EN INTRODUKSJON TIL HYPERBOLSK GEOMETRI MAT4510/3510 HØSTEN 2004

BJØRN JAHREN

(Versjon: 13. oktober 2004)

1. Innledning. Stereografisk projeksjon

Med *hyperbolsk geometri* skal vi mene en insidensgeometri med mellomhet og kongruens, der Hilberts aksiomsystem gjelder — bortsett fra at parallellaksiomet er erstattet med

H: Gitt en linje l og et punkt $P \notin l$, så finnes minst to linjer m gjennom P som ikke skjærer l.

Da vi gikk gjennom Hilberts aksiomsystem, observerte vi at en åpen, konveks delmengde K av det Euklidske plan er en kandidat til en slik geometri, hvis vi lar 'linjer' bety snitt mellom K og Euklidske linjer, og mellomhet er definert som i $\mathbb{R}^2$. Da vil insidens- og mellomhetsaksiomene være oppfylt, og likeledes Dedekinds aksiom. Derimot finnes det åpenbart *uendelig mange* linjer parallelle med en gitt linje og gjennom et gitt punkt utenfor linjen, så aksiom H vil gjelde i stedet for aksiom P.

Det som mangler for at vi skal ha en *hyperbolsk geometri* er derfor et *kongruensbegrep* som tilfredsstiller Hilberts aksiomer C1–C6. Prøver vi oss med det vanlige (Euklidske) kongruensbegrepet, er riktignok C2–C6 fremdeles oppfylt, men det fundamentale C1 bryter sammen. Vi skal nå se på en spesiell slik K, nemlig det indre av en sirkelskive med radius 1, og vise at vi i dette tilfellet kan utvikle et ny kongruensrelasjon som er forenlig med resten av den strukturen vi har, og som tilfredsstiller alle aksiomene. Dette er *Beltrami–Kleins modell* $\mathbb{K}$ for det hyperbolske plan og er kanksje den modellen som er enklest å visualisere og enklest med hensyn på insidens- og mellomhetsrelasjonene. Det er likevel ikke denne vi skal studere mest. I dette innledningskapitlet skal vi se at $\mathbb{K}$ er ekvivalent (mhp. den strukturen vi har definert så langt) med to andre modeller: *Poincarés øvre halvplan* $\mathbb{H}$ og *Poincarés disk* $\mathbb{D}$. Disse kan virke mindre naturlige ved første øyekast, men de har mange egenskaper som gjør dem mye enklere å studere enn $\mathbb{K}$.

Litt notasjon: I de forskjellige modellene $(\mathbb{K}, \mathbb{B}, \mathbb{D}, \mathbb{H})$ vil det være forskjellig typer kurver som utgjør 'linjene' i geometrien. Vi skal da bruke betegnelsen $\mathbb{K}$–linjer, $\mathbb{B}$–linjer etc., eller evt. *hyperbolske* linjer hvis modellen er underforstått, eller hvis det ikke spiller noen rolle hvilken modell vi bruker. For eksempel er $\mathbb{K}$–linjene da de åpne kordene i enhetssirkelen i planet.

Som mengde er altså $\mathbb{K}$ det indre av enhetsdisken i $\mathbb{R}^2$: $\mathbb{K} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$. Betrakt $\mathbb{R}^2$ som underrommet av $\mathbb{R}^3$ der siste koordinat er 0, og la $\mathbb{B}$ være den nedre halvsfæren $\{(x, y, z) \in R^3 \mid x^2 + y^2 + z^2 = 1, z < 0\}$. Vertikal projeksjon gir da en homeomorfi $\mathbb{K} \approx \mathbb{B}$ som avbilder kordene i $\mathbb{K}$ på

halvsirkler i $\mathbb{B}$ som ender ortogonalt på randkurven $\{(x,y) \in \mathbb{R}^2 \,|\, x^2+y^2 = 1\}$. (Det er kanskje lettest å innse dette hvis vi ser på bildet av en korde som snittkurven mellom $\mathbb{B}$ og planet som inneholder korden og er parallelt med z-aksen. Da får vi halvparten av snittkurven med hele enhetssfæren $\mathbb{S}^2$, som åpenbart er en sirkel.) Med slike halvsirkler som $\mathbb{B}$–linjer blir derfor også $\mathbb{B}$ en modell med de samme egenskapene som $\mathbb{K}$.

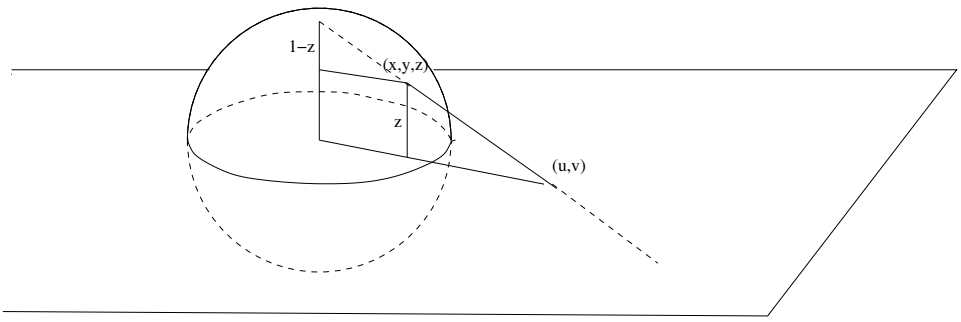
Vi skal nå bruke *stereografisk projeksjon* til å avbilde $\mathbb{B}$ tilbake til $\mathbb{R}^2$ igjen. Stereografisk projeksjon er en en–entydig korrespondanse $\mathbb{S}^2 - (0,0,1) \approx \mathbb{R}^2$ definert som følger: Hvis P er et punkt i $\mathbb{S}^2 - (0,0,1)$, finnes en entydig bestemt rett linje gjennom P og (0,0,1), og denne linjen treffer $\mathbb{R}^2$ i et entydig punkt. Dette definerer en funksjon $\Phi : \mathbb{S}^2 - (0,0,1) \to \mathbb{R}^2$, som åpenbart er både surjektiv og injektiv. En enkelt argument med likedannede trekanter (se fig.1) viser at Φ er gitt ved formelen

$$\Phi(x,y,z) = (\frac{x}{1-z}, \frac{y}{1-z}) \,, \tag{1.1}$$

og den omvendte funksjonen er gitt ved

$$\Phi^{-1}(u,v) = (\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{u^2+v^2-1}{u^2+v^2+1}) \,. \tag{1.2}$$

Disse funksjonene er begge kontinuerlige, og er derfor omvendte *homeomorfier*.



FIGUR 1. Stereografisk projeksjon

De to følgende lemmaene handler om viktige egenskaper til stereografisk projeksjon.

Lemma 1.1. *La C være en sirkel på $\mathbb{S}^2$.*

(i) *Hvis $(0,0,1) \notin C$, så er $\Phi(C)$ en sirkel i $\mathbb{R}^2$.*
(ii) *Hvis $(0,0,1) \in C$, så er $\Phi(C - (0,0,1))$ en rett linje i $\mathbb{R}^2$*

Bevis. Observer først at hvis $(x,y,z) \in \mathbb{S}^2 - (0,0,1)$ og $(u,v) = \Phi(x,y,z)$, så er, etter formel (1.1):

$$u^2 + v^2 = \frac{x^2+y^2}{(1-z)^2} = \frac{1-z^2}{(1-z)^2} = \frac{1+z}{1-z} \,.$$

Sirkelen C er skjæringskurven mellom $\mathbb{S}^2$ og et plan som vi kan anta har ligningen $ax + by + cz = d$. Hvis $(x,y,z) \neq (0,0,1)$ er et punkt på C, må vi

da ha

$$\begin{aligned} au + bv = \frac{ax + by}{1 - z} &= \frac{d - cz}{1 - z} = \\ &= \frac{d + c}{2} + \frac{d - c}{2}\left(\frac{1 + z}{1 - z}\right) \\ &= \frac{d + c}{2} + \frac{d - c}{2}(u^2 + v^2) \end{aligned}$$

I tilfelle (i) er $d \neq c$, og dette er ligningen for en sirkel. I tilfelle (ii) er $d = c$, og vi har ligningen for en rett linje. □

Lemma 1.2. *Φ bevarer vinkler — dvs. hvis de to storsirklene $\mathcal{C}$ og $\mathcal{C}'$ skjærer hverandre under vinkelen θ i (x_0, y_0, z_0) så skjærer $\Phi(\mathcal{C})$ og $\Phi(\mathcal{C}')$ hverandre under samme vinkel.*

Bevis. La de to sirklene være snittkurvene med planene $ax + by + cz = 0$ og $a'x + b'y + c'z = 0$, der vi kan anta at $a^2 + b^2 + c^2 = a'^2 + b'^2 + c'^2 = 1$. Vektorene $\pm(a, b, c)$ og $\pm(a', b', c')$ er enhetsnormaler til disse planene, og valg av tegn svarer til valg av omløpsretning i sirklene ('høyrehåndsregelen'). Da er θ lik vinkelen mellom normalene, så vi har $\cos(\theta) = aa' + bb' + cc'$. På grunn av rotasjonssymmetri om z-aksen er det tilstrekkelig å betrakte et punkt der $y_0 = 0$. Videre kan vi da sammenligne hver av de to storsirklene med den gitt ved $y = 0$, så vi kan anta at $(a', b', c') = (0, 1, 0)$, slik at $\cos(\theta) = b$ og $\Phi(\mathcal{C}')$ er linjen $v = 0$ med normal $(0, 1, 0)$. $\Phi(\mathcal{C})$ er gitt ved ligningen (fra lemma 1.1) $\frac{c}{2}(u^2 + v^2) + au + bv = \frac{c}{2}$, og i punktet (u_0, v_0) har denne normalvektoren $(cu_0 + a, cv_0 + b) = (cu_0 + a, b)$. Vinkelen mellom $\Phi(\mathcal{C})$ og $\Phi(\mathcal{C}')$ er derfor gitt ved

$$\cos(\theta') = \frac{b \cdot 1}{\sqrt{(cu_0 + a)^2 + b^2 \cdot 1}} = \frac{b}{\sqrt{c^2u_0^2 + 2cu_0 + a^2 + b^2}} = b\,,$$

som vi får ved å sette inn $(u_0, 0)$ i ligningen for $\mathcal{C}$ og bruke at $a^2 + b^2 + c^2 = 1$. □

(Du finner instruktive geometriske bevis for begge disse lemmaene i Hilbert-Cohn Vossens *"Geometry and the imagination"*, s. 248ff.)

Det er naturligvis mange måter å identifisere $\mathbb{B}$ med en åpen halvkule i $\mathbb{S}^2$ på, og vi får et en–entydig bilde i planet (og derfor en ny modell) bare vi unngår punktet (0,0,1). Lemmaene ovenfor gir at alle slike modeller vil være områder begrenset av enten sirkler eller rette linjer i $\mathbb{R}^2$, og 'linjene' vil svare til sirkelbuer eller rette linjer som står ortogonalt på begrensningskurven. Vi skal bruke to slike modeller, som begge har navn etter den store franske matematikeren Henri Poincaré:

1. *Poincarés diskmodell* $\mathbb{D}$ får vi ved å la $\mathbb{B}$ være den nedre halvkulen, som før. Da blir bildet det indre av enhetssirkelen om origo igjen, men linjene blir nå enten diametere eller sirkelbuer som står ortogonalt på randkurven.

2. Det *øvre halvplanet* $\mathbb{H} = \{(x, y) \mid y > 0\}$ er bildet av halvkulen $\{(x, y, z) \in \mathbb{S}^2 | y > 0\}$. $\mathbb{H}$–linjene blir enten halvsirkler med diameter langs x-aksen eller rette linjer parallelle med y-aksen.

I begge modellene skal vi heretter identifisere $\mathbb{R}^2$ med det komplekse plan $\mathbb{C}$ og gjøre utstrakt bruk av den ekstra strukturen vi har der (kompleks

multiplikasjon, kompleks funksjonsteori etc.). Som *mengder* er altså $\mathbb{D} = \{z \in \mathbb{C} \,|\, |z| < 1\}$ og $\mathbb{H} = \{z \in \mathbb{C} \,|\, \operatorname{Im} z > 0\}$.

$\mathbb{D}$ er mest symmetrisk, og derfor den som ofte egner seg best til geometriske argumenter. Men, som vi skal se, så er $\mathbb{H}$ den som egner seg best til å analysere og beskrive *kongruensbegrepet*. Derfor er det den vi skal begynne med å studere.

Oppgaver.

1.1. Utled formlene for Φ og dens inverse.

1.2. Vi kan også definere stereografisk projeksjon ved å projisere fra $(0, 0, -1)$ i stedet for $(0, 0, 1)$. La Φ_- være avbildningen vi da får.

Bestem $\Phi_- \circ \Phi^{-1}$. (Vi identifiserer $\mathbb{R}^2$ med $\mathbb{C}$.)

1.3. Vi kan få en homeomorfi mellom de to modellene $\mathbb{H}$ og $\mathbb{D}$ som $\Phi \circ F \circ \Phi^{-1}$, der F er en identifikasjon mellom de to halvkulene vi brukte i definisjonene av disse modellene. Finn en formel for en slik homeomorfi. (Velg F enklest mulig.)

2. Kongruens i $\mathbb{H}$. Möbiustransformasjoner.

Vi ønsker å definere kongruens som ekvivalens under en gruppe av "rigide" bevegelser. Det vil si at vi vil prøve å finne en passende gruppe G av homeomorfier av det øvre halvplan slik at vi kan definere kongruensrelasjonen $\cong$ ved

Segmenter: $AB \cong A'B' \iff$ det finnes en $g \in G$ slik at $g(AB) = A'B'$
Vinkler: $\angle BAC \cong \angle B'A'C' \iff$ det finnes en $g \in G$ slik at $g(\overrightarrow{AB}) = \overrightarrow{A'B'}$ og $g(\overrightarrow{AC}) = \overrightarrow{A'C'}$

Undersøker vi Hilberts kongruensaksiomer, ser vi at for at dette skal kunne gi en fornuftig definisjon, så må G bl. a. oppfylle:

> *G må avbilde $\mathbb{H}$–linjer på $\mathbb{H}$–linjer, og hvis A og A' er punkter på henholdsvis $\mathbb{H}$-linjene l og l', så finnes $g \in G$ slik at $g(l) = l'$ og $g(A) = A'$.*

Vi skal finne en slik gruppe, og den vil bestå av såkalte *Möbiustransformasjoner*, restriktert til det øvre halvplan.

Fra kompleks funksjonsteori vet vi at meromorfe funksjoner kan betraktes som funksjoner $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$, der $\overline{\mathbb{C}}$ er det *utvidede komplekse plan* eller *Riemannsfæren* $\mathbb{C} \cup \{\infty\}$, med topologi slik at stereografisk projeksjon utvides til en homeomorfi $\overline{\Phi} : \mathbb{S}^2 \approx \overline{\mathbb{C}}$. Lemma 1 sier at $\overline{\Phi}$ avbilder sirkler på kurver i $\overline{\mathbb{C}}$ som enten er sirkler i $\mathbb{C}$ eller av formen $l \cup \{\infty\}$, der l er en (reell) linje i $\mathbb{C}$. Vi skal kalle alle disse kurvene *sirkler i* $\overline{\mathbb{C}}$.

Hvis to slike sirkler skjærer hverandre i to punkter i $\mathbb{C}$, så er skjæringsvinklene i de to punktene den samme. Hvis de har ett skjæringspunkt i $\mathbb{C}$ og et i ∞ (dvs. vi har to rette, reelle linjer i $\mathbb{C}$ som skjærer hverandre), så er det derfor naturlig å si at skjæringsvinkelen i ∞ er den samme som i det endelige skjæringspunktet. Hvis de bare skjærer hverandre i ∞, så har vi to parallelle rette linjer i $\mathbb{C}$, og vi sier at de *tangerer* hverandre i ∞, og skjæringsvinkelen er 0.

For at en meromorf $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ skal være en homeomorfi, må den ha nøyaktig ett nullpunkt og én pol, så den må nødvendigvis ha formen

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

der $a, b, c, d \in \mathbb{C}$. Løser vi ligningen $w = f(z)$ med hensyn på z, så får vi

$$z = g(w) = \frac{dw-b}{-cw+a}\,,$$

og vi ser ved utregning at

$$f(g(w)) = \frac{(ad-bc)w}{ad-bc} \text{ og } g(f(z)) = \frac{(ad-bc)z}{ad-bc}\,.$$

Derfor er f invertibel med g som invers hvis $ad-bc \neq 0$. Hvis $ad-bc = 0$ har disse uttrykkene ingen mening, men vi sjekker lett at da er $f(z)$ er konstant. Vi har altså:

Funksjonen $f(z) = \dfrac{az+b}{cz+d}$ *definerer en homeomorfi* $\overline{\mathbb{C}} \approx \overline{\mathbb{C}}$ *hvis og bare hvis* $ad-bc \neq 0$.

Slike funksjoner skal vi kalle *lineære brøktransformasjoner* — forkortet LBT.

Vi begynner med å gi noen viktige egenskaper til LBT-er.

Lemma 2.1. (i) *En LBT avbilder sirkler i* $\overline{\mathbb{C}}$ *på sirkler i* $\overline{\mathbb{C}}$.
(ii) *En LBT bevarer vinkler mellom sirkler.*

Bevis. (i) Merk at vi kan skrive ligninger for både sirkler og rette linjer i $\mathbb{C}$ på formen $\lambda(x^2+y^2) + \alpha x + \beta y + \gamma = 0$, der $\lambda, \alpha, \beta, \gamma$ er relle tall og $z = x+iy$. $\lambda = 0$ gir en linje, og $\lambda \neq 0$ en sirkel. Bruker vi nå at $x^2+y^2 = z\overline{z}$, $x = (z+\bar{z})/2$ og $y = (z-\bar{z})/2i = (\bar{z}-z)i/2$, så kan vi skrive ligningen som

$$\lambda z\bar{z} + \mu z + \bar{\mu}\bar{z} + \gamma = 0\,,$$

der $\mu = (\alpha - i\beta)/2$. Vi trenger altså å vise at hvis z tilfredsstiller en slik ligning og $w = f(z)$ for en LBT f, så tilfredsstiller w en tilsvarende ligning. Dette kan vi sjekke ved å skrive $z = f^{-1}(w) = \dfrac{aw+b}{cw+d}$ og sette inn:

$$\lambda\frac{aw+b}{cw+d}\cdot\frac{\overline{aw+b}}{\overline{cw+d}} + \mu\frac{aw+b}{cw+d} + \overline{\mu}\frac{\overline{aw+b}}{\overline{cw+d}} + \gamma = 0\,.$$

Multipliserer vi hele denne ligningen med $(cw+d)(\overline{cw+d})$ og ordner, så får vi etter litt regning nettopp et slikt uttrykk som vi ønsker.

(ii) Vi begrenser oss her til å betrakte skjæringspunkter utenom ∞ og eventuelle poler. Da er $f(z) = \dfrac{az+b}{cz+d}$ analytisk med derivert $f'(z) = \dfrac{ad-bc}{(cz+d)^2}$, som er forskjellig fra 0. Hvis $t \mapsto \omega(t)$ er en differensiabel kurve i $\mathbb{C}$, så angir $\omega'(t_0)$ retningen til kurven i punktet $\omega(t_0)$, og på samme måte angir $\frac{d}{dt}f(\omega(t))|_{t_0} = f'(\omega(t_0))\cdot\omega'(t_0)$ retningen til kurven $f(\omega(t))$ i $f(\omega(t_0))$. Men multiplikasjon med det komplekse tallet $f'(\omega(t_0))$ er det samme som dreining med en fast vinkel $\arg f'(\omega(t_0))$ og multiplikajon med det positive tallet $|f'(\omega(t_0))|$, og begge disse operasjonen bevarer vinkler. □

(En avbildning som bevarer vinkler kalles *konform,* og beviset gir at en kompleks analytisk avbildning er konform der den deriverte er forskjellig fra 0. Lemma 1.2 sier at også sterografisk projeksjon er konform.)

Ordet 'lineær' i LBT har sammenheng med til følgende bemerkelsesverdige observasjon:

La $f(z) = \dfrac{az+b}{cz+d}$ og $g(z) = \dfrac{a'z+b'}{c'z+d'}$. Da blir (regn ut)

$$(f \circ g)(z) = f(g(z)) = \frac{a\,g(z)+b}{c\,g(z)+d} = \frac{(aa'+bc')z+(ab'+bd')}{(ca'+dc')z+(cb'+dd')}$$

Denne formelen sier oss to ting. For det første betyr den at sammensetningen av to LBT-er er en ny LBT. Vi har vist ovenfor at den inverse av en LBT er en ny LBT, og vi har trivielt at identitesavbildningen også er. ($z = \frac{1z+0}{0z+1}$.) Derfor danner mengden av lineære brøktransformasjoner en *gruppe* under sammensetning. Denne gruppen skal vi kalle $M\ddot{o}b^+(\mathbb{C})$.

Det andre er at det er mulig å regne med lineære brøktransformasjoner som med *matriser*: det er uten videre klart at matrisen $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ bestemmer f fullstendig, og betingelsen $ad-bc \neq 0$ betyr ikke noe annet enn at denne matrisen er invertibel. På samme måte er g bestemt av $\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$, og utregningen vi nettopp gjorde viser at $f \circ g$ er bestemt av matrisen $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$.

Mengden av invertible 2×2–matriser over $\mathbb{C}$ danner en gruppe — *den generelle lineære gruppen* $GL_2(\mathbb{C})$. Vi har altså en surjektiv homomorfi av grupper fra $GL_2(\mathbb{C})$ på $M\ddot{o}b^+(\mathbb{C})$. Denne homomorfien er ikke injektiv, for hvis $k \neq 0$ vil $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ og $\begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix} = k \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ bestemme samme avbildning. Imidlertid er dette den eneste flertydigheten vi har, så vi får en *isomorfi* mellom gruppen $M\ddot{o}b^+(\mathbb{C})$ av lineære brøktransformasjoner og kvotientgruppen $PGL_2(\mathbb{C}) = GL_2(\mathbb{C})/D$, der $D = \{uI | u \in \mathbb{C}-\{0\}$ og I er identitesmatrisen$\}$.

Det neste resultatet beskriver en viktig måte å bestemme lineære brøktransformasjoner på.

Lemma 2.2. *Gitt tre distinkte punkter z_1, z_2 og z_3 i $\overline{\mathbb{C}}$. Da finnes en entydig bestemt LBT f slik at $f(z_1) = 0$, $f(z_2) = 1$ og $f(z_3) = \infty$.*

Bevis. Eksistens: Anta først at ingen av z_i-ene er ∞. Da kan vi sette

$$f(z) = \frac{z-z_1}{z-z_3} \cdot \frac{z_2-z_3}{z_2-z_1}$$

I de andre tilfellene kan vi sette:

$$\text{hvis } z_1 = \infty\text{:} \qquad f(z) = \frac{z_2-z_3}{z-z_3}$$

$$\text{hvis } z_2 = \infty\text{:} \qquad f(z) = \frac{z-z_1}{z-z_3}$$

$$\text{hvis } z_3 = \infty\text{:} \qquad f(z) = \frac{z-z_1}{z_2-z_1}$$

Entydighet: Anta $g(z)$ har samme egenskap, og betrakt $h = g \circ f^{-1}$. Dette blir en ny LBT med $h(0) = 0$, $h(1) = 1$ og $h(\infty) = \infty$. Den siste betingelsen

gir at h må ha formen $h(z) = az + b$, og de to første gir da at $a = 1$ og $b = 0$. Dvs. $(g \circ f^{-1})(z) = z$ for alle z, så $f = g$. □

Korollar 2.1. *Gitt to tripler (z_1, z_2, z_3) og (w_1, w_2, w_3) av distinkte punkter i $\overline{\mathbb{C}}$. Da finnes entydig LBT f slik at $f(z_i) = w_i,\ i = 1, 2, 3$. Hvis alle de seks punktene ligger i $\overline{\mathbb{R}}$, så kan f skrives med reelle koeffisienter.*

Eksistensen her uttrykker vi ved å si at $M\ddot{o}b^+(\mathbb{C})$ virker *transitivt* på mengden av slike tripler. *Entydigheten* betyr at hvis to lineære brøktransformasjoner stemmer overens på tre punkter, så er de like. Spesielt er en LBT som fikserer tre punkter lik identitetsavbildningen.

Bevis. Etter lemma 2.2 kan vi finne entydig bestemte h og g slik at $h(z_1) = 0,\ h(z_2) = 1$ og $h(z_3) = \infty$, og $g(w_1) = 0,\ g(w_2) = 1$ og $g(w_3) = \infty$. La $f = g^{-1}h$

Anta at f' også avbilder hver z_i på w_i. Da er gf og gf' begge som i lemma 2.2, og på grunn av entydigheten er $gf = gf'$. Derfor er også $f = f'$.

Den siste påstanden følger av formlene i beviset for 2.2. De viser at h og g har reelle koeffisienter, og derfor må også $f = g^{-1}h$ ha det. □

Vår neste observasjon er at $M\ddot{o}b^+(\mathbb{C})$ også virker transitivt på mengden av *sirkler* i $\overline{\mathbb{C}}$. Dette bygger på at tre distinkte punkter i $\overline{\mathbb{C}}$ bestemmer *entydig* en sirkel som går gjennom alle tre.

Korollar 2.2. *Gitt to sirkler $\mathcal{C}_1$ og $\mathcal{C}_2$ i $\overline{\mathbb{C}}$. Da finnes en lineær brøktransformasjon f slik at $f(\mathcal{C}_1) = \mathcal{C}_2$.*

Bevis. Velg tre distinkte punkter (z_1, z_2, z_3) på $\mathcal{C}_1$ og (w_1, w_2, w_3) på $\mathcal{C}_2$, og la f være som i korollaret ovenfor. Da er $f(\mathcal{C}_1)$ en sirkel i $\overline{\mathbb{C}}$ som inneholder w_1, w_2, og w_3 — dvs $\mathcal{C}_2$. □

Vi er interessert i å bestemme de lineære brøktransfomasjonene f som restrikterer til en homeomorfi av det øvre halvplan. En slik f er karakterisert ved at $f(\overline{\mathbb{R}}) = \overline{\mathbb{R}}$, og $\operatorname{Im} f(z) > 0$ hvis $\operatorname{Im} z > 0$. Her betyr $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \subset \overline{\mathbb{C}}$.

Proposisjon 2.1. *En lineær brøktransformasjon restrikterer til en homeomorfi av $\mathbb{H}$ hvis og bare hvis den kan skrives på formen $f(z) = \dfrac{az + b}{cz + d}$, der a, b, c, d er reelle og $ad - bc = 1$.*

En slik LBT bevarer $\mathbb{H}$–linjer.

Bevis. Korollar 2.1 sier at hvis $f(z) = \dfrac{az + b}{cz + d}$ restrikterer til en homeomorfi av $\mathbb{H}$, så kan a, b, c, d velges reelle, siden $f(\overline{\mathbb{R}}) = \overline{\mathbb{R}}$. Omvendt vil åpenbart $f(\overline{\mathbb{R}}) = \overline{\mathbb{R}}$ hvis a, b, c, d er reelle.

Litt regning gir

$$f(z) = \frac{(az + b)(c\bar{z} + d)}{(cz + d)(c\bar{z} + d)} = \frac{ac|z|^2 + (ad + bc)\operatorname{Re} z + bd}{|cz + d|^2} + \frac{(ad - bc)\operatorname{Im} z}{|cz + d|^2}\, i\,. \tag{2.1}$$

Det følger at f bevarer det øvre halvplanet hvis og bare hvis $ad - bc > 0$. Multipliserer vi derfor a, b, c og d med $\frac{1}{\sqrt{ad-bc}}$, får vi f på den formen vi ønsker.

Det siste utsagnet følger umiddelbart av at lineære brøktransformasjoner bevarer sirkler og vinkler i $\overline{\mathbb{C}}$. Enhver $\mathbb{H}$–linje bestemmer en sirkel i $\overline{\mathbb{C}}$ som står ortogonalt på $\mathbb{R}$, og siden f bevarer vinkler og $f(\overline{\mathbb{R}}) = \overline{\mathbb{R}}$, vil bildene av disse sirklene også stå ortogonalt på $\mathbb{R}$. □

De lineære brøktransformasjonene som restrikterer til homeomorfier av $\mathbb{H}$ utgjør åpenbart en undergruppe av $\mathit{Möb}^+(\mathbb{C})$, som vi kaller $\mathit{Möb}^+(\mathbb{H})$. Den kan også beskrives ved matriser, som følger:

La $SL_2(\mathbb{R})$ være *den spesielle lineære gruppen* — gruppen av 2×2–matriser med determinant 1. De eneste multipla av identitesmatrisen her er $\pm I$, så som ovenfor får vi en isomorfi mellom $\mathit{Möb}^+(\mathbb{H})$ og kvotientgruppen $PSL_2(\mathbb{R}) = SL_2(\mathbb{R})/(\pm I)$.

Det er ganske klart at vi ikke har fått med oss *alle* sirkelbevarende homeomorfier av $\overline{\mathbb{C}}$. *Kompleks konjugasjon* er jo også åpenbart sirkelbevarende, men er ikke engang kompleks analytisk. Vi definerer *gruppen av komplekse Möbiustransformasjoner,* $\mathit{Möb}(\mathbb{C})$, som gruppen av homeomorfier generert av de lineære brøktransformasjonene $\mathit{Möb}^+(\mathbb{C})$ og kompleks konjugasjon.

Proposisjon 2.2. (1) *Komplekse Möbiustransformasjoner kan skrives på en og bare en av formene*

$$f(z) = \frac{az+b}{cz+d} \quad \textit{eller} \quad f(z) = \frac{a\bar{z}+b}{c\bar{z}+d}, \quad \textit{der} \quad a, b, c, d \in \mathbb{C} \ \textit{og} \ ad - bc = 1\,.$$

(2) *De komplekse Möbiustransformasjonene som bevarer* $\mathbb{H}$*, kan skrives enten som*

$$\text{(i)} \quad f(z) = \frac{az+b}{cz+d}, \quad \textit{der} \quad a, b, c, d \in \mathbb{R} \quad \textit{og} \quad ad - bc = 1\,, \ \textit{eller}$$

$$\text{(ii)} \quad f(z) = \frac{a\bar{z}+b}{c\bar{z}+d}, \quad \textit{der} \quad a, b, c, d \in \mathbb{R} \quad \textit{og} \quad ad - bc = -1\,.$$

Bevis. (i) Det er klart at disse to uttrykkene begge ligger i $\mathit{Möb}(\mathbb{C})$. Men det er lett å sjekke at sammensetningen av to slike uttrykk gir et nytt uttrykk på en av disse formene, og den inverse av en funksjon på en slik form igjen har samme form. At determinanten kan velges lik 1, kommer igjen av at vi kan dividere alle koeffisientene med en kvadratrot av $ad - bc$.

(ii) Som i beviset for Proposisjon 2.1 ser vi at hvis de to uttrykkene i (i) bevarer $\overline{\mathbb{R}}$, så kan vi velge a, b, c, og d reelle. Men det samme argumentet med imaginærdelen til z gir at for at $\mathbb{H}$ skal bevares, må $ad - bc$ i det første tilfellet være positiv (som før), men i det andre tilfellet *negativ*. Normering gir derfor at den kan velges henholdsvis $+1$ og -1 som påstått. □

Vi lar $\mathit{Möb}(\mathbb{H})$ bestå av de Möbiustransformasjoner som restrikter til homeomorfier av $\mathbb{H}$. Alle elementene der kan altså skrives som en av de to typene under (ii) i proposisjon 2.2, og derfor kaller vi også elementene i $\mathit{Möb}(\mathbb{H})$ de *reelle Möbiustransformasjonene.* Legg merke til at kompleks konjugasjon, dvs. refleksjon om den reelle aksen, selv ikke er i $\mathit{Möb}(\mathbb{H})$, men $f(z) = -\bar{z}$, refleksjon om den imaginære aksen, er. Vi ser at $\mathit{Möb}(\mathbb{H})$ er generert av denne refleksjonen og $\mathit{Möb}^+(\mathbb{H})$.

Vi skal bruke notasjonen $M\ddot{o}b^-(\mathbb{H})$ for elementene i $M\ddot{o}b(\mathbb{H})$ av type (ii). $M\ddot{o}b(\mathbb{H})$ er altså en disjunkt union av $M\ddot{o}b^+(\mathbb{H})$ og $M\ddot{o}b^-(\mathbb{H})$.

Det er gruppen $M\ddot{o}b(\mathbb{H})$ vi skal bruke til å definere kongruens i $\mathbb{H}$, men før vi viser at Hilberts aksiomer C1–6 blir oppfylt, skal vi se litt nærmere på elementene i $M\ddot{o}b(\mathbb{H})$ og vise at de kan klassifiseres i noen få typer som er enklere å beskrive.

Oppgaver.

2.1. Hvilke betingelser må λ, μ, γ oppfylle for at ligningen $\lambda z\bar{z} + \mu z + \bar{\mu}\bar{z} + \gamma = 0$ skal definere en $\mathbb{H}$–linje?

2.2. La sirkelen $\mathcal{C}$ være gitt ved ligningen $|z - z_0| = r$, og la f være funksjonen $f(z) = 1/z$. Hva er betingelsen for at $f(\mathcal{C})$ skal være en rett linje $\cup\{\infty\}$?

2.3. Vis at $\dfrac{az+b}{cz+d} = \dfrac{a'z+b'}{c'z+d'}$ for alle z hvis og bare hvis det finnes en $k \neq 0$ slik at $\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = k \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

2.4. Finn en eksplisitt lineær brøktransformasjon som avbilder $\mathbb{H}$ på $\mathbb{D}$.

2.5. Beskriv alle elementer i $M\ddot{o}b(\mathbb{H})$ som avbilder den imaginære aksen på seg selv.

2.6. Vis at $M\ddot{o}b(\mathbb{H})$ virker transitivt på mengden av distinkte tripler i $\overline{\mathbb{R}}$. Gjør $M\ddot{o}b^+(\mathbb{H})$ også?

2.7. Vis at $M\ddot{o}b(\mathbb{H})$ er isomorf med gruppen $PGL_2(R) = GL_2(R)/D$, der nå $D = \{uI | u \in \mathbb{R} - \{0\}\}$.

3. Klassifikasjon av reelle Möbiustransformasjoner

Vi betrakter først undergruppen $M\ddot{o}b^+(\mathbb{H})$ av $M\ddot{o}b(\mathbb{H})$. Et element her kan altså skrives $f(z) = \dfrac{az+b}{cz+d}$, der $ad - bc = 1$. Nøkkelen til klassifikasjonen ligger i å undersøke *fikspunktene* til f betraktet som et element i $M\ddot{o}b^+(\mathbb{C})$ — dvs. løsningene i $\overline{\mathbb{C}}$ til ligningen $z = f(z)$.

Hvis $c = 0$, er ∞ et fikspunkt. Hvis i tillegg $a/d = 1$, er dette det eneste fikspunktet — ellers har vi et til, nemlig $z = -b/(a-d)$, som er et reelt tall. I denne situasjonen har vi altså enten ett eller to fikspunkter, og de ligger begge i $\overline{\mathbb{R}}$.

Hvis $c \neq 0$, leder ligningen $z = \dfrac{az+b}{cz+d}$ til annengradsligningen

$$cz^2 - (a-d)z - b = 0\,, \text{ med røtter } z = \frac{a-d \pm \sqrt{(a-d)^2 + 4bc}}{2c}\,.$$

Bruker vi nå at $ad - bc = 1$, kan vi forenkle uttrykket under rottegnet og skrive

$$z = \frac{a-d \pm \sqrt{(a+d)^2 - 4}}{2c}\,.$$

Vi ser at igjen kan vi ha ett eller to fikspunkter, begge i $\overline{\mathbb{R}}$, men vi kan også ha et par av konjugerte komplekse røtter. La oss se nærmere på de tre tilfellene:

(1) *Ett fikspunkt i* $\overline{\mathbb{R}}$. Dette har vi altså når enten $c = 0$ og $a/d = 1$ eller når $c \neq 0$ og $(a+d)^2 = 4$. Men hvis $c = 0$, gir ligningene $ad - bc = 1$ og $(a+d)^2 = 4$ at $a = d = \pm 1$, så vi ser at begge tilfellene kan uttrykkes ved betingelsen $(a+d)^2 = 4$.

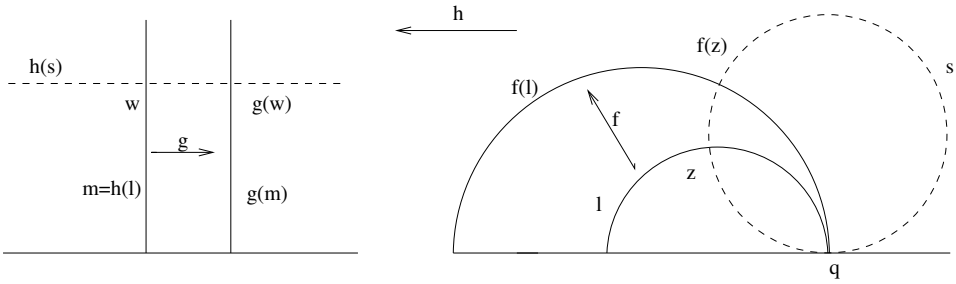
Hvis $c = 0$ kan vi skrive $f(z) = z+\beta$, dvs. f er en translasjon i x-retningen. Hvis $c \neq 0$ er det entydige fikspunktet $q = (a-d)/2c \in \mathbb{R}$, og vi kan finne en $h \in M\ddot{o}b^+(\mathbb{H})$ som avbilder q på ∞. (Velg f. eks. $h(z) = -1/(z-q)$.) Da blir sammensetningen $g = h \circ f \circ h^{-1}$ også er et element i $M\ddot{o}b^+(\mathbb{H})$, og g har ∞ som entydig fikspunkt. Det følger derfor, som over, at g har formen $g(z) = z + \gamma$ for et reelt tall γ. I dette tilfellet sier vi at f er av *parabolsk* type.

f er altså *konjugert* i gruppen $M\ddot{o}b^+(\mathbb{H})$ med en translasjon i x–retningen. En slik konjugasjon kan vi tenke på som et koordinatskifte: Skriver vi i stedet $f = h^{-1} \circ g \circ h$, ser vi at vi kan beskrive f ved at vi først flytter oss til et annet sted ved h, translaterer ved g, og så flytter tilbake igjen ved h^{-1}.

g fikserer punktet ∞ og parallellforskyver alle de $\mathbb{H}$–linjene som ender i ∞, nemlig de rette linjene ortogonalt på den reelle aksen. Siden h og h^{-1} avbilder $\mathbb{H}$–linjer på $\mathbb{H}$–linjer, ser vi at f må avbilde $\mathbb{H}$–linjer som ender i q på $\mathbb{H}$–linjer av samme type.

Figur 2 illustrerer mer detaljert hva som skjer. At g parallellforskyver vertikale linjer betyr også at den bevarer de horisontale linjene (stiplet på figuren til venstre). Avbilder vi tilbake med h^{-1}, blir bildet av disse også sirkler, men disse er nå *tangentielle* med x–aksen — dvs. som i figuren til høyre. Hvis z ligger på en slik sirkel, må derfor også $f(z)$ ligge på samme sirkel.

(Slike kurver, dvs. sirkler som tangerer $\overline{\mathbb{R}}$ i et punkt, kalles *horosirkler*. De er også karakterisert ved at de står ortogonalt på alle hyperbolske linjer som ender i dette tangeringspunktet.)



FIGUR 2. Parabolsk Möbiustransformasjon

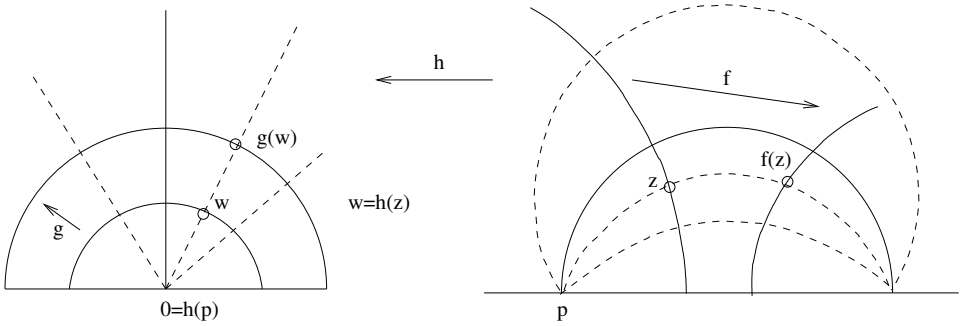
(2) Hvis f har *to fikspunkter i* $\overline{\mathbb{R}}$, sier vi at den er av *hyperbolsk* type. Hvis $c \neq 0$ skjer dette når $(a+d)^2 > 4$. Hvis $c = 0$ og $(a+d)^2 > 4$, må $ad = 1$, siden vi har antatt $ad - bc = 1$. Men da kan vi ikke ha $a = d$, så også i dette tilfellet får vi to fikspunkter. Nødvendig og tilstrekkelig betingelse er altså at $(a+d)^2 > 4$.

La nå $h \in M\ddot{o}b^+(\mathbb{H})$ avbilde de to fikspunktene på henholdsvis 0 og ∞. Da er $g = h \circ f \circ h^{-1} \in M\ddot{o}b^+(\mathbb{H})$, og den har 0 og ∞ som fiksunkter. Den vil også avbilde den imaginære aksen på en $\mathbb{H}$–linje l, men siden den må

inneholde 0 og ∞ må l også være den imaginære aksen. Spesielt finnes et positivt tall η slik at $g(i) = \eta i$, og siden vi vet at g er entydig bestemt av verdiene i de tre punktene 0, i og ∞, ser vi at vi må ha $g(z) = \eta z$ for alle z. Altså: opp til et koordinatskifte er f nå bare multiplikasjon med et reelt tall.

Hyperbolske transformasjoner er illustrert i figur 3. g bevarer rette linjer gjennom 0, og disse avbildes ved h^{-1} på sirkelbuer gjennom fikspunktene til f, men det er bare bildet av den imaginære aksen som står ortogonalt på x–aksen. Dette blir da en $\mathbb{H}$–linje mellom de to fikspunktene til f, og den avbildes på seg selv ved f. La oss kalle den *aksen* til f.

En transformasjon av hyperbolsk type med akse l, kalles ofte en "translasjon langs l".



FIGUR 3. Hyperbolsk Möbiustransformasjon

(3) Det siste tilfellet er når vi har *to komplekse fikspunkter,* og dette skjer når $(a+d)^2 < 4$. (Da kan c ikke være 0.) Siden de to fikspunktene er røttene i en ligning med reelle koeffisienter, er de konjugerte, og vi har nøyaktig ett i det øvre halvplan og ingen i $\overline{\mathbb{R}}$. Kall dette fikspunktet p og la $h(z) = \dfrac{z - \operatorname{Re} p}{\operatorname{Im} p}$, slik at $h \in \mathit{Möb}^+(\mathbb{H})$ og $h(z_0) = i$. Denne gangen vil $g = h \circ f \circ h^{-1}$ være en reell lineær brøktransformasjon som fikserer i. La oss finne ut hvordan en slik g må se ut.

Sett $g(z) = \dfrac{\alpha z + \beta}{\gamma z + \delta}$, der α, β, γ, $\delta \in \mathbb{R}$. $g(i) = i$ betyr da at $\alpha i + \beta = -\gamma + \delta i$ — dvs. $\alpha = \delta$ og $\beta = -\gamma$. Setter vi dette inn i ligningen $\alpha\delta - \beta\gamma = 1$, ser vi at $\alpha^2 + \beta^2 = 1$. Da kan vi skrive $\alpha = \cos(\theta)$, $\beta = \sin(\theta)$, for et reelt tall θ, og vi har

$$g(z) = \frac{\cos(\theta) z + \sin(\theta)}{-\sin(\theta) z + \cos(\theta)}\,.$$

I dette tilfellet sier vi at f er av *elliptisk* type.

Matrisen $\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$ beskriver rotasjon med en vinkel θ i $\mathbb{R}^2$. Vi kan tenke på g også som en slags rotasjon, for den vil jo avbilde $\mathbb{H}$–linjene gjennom i på andre linjer av samme type. Når vi senere har innført Poincarés diskmodell, skal vi se analogien enda tydeligere. Her nøyer vi oss med å bemerke at mens vi for rotasjoner i $\mathbb{R}^2$ kommer tilbake til identiteten når θ har økt med 2π, så er det her nok å legge til π. Det kommer av at

$$\frac{\cos(\theta+\pi)z+\sin(\theta+\pi)}{-\sin(\theta+\pi)z+\cos(\theta+\pi)} = \frac{\cos(-\theta)z+\sin(-\theta)}{-\sin(-\theta)z+\cos(-\theta)}$$
$$= \frac{-\cos(\theta)z+(-\sin(\theta))}{\sin(\theta)z+(-\cos(\theta))} = \frac{\cos(\theta)z+\sin(\theta)}{-\sin(\theta)z+\cos(\theta)}$$

For å oppsummere så langt, har vi altså:

Hvis $f(z)=\dfrac{az+b}{cz+d}$, *der* a, b, c, $d\in\mathbb{R}$ *og* $ad-bc=1$, *så er* f *av*

- *Parabolsk type hvis* $(a+d)^2=4$
- *Hyperbolsk type hvis* $(a+d)^2>4$
- *Elliptisk type hvis* $(a+d)^2<4$

Størrelsen $a+d$ kalles som kjent *sporet* til matrisen $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Enhver $f\in M\ddot{o}b^+(\mathbb{H})$ er av en av disse typene, og en viktig konsekvens er at hvis to slike funksjoner er like i to punkter i $\mathbb{H}$, så er de like overalt!

Vi betrakter nå et element i $M\ddot{o}b^-(\mathbb{H})$ — dvs. en avbildning på formen $f(z)=\dfrac{a\bar{z}+b}{c\bar{z}+d}$, der a, b, c, $d\in\mathbb{R}$ og $ad-bc=-1$.

Hvis $c=0$ er $ad=-1$, og vi får $f(z)=-a^2\bar{z}-ab=a^2(-\bar{z}+b/a)$. Dette er en sammensetning av refleksjon om $\mathbb{H}$–linjen $x=b/2$ og multiplikasjon med det positive tallet a^2.

Hvis $c\neq 0$, så kan vi skrive (utregning)

$$f(z)=\frac{a}{c}+\frac{\frac{1}{c}}{c\bar{z}+d}=\frac{a+d}{c}+\left(-\frac{d}{c}+\frac{1}{c^2}\frac{z-(-\frac{d}{c})}{|z-(-d/c)|^2}\right) \tag{3.1}$$

Uttrykket i en siste parentesen ser kanskje litt komplisert ut, men funksjonen er i virkeligheten ganske enkel. Skriver vi mer generelt

$$w=m+r^2\frac{z-m}{|z-m|^2}=m+\frac{r^2}{\overline{z-m}}\,,$$

ser vi at w ligger et sted på strålen fra m i retning z, og slik at produktet $|w-m||z-m|$ er lik r^2. Dette er en viktig geometrisk konstruksjon som kalles "inversjon i sirkelen med sentrum m og radius r". Den har spesielt den egenskapen at den avbilder punkter utenfor sirkelen på punkter innenfor sirkelen og omvendt, og den holder sirkelen selv i ro.

Vi har altså fått at opp til en translasjon, er f gitt som inversjon i sirkelen med sentrum i $-d/c$ og radius $1/|c|$. Men dette er en $\mathbb{H}$–linje. Inversjonen fikserer denne $\mathbb{H}$–linjen og gir en homeomorfi mellom de to komponentene den deler $\mathbb{H}$ i.

Etter eventuelt å ha skiftet tegn på a, b, c og d, kan vi anta $c>0$. Da definerer $h(z)=cz+d$ et element i $M\ddot{o}b^+(\mathbb{H})$, og vi kan skrive denne inversjonen som $h^{-1}(z)J(z)h(z)$, der $J(z)=1/\bar{z}=\bar{z}/|z|^2$ er inversjon i enhetssirkelen $|z|=1$. (Sjekk ved utregning.) Derfor er alle inversjoner konjugerte med J, og dermed også konjugerte innbyrdes.

Vi kan også oppfatte Möbiustransformasjonen $z \mapsto -\bar{z}$ som en inversjon. La nemlig $g(z) = \dfrac{z+1}{-z+1}$, som tar $\mathbb{H}$–linjen $|z| = 1$ på den imaginære aksen. Da er det lett å sjekke ved utregning at $g(J(g^{-1}))(z) = -\bar{z}$, så $z \mapsto -\bar{z}$ er også konjugert med inversjonen J. Tilsvarende får vi naturligvis for refleksjoner i alle andre vertikale linjer. Vi kan derfor også kalle disse for inversjoner. Da har vi definert en inversjon i alle $\mathbb{H}$-linjer, og de er alle konjugerte. Vi merker oss følgende, som er implisitt i det vi har gjort:

- Et element i $M\ddot{o}b^-(\mathbb{H})$ er en ren inversjon hvis og bare hvis sporet $a+d=0$. (Denne betingelsen er uavhengig av om vi har normalisert koeffisientene eller ikke.)
- En inversjon i en $\mathbb{H}$–linje l er karakterisert, som element i $M\ddot{o}b(\mathbb{H})$, ved at den har hele l som fikspunktmengde.

Oppgaver.

3.1. Klassifiser følgende avbildninger og skriv dem eksplisitt som konjugerte til avbildninger på standard form.

$$\frac{4z-3}{2z-1}, \qquad -\frac{1}{z-1}, \qquad \frac{z}{z+1}$$

3.2. Beskriv klassifikasjonen av Möbiustransformasjonene ved hjelp av matriserepresentasjonen.

3.3. Forklar hva en hyperbolsk transformasjon f gjør med horosirklene til aksen til f, og også med de andre $\mathbb{H}$–linjene som har et endepunkt felles med denne aksen.

3.4. Vis at $f \in M\ddot{o}b^-(\mathbb{H})$ er en inversjon hvis og bare hvis den har et fikspunkt i $\mathbb{H}$.

3.5. Vis at en inversjon i en sirkel $\mathcal{C} \subset \mathbb{C}$, betraktet som en avbildning på $\mathbb{C}$ minus sentrum i $\mathcal{C}$, har følgende egenskaper:

(a) Den avbilder linjer utenfor $\mathcal{C}$ på sirkler innenfor $\mathcal{C}$ og gjennom sentrum.
(b) Den avbilder sirkler som skjærer $\mathcal{C}$ ortogonalt, på seg selv.

(Dette er viktige reultater om inversjoner som vanligvis gis rent geometriske bevis. Her skal de følge de ganske enkelt av det vi nå vet om Möbiustransformasjoner.)

4. Hilberts kongruensaksiomer i $\mathbb{H}$

Vi trenger først litt mer presis notasjon for linjer, stråler etc. Hvis z_1, z_2 er to punkter i $\mathbb{H}$, så skriver vi $\overleftrightarrow{z_1 z_2}$ for den entydig bestemte hyperbolske linjen som går gjennom dem, $\overrightarrow{z_1 z_2}$ for strålen fra z_1 gjennom z_2, og $[z_1, z_2]$ for segmentet med z_1 og z_2 som endepunkter ($= \overrightarrow{z_1 z_2} \cap \overrightarrow{z_2 z_1}$). En $\mathbb{H}$–linje l er entydig bestemt av sine endepunkter p og q i $\overline{\mathbb{R}}$, og vi vil derfor også skrive $l = (p, q)$. F. eks. vil da identiteten $\overleftrightarrow{z_1 z_2} = (p, q)$ bety at den entydig bestemte $\mathbb{H}$–linjen som går gjennom z_1 og z_2 har p og q som endepunkter. På lignende måte vil vi også skrive $[z_1, q) = \overrightarrow{z_1 q} = \overrightarrow{z_1 z_2}$, dersom q er endepunktet for strålen $\overrightarrow{z_1 z_2}$. Vi vil si at z_1 er *hjørnet* og q *endepunktet* for strålen $[z, q)$. En *vinkel* er da et uordnet par av stråler med samme hjørne, der de to strålene

ligger på samme linje. Vi skal bruke notasjonen $\angle uzv$, der $z \in \mathbb{H}$ og u, v er enten i $\mathbb{H}$ eller $\overline{\mathbb{R}}$.

Vi minner om at vi nå definerer kongruens ved

Kongruens av segmenter: $[z_1, z_2] \cong [w_1, w_2] \iff$ det finnes en $g \in \mathit{Möb}(\mathbb{H})$ slik at $g([z_1, z_2]) = [w_1, w_2]$.

Kongruens av vinkler: $\angle uzv \cong \angle u'z'v' \iff$ det finnes en $g \in \mathit{Möb}(\mathbb{H})$ slik at $g(\overrightarrow{zu}) = \overrightarrow{z'u'}$ og $g(\overrightarrow{zv}) = \overrightarrow{z'v'}$. (Vi vil da skrive $g(\angle uzv) = \angle u'z'v'$.)

Når vi nå skal prøve å vise kongruensaksiomene, vil vi se at eksistensutsagnene krever at vi har nok Möbiustransformasjoner til å flytte vinkler og segmenter fritt rundt i $\mathbb{H}$, mens entydigheten er enklest dersom det ikke er for mange valg av slike transformasjoner. De resultatene vi trenger, er formulert i følgende lemmaer:

Lemma 4.1. *Anta z_j ligger på en $\mathbb{H}$–linje l_j, med endepunkter p_j og q_j, for $j = 1, 2$. Da finnes entydig bestemt $f \in \mathit{Möb}^+(\mathbb{H})$ slik at $f(p_1) = p_2$, $f(q_1) = q_2$ og $f(z_1) = z_2$.*

Spesielt får vi at, f. eks., $f([z_1, q_1)) = [z_2, q_2)$. På den annen side bestemmer en stråle fullstendig linjen den er en del av, så vi får

Korollar 4.1. *$\mathit{Möb}^+(\mathbb{H})$ virker transitivt på mengden av stråler: Gitt to stråler σ_1 og σ_2 med hjørner z_1 og z_2, så fins entydig $f \in \mathit{Möb}^+(\mathbb{H})$ slik at $f(z_1) = z_2$ og $f(\sigma_1) = \sigma_2$.*

Lemma 4.2. *Anta at segmentene $[z_1, z_2]$ og $[w_1, w_2]$ er kongruente. Da finnes entydig bestemt $f \in \mathit{Möb}^+(\mathbb{H})$ slik at $f(z_1) = w_1$ og $f(z_2) = w_2$.*

Lemma 4.3. *Gitt to stråler σ_1 og σ_2 med samme hjørne z_0. Da fins entydig bestemt inversjon g slik at $g(z_0) = z_0$, $g(\sigma_1)) = \sigma_2$ og $g(\sigma_2) = \sigma_1$.*

Bevis for lemma 4.1. Vi har allerede vist (korollar 2.1 at det finnes en entydig bestemt LBT f med egenskaper som påstått, så vi må bare vise at den kan velges i $\mathit{Möb}^+(\mathbb{H})$.

Anta først at $(p_2, z_2, q_2) = (0, i, \infty)$, slik at l_2 er den imaginære aksen. Hvis verken p_1 eller q_1 er ∞, så kan vi sette

$$f(z) = \frac{z - p_1}{z - q_1} \frac{z_1 - q_1}{z_1 - p_1} i\,.$$

Siden z_1 ligger på en halvsirkel med diameter utspent av p_1 og q_1, så er $z_1 - p_1$ og $z_1 - q_1$ ortogonale. Derfor er $\dfrac{z_1 - p_1}{z_1 - q_1}$ rent imaginær og $\dfrac{z_1 - p_1}{z_1 - q_1}\, i$ reell. Det følger at $f \in \mathit{Möb}^+(\mathbb{H})$.

Hvis $p_1 = \infty$ må $\operatorname{Re} z_1 = q_1$, så vi kan sette $f(z) = -\operatorname{Im} z_1/(z - q_1)$.

Hvis $q_1 = \infty$, kan vi på tilsvarende måte sette $f(z) = (z - p_1)/\operatorname{Im} z_1$.

I det generelle tilfellet kan vi finne f_1 og f_2 slik at $f_j(p_j, z_j, q_j) = (0, i, \infty)$ for $j = 1, 2$. Sett $f = f_2^{-1} \circ f_1$. □

Bevis for lemma 4.2. Entydighet følger av at et element i $\mathit{Möb}^+(\mathbb{H})$ er fullstendig bestemt av verdien i to punkter.

Vi har gitt at $g([z_1, z_2]) = [w_1, w_2]$ for en $g \in \mathit{Möb}(\mathbb{H})$. Hvis $g \in \mathit{Möb}^-(\mathbb{H})$, så erstatter vi g med $k \circ g$, der k er inversjon i $\mathbb{H}$–linjen $\overleftrightarrow{w_1 w_2}$. Derfor kan vi anta $g \in \mathit{Möb}^+(\mathbb{H})$.

Problemet nå er at vi kan ha $g(z_1) = w_2$ og $g(z_2) = w_1$. Velg nå en $h \in M\ddot{o}b^+(\mathbb{H})$ slik at $h(\overleftrightarrow{w_1 w_2})$ er den imaginære aksen, og skriv $h(w_1) = \omega_1 i,\ h(w_2) = \omega_2 i$. Setter vi $k(z) = -\omega_1\omega_2/z$, ser vi at k bytter om $\omega_1 i$ og $\omega_2 i$. Da vil $h^{-1}kh$ bytte om w_1 og w_2, og vi kan sette $f = h^{-1}khg$. □

Bevis for lemma 4.3. La de to strålene ha endepunktene q_1 og q_2. Strålene bestemmer $\mathbb{H}$–linjer l_1 og l_2, og de andre endepunktene for disse kaller vi henholdsvis p_1 og p_2. Det finnes da, ved lemma 4.1, en $h \in M\ddot{o}b^+(\mathbb{H})$ slik at $h(p_1) = 0,\ h(q_1) = \infty$ og $h(z_0) = i$. ($h(l_1) = l_1'$ blir den imaginære aksen.) Sett $h(p_2) = p$ og $h(q_2) = q$. $h(l_2) = l_2' = (p, q)$ blir da en annen $\mathbb{H}$–linje gjennom i. Det betyr spesielt at verken p eller q er ∞.

Vi setter nå $f(z) = \dfrac{q\bar{z} - qp}{\bar{z} - q}$. Sporet til denne representasjonen av f er $q + (-q) = 0$, så for at f skal være en inversjon på $\mathbb{H}$, trenger vi bare at determinanten $-q^2+qp$ er negativ. Men siden (q, p) er en $\mathbb{H}$–linje som krysser den imaginære aksen (i i), må p og q ha motsatt tegn. Derfor er $qp < 0$, og dermed også $-q^2 + qp < 0$.

Det er lett å sjekke at $f(q) = \infty$ og $f(\infty) = q$, og at $f(p) = 0$ og $f(0) = p$. Siden f avbilder $\mathbb{H}$–linjer på $\mathbb{H}$–linjer, må derfor $f(l_1') = l_2'$ og $f(l_2') = l_1'$. Men da må også $f(i) = f(l_1' \cap l_2') = l_1' \cap l_2' = i$, så vi får i alt at $f([i, q)) = [i, \infty)$ og $f([i, \infty)) = [i, q)$.

$g = h^{-1}fh$ vil da være en inversjon som bytter om $\sigma_1 = [z_0, q_1)$ og $\sigma_2 = [z_0, q_2)$.

Entydighet: Anta g' er en annen inversjon med de samme egenskapene. Da er $g^{-1}g'$ et element i $M\ddot{o}b^+(\mathbb{H})$ som avbilder både σ_1 og σ_2 på seg selv. Derfor vil den ha de tre fikspunktene z_0, q_1 og q_2 (i $\overline{\mathbb{C}}$) og er derfor lik identitetsavbildningen. Det følger at $g = g'$.

□

Vi er nå klar til å vise at Hilberts kongruensaksiomer C1-6 gjelder. La oss først repetere aksiomene.

Aksiomene for kongruens av segmenter er:

- **C1:** Gitt et segment $[z_1, z_2]$ og en stråle σ med hjørne w_1, så fins et entydig bestemt punkt w_2 på σ slik at $[w_1, w_2] \cong [z_1, z_2]$.
- **C2:** $\cong$ er en ekvivalensrelasjon på mengden av segmenter.
- **C3:** Hvis $z_1 * z_2 * z_3$ og $w_1 * w_2 * w_3$ og vi har både $[z_1, z_2] \cong [w_1, w_2]$ og $[z_2, z_3] \cong [w_2, w_3]$, så er også $[z_1, z_3] \cong [w_1, w_3]$.

Aksiomene for kongruens av vinkler er:

- **C4:** Gitt stråle $[w, q)$ og vinkel $\angle uzv$, så fins entydig bestemte vinkler $\angle p_1wq$ og $\angle p_2wq$ på motsatte sider av $[w, q)$ slik at $\angle p_1wq \cong \angle p_2wq \cong \angle uzv$.
- **C5:** $\cong$ er en ekvivalensrelasjon på mengden av vinkler.
- **C6:** (SAS) Gitt trekanter $z_1z_2z_3$ og $w_1w_2w_3$. Hvis $[z_1, z_2] \cong [w_1, w_2]$, $[z_1, z_3] \cong [w_1, w_3]$ og $\angle z_2z_1z_3 \cong \angle w_2w_1w_3$, så er de to trekantene kongruente — dvs. vi har også $[z_2, z_3] \cong [w_2, w_3]$, $\angle z_1z_2z_3 \cong \angle w_1w_2w_3$ og $\angle z_2z_3z_1 \cong \angle w_2w_3w_1$.

C2 og **C5** følger umiddelbart, siden vi har definert kongruens ved en gruppevirkning.

C1. Segmentet $[z_1, z_2]$ definerer stråle $\overrightarrow{z_1 z_2}$, og korollar 4.1 gir at det finnes en $f \in M\ddot{o}b^+(\mathbb{H})$ slik at $f(z_1) = w_1$ og $f(\overrightarrow{z_1 z_2}) = \sigma$. Setter vi $w_2 = f(z_2)$, får vi åpenbart $[w_1, w_2] \cong [z_1, z_2]$.

Anta at w_2' er et annet punkt på σ slik at $[w_1, w_2'] \cong [z_1, z_2]$. Etter Lemma 4.2 kan vi finne $h \in M\ddot{o}b^+(\mathbb{H})$ slik at $h(z_1) = w_1$ og $h(z_2) = w_2'$. Men da er også $h(\overrightarrow{z_1 z_2}) = \overrightarrow{w_1 w_2'} = \sigma$, og på grunn av entydigheten i korollar 4.1, må vi ha $h = f$. Derfor er også $w_2 = f(z_2) = h(z_2) = w_2'$.

C3. Her bruker vi lemma 4.2, som sier at det finnes $g \in M\ddot{o}b^+(\mathbb{H})$ slik at $g(z_1) = w_1$ og $g(z_2) = w_2$. Da ligger $w_3' = g(z_3)$ på strålen $\overrightarrow{w_2 w_3}$, og g definerer kongruenser $[z_2, z_3] \cong [w_2, w_3']$ og $[z_1, z_3] \cong [w_1, w_3']$. Men på grunn av entydigheten i C1, må $w_3' = w_3$.

C4. Vi kan anta at u og v er endepunktene for strålene i $\angle uzv$. Å konstruere vinklene er enkelt: la $g \in M\ddot{o}b^+(\mathbb{H})$ være slik at $g([z, u)) = [w, q)$, og definer $\angle p_1 w q$ som $g(\angle uzv)$. La $j \in M\ddot{o}b^-(\mathbb{H})$ være inversjon i den $\mathbb{H}$–linjen l som $[w, q)$ er en del av, og sett $\angle p_2 w q = jg(\angle uzv) = j(\angle p_1 w q)$. Siden j bytter om de to sidene av l, vil $\angle p_1 w q$ og $\angle p_2 w q$ ligge på hver sin side av $[w, q)$.

Det gjenstår å vise entydighet. Anta $\angle pwq = h(\angle uzv)$, der $h \in M\ddot{o}b(\mathbb{H})$. Dette betyr at $h([z, u))$ er enten $[w, q)$ eller $[w, p)$, men etter lemma 4.3 kan vi eventuelt sette sammen h med en inversjon som bytter om $[w, q)$ og $[w, p)$, så vi kan anta at $h([z, u)) = [w, q)$. Hvis $h \in M\ddot{o}b^+(\mathbb{H})$, må da $h = g$, på grunn av entydigheten i korollar 4.1 igjen, og $[w, p) = [w, p_1)$. Hvis $h \in M\ddot{o}b^-(\mathbb{H})$, så er $jh \in M\ddot{o}b^+(\mathbb{H})$, og siden $jh([z, u)) = j([w, q)) = [w, q)$, er nå $jh = g$. Derfor er $h = jg$, og $[w, p) = [w, p_2)$.

C6. Ved å bruke lemma 4.3, kan vi anta at $\angle w_2 w_1 w_3 = g(\angle z_2 z_1 z_3)$ med $g \in M\ddot{o}b^+(\mathbb{H})$, og etter eventuelt å ha byttet om w_2 og w_3 i notasjonen, kan vi anta at $g(\overrightarrow{z_1 z_2}) = \overrightarrow{w_1 w_2}$ og $g(\overrightarrow{z_1 z_3}) = \overrightarrow{w_1 w_3}$. Men entydigheten i C1 betyr da at $g(z_2) = w_2$ og $g(z_3) = w_3$. Da er åpenbart også $g([z_2, z_3]) = [w_2, w_3]$, $g(\angle z_1 z_2 z_3) = \angle w_1 w_2 w_3$ og $g(\angle z_2 z_3 z_1) = \angle w_2 w_3 w_1$.

□

5. Avstandsbegrepet i $\mathbb{H}$

Vi skal definere en *metrikk* på $\mathbb{H}$ — dvs. en funksjon $d : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ slik at

(d1) $d(z, w) \geq 0$, og $d(z, w) = 0$ hvis og bare hvis $z = w$
(d2) $d(z, w) = d(w, z)$ for alle $z, w \in \mathbb{H}$
(d3) $d(z, w) \leq d(z, u) + d(u, w)$ for alle $z, u, w \in \mathbb{H}$

Metrikken skal naturligvis også ha noe med geometrien å gjøre, så i tillegg skal den ha egenskapene

(d4) Hvis z, u, w er distinkte punkter i $\mathbb{H}$, så er $d(z, w) = d(z, u) + d(u, w)$ hvis og bare hvis $z * u * w$. ("Avstand måles langs $\mathbb{H}$–linjer".)
(d5) $d(z, w) = d(z', w')$ hvis og bare hvis det finnes $g \in M\ddot{o}b(\mathbb{H})$ slik at $g(z) = z'$ og $g(w) = w'$.
("Möbiustransformasjonene er isometrier".)

Først skal vi se at (d4) og (d5) bestemmer metrikken nesten entydig. Gitt z, w, så finnes entydig $g \in M\ddot{o}b^+(\mathbb{H})$ slik at $g(z) = i$ og $g(w) = yi$, der $y \geq 1$.

(Dette er essensielt lemma 4.1 anvendt på punktet i på den imaginære aksen og punktet z på $\mathbb{H}$–linjen gjennom z og w. Hvis $g(w)$ skulle ha imaginærdel mindre enn 1, så kan vi erstatte g med kg, der $k(z) = -1/z$.) Da må, på grunn av (d5), $d(z,w) = d(i,yi)$, så d er fullstendig bestemt av funksjonen $f : [1,\infty) \to [0,\infty)$ definert ved $f(t) = d(i,ti)$.

Vi kan si mer om funksjonen f. La s og t være to tall større eller lik 1. Da ligger si mellom i og sti, så etter (d4) har vi:

$$f(st) = d(i,sti) = d(i,si) + d(si,sti) = d(i,si) + d(i,ti) = f(s) + f(t)\,.$$

(Den tredje likheten kommer av at etter (d5) skal vi ha $d(h(i),h(ti)) = d(i,ti)$, der $h(z) = sz$.) Men man kan vise at da må $f(t) = C\ln(t)$, for en reell konstant C. (Se oppgave 5.1 for det tilfellet at f er deriverbar i 1.)

Ethvert tall $C > 0$ kunne brukes i det følgende, men vi velger å sette $C = 1$. Hvis ti og si er to punkter på den imaginære aksen med $s < t$, så får vi da

$$d(si,ti) = d(i,\frac{t}{s}i) = \ln(\frac{t}{s})\,.$$

Bytter vi om si og ti, får vi også

$$d(ti,si) = d(si,ti) = \ln(\frac{t}{s}) = -\ln(\frac{s}{t})\,,$$

slik at vi alltid kan definere $d(si,ti)$ som $|\ln(\frac{t}{s})|$ for to punkter på den imaginære aksen. Spesielt er $d(i,ti) = |\ln(t)|$ for *alle* $t > 0$.

Vår definisjon av funksjonen d er altså

$$d(z,w) = |\ln(|\frac{g(z)}{g(w)}|)| = |\ln(|\frac{i}{g(w)}|)| = |\ln(|g(w)|)|\,,$$

der $g \in \mathit{Möb}^+(\mathbb{H})$ avbilder z på i og w på et annet punkt på den imaginære aksen.

I beviset for lemma 4.1 fant vi et eksplisitt uttrykk for en slik g. Vi kan formulere resultatet i denne situasjonen slik:

Hvis $\operatorname{Re} z = \operatorname{Re} w$, så er enten $g(w) = -\operatorname{Im} z/(w - \operatorname{Re} z) = -\operatorname{Im} z/(\operatorname{Im} wi) = (\operatorname{Im} z/\operatorname{Im} w)i$ eller $g(w) = (w - \operatorname{Re} w)/\operatorname{Im} z = (\operatorname{Im} w/\operatorname{Im} z)i$. I begge tilfeller får vi at vi kan skrive $d(z,w) = |\ln(|\operatorname{Im} w/\operatorname{Im} z|)|$.

Hvis $\operatorname{Re} z \neq \operatorname{Re} w$, så har $\mathbb{H}$–linjen gjennom z og w endepunkter p og q i $\mathbb{R}$, og vi har enten

$$g(w) = \frac{w-p}{w-q}\,\frac{z-q}{z-p}\,i \quad \text{eller} \quad g(w) = \frac{w-q}{w-p}\,\frac{z-p}{z-q}\,i\,.$$

Koeffisientene foran i må være reelle tall, siden de begge er produkt av to kvosienter som vi observerte var rent imaginære. (Se beviset for lemma 4.1.) Men igjen ser vi at de to mulige kvosientene da er inverse av hverandre, så vi har

$$d(z,w) = |\ln(|\frac{w-p}{w-q}\,\frac{z-q}{z-p}|)|$$

Dermed har vi eksplisitte uttrykk for $d(z,w)$ i de to situasjonene $\operatorname{Re} z = \operatorname{Re} w$ og $\operatorname{Re} z \neq \operatorname{Re} w$. Vi må nå vise at funksjonen d tilfredsstiller (d1-5).

(d1) er opplagt oppfylt, og (d2) følger av formlene vi nettopp utledet. La oss dernest se på (d5):

La først f være et element i $M\ddot{o}b(\mathbb{H})$ og betrakt $d(f(z), f(w))$, der vi antar $z \neq w$. (Tilfellet $z = w$ er trivielt.) Ved eventuelt å sette sammen med inversjon i $\mathbb{H}$–linjen gjennom $f(z)$ og $f(w)$, kan vi anta $f \in M\ddot{o}b^+(\mathbb{H})$. Hvis nå $g \in M\ddot{o}b^+(\mathbb{H})$ tar $f(z)$ på i og $f(z)$ på den imaginære aksen, har vi $d(f(z), f(w)) = |\ln(|gf(w)|)|$. Men $gf \in M\ddot{o}b^+(\mathbb{H})$ og tar z på i og w på den imaginære aksen, så vi har også $d(z, w) = |\ln(|gf(w)|)| = d(f(z), f(w))$.

Anta så at $d(z, w) = d(z', w')$. La $g, g' \in M\ddot{o}b^+(\mathbb{H})$ slik at $g(z) = g'(z') = i,\ g(w) = ti$ og $g'(w') = t'i$, der t og t' er ≥ 1. Da er

$$\ln(t) = d(z, w) = d(z'w') = \ln(t')\,,$$

så $t = t'$. Sett $f = g^{-1}g'$. Da er $f \in M\ddot{o}b^+(\mathbb{H})$ og $f(z') = z,\ f(w') = w$.

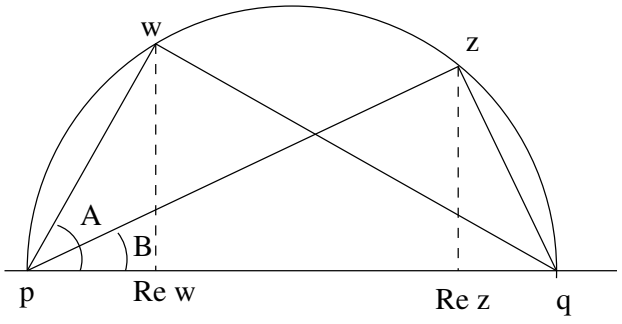
Det gjenstår å vise trekantulikheten (d3) med den ekstra egenskapen (d4). Vi skal få bruk for følgende lemma:

Lemma 5.1. *For alle z, w gjelder $d(z, w) \geq |\ln(\operatorname{Im} z/\operatorname{Im} w)|$, med likhet hvis og bare hvis $\operatorname{Re} z = \operatorname{Re} w$.*

Bevis. Hvis $\operatorname{Re} z = \operatorname{Re} w$ har vi sett at $d(z, w) = |\ln(\operatorname{Im} z/\operatorname{Im} w)|$, så vi antar $\operatorname{Re} z \neq \operatorname{Re} w$. Ulikheten er trivielt oppfylt hvis $\operatorname{Im} z = \operatorname{Im} w$ (ekte ulikhet hvis $z \neq w$), så vi kan anta $\operatorname{Im} z \neq \operatorname{Im} w$, og siden $d(z, w) = d(w, z)$, kan vi også anta $\operatorname{Im} w > \operatorname{Im} z$. Da er

$$d(z, w) = |\ln(|\frac{w-p}{w-q}\,\frac{z-q}{z-p}|)| = |\ln(\frac{|w-p|}{|w-q|}/\frac{|z-p