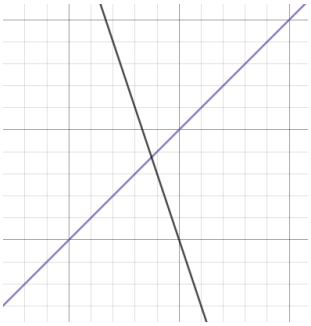
Löser vi detta i Mathematica med hjälp av Resolve får vi ut ett uttryck på kvantifikatorfri form som ser ut såhär:

$$a = 0 \wedge c \neq 0 \vee a \neq 0 \wedge a - c \neq 0 \vee b - d = 0 \tag{29}$$

Detta klarade Mathematica utan problem.



Figur 9:



Figur 10:

# 5 Avslutning

Mathematica klarar av att eliminera de problem som jag testade den på. Dock är inte alltid svaren lätta att tolka och blir mera svåra att tolkade ju fler variabler man har. Men även med få variabler returnerar den på konstiga former istället för logiskt. Den har även en tendens att ha med överflödiga data som ett krav att talen i kvadrat ska vara större än noll när vi ha angett att de är reella, samt har med vissa onödiga villkor som tidigare villkor redan täcker. Detta är troligvis en av anledningarna till att det eskalerade med fler variabler. Min slutsats är att Mathematica klarar av att eliminationen men formen den ger uttrycken är svårtolkade.

# Referenser


[1] Erwin Engeler *Foundations of Mathematics Questions of Analysis, Geometry and Algorithmics* Springer-Verlag(1993)

[2] B.F. Caviness and J.R. Johnson (eds.) *Quantifier Elimination and Cylindrical Algebraic Decomposition* SpringerWienNewYork (1998)

[3] Alfred Tarski *What is elementary geometry?* University of California (1959)

[4] Trygve Nagell *lärobok i algebra* Uppsala (1949)

# 6 Bilagor

### Bilaga 1:

```
In[9]:= Resolve[Exists[{x, y}, (x)^2 + (y)^2 == r^2 && (x - c)^2 + (y - d)^2 == s^2], Reals]
Out[9]= (d ≠ 0 && c^2 + d^2 == 0 && c^2 + d^2 + r^2 - s^2 ≠ 0 &&
    c^8 + 4 c^6 d^2 + 6 c^4 d^4 + 4 c^2 d^6 + d^8 - 4 c^6 r^2 - 20 c^4 d^2 r^2 - 28 c^2 d^4 r^2 - 12 d^6 r^2 + 6 c^4 r^4 -
      36 c^2 d^2 r^4 - 26 d^4 r^4 - 4 c^2 r^6 - 12 d^2 r^6 + r^8 - 4 c^6 s^2 - 12 c^4 d^2 s^2 - 12 c^2 d^4 s^2 - 4 d^6 s^2 +
      4 c^4 r^2 s^2 + 24 c^2 d^2 r^2 s^2 + 20 d^4 r^2 s^2 + 4 c^2 r^4 s^2 + 20 d^2 r^4 s^2 - 4 r^6 s^2 + 6 c^4 s^4 +
      12 c^2 d^2 s^4 + 6 d^4 s^4 + 4 c^2 r^2 s^4 - 4 d^2 r^2 s^4 + 6 r^4 s^4 - 4 c^2 s^6 - 4 d^2 s^6 - 4 r^2 s^6 + s^8 ≤ 0) ||
  (c^2 + d^2 > 0 && c^2 d + d^3 + d r^2 - d s^2 ≥ 0 &&
    c^4 + 2 c^2 d^2 + d^4 + 2 c^2 r^2 - 2 d^2 r^2 + r^4 - 2 c^2 s^2 - 2 d^2 s^2 - 2 r^2 s^2 + s^4 == 0 &&
    c^6 + 2 c^4 d^2 + c^2 d^4 - 2 c^4 r^2 - 2 c^2 d^2 r^2 + c^2 r^4 - 2 c^4 s^2 - 2 c^2 d^2 s^2 - 2 c^2 r^2 s^2 + c^2 s^4 ≤ 0) ||
  (c^2 + d^2 > 0 && c^6 + 2 c^4 d^2 + c^2 d^4 - 2 c^4 r^2 - 2 c^2 d^2 r^2 + c^2 r^4 - 2 c^4 s^2 - 2 c^2 d^2 s^2 -
      2 c^2 r^2 s^2 + c^2 s^4 ≤ 0 && c^6 + c^4 d^2 - c^2 d^4 - d^6 + 2 c^4 r^2 + 4 c^2 d^2 r^2 + 2 d^4 r^2 +
      c^2 r^4 - d^2 r^4 - 2 c^4 s^2 + 2 d^4 s^2 - 2 c^2 r^2 s^2 + 2 d^2 r^2 s^2 + c^2 s^4 - d^2 s^4 ≥ 0) ||
  (c^2 + d^2 > 0 && c^2 d + d^3 + d r^2 - d s^2 ≤ 0 &&
    c^4 + 2 c^2 d^2 + d^4 + 2 c^2 r^2 - 2 d^2 r^2 + r^4 - 2 c^2 s^2 - 2 d^2 s^2 - 2 r^2 s^2 + s^4 == 0 &&
    c^6 + 2 c^4 d^2 + c^2 d^4 - 2 c^4 r^2 - 2 c^2 d^2 r^2 + c^2 r^4 - 2 c^4 s^2 - 2 c^2 d^2 s^2 - 2 c^2 r^2 s^2 + c^2 s^4 ≤ 0) ||
  (c^2 + d^2 - 2 d r + r^2 - s^2 == 0 && c^6 d + 3 c^4 d^3 + 3 c^2 d^5 + d^7 + 3 c^4 d r^2 + 10 c^2 d^3 r^2 +
      7 d^5 r^2 + 3 c^2 d r^4 + 7 d^3 r^4 + d r^6 - 3 c^4 d s^2 - 6 c^2 d^3 s^2 - 3 d^5 s^2 - 6 c^2 d r^2 s^2 -
      10 d^3 r^2 s^2 - 3 d r^4 s^2 + 3 c^2 d s^4 + 3 d^3 s^4 + 3 d r^2 s^4 - d s^6 ≥ 0) ||
  (c^2 + d^2 - 2 d r + r^2 - s^2 == 0 && c^6 d + 3 c^4 d^3 + 3 c^2 d^5 + d^7 + 3 c^4 d r^2 + 10 c^2 d^3 r^2 +
      7 d^5 r^2 + 3 c^2 d r^4 + 7 d^3 r^4 + d r^6 - 3 c^4 d s^2 - 6 c^2 d^3 s^2 - 3 d^5 s^2 - 6 c^2 d r^2 s^2 -
      10 d^3 r^2 s^2 - 3 d r^4 s^2 + 3 c^2 d s^4 + 3 d^3 s^4 + 3 d r^2 s^4 - d s^6 ≤ 0) ||
  (c^2 + d^2 + 2 d r + r^2 - s^2 == 0 && c^6 d + 3 c^4 d^3 + 3 c^2 d^5 + d^7 + 3 c^4 d r^2 + 10 c^2 d^3 r^2 +
      7 d^5 r^2 + 3 c^2 d r^4 + 7 d^3 r^4 + d r^6 - 3 c^4 d s^2 - 6 c^2 d^3 s^2 - 3 d^5 s^2 - 6 c^2 d r^2 s^2 -
      10 d^3 r^2 s^2 - 3 d r^4 s^2 + 3 c^2 d s^4 + 3 d^3 s^4 + 3 d r^2 s^4 - d s^6 ≥ 0) ||
  (c^2 + d^2 + 2 d r + r^2 - s^2 == 0 && c^6 d + 3 c^4 d^3 + 3 c^2 d^5 + d^7 + 3 c^4 d r^2 + 10 c^2 d^3 r^2 +
      7 d^5 r^2 + 3 c^2 d r^4 + 7 d^3 r^4 + d r^6 - 3 c^4 d s^2 - 6 c^2 d^3 s^2 - 3 d^5 s^2 -
      6 c^2 d r^2 s^2 - 10 d^3 r^2 s^2 - 3 d r^4 s^2 + 3 c^2 d s^4 + 3 d^3 s^4 + 3 d r^2 s^4 - d s^6 ≤ 0)
```

Figur 11: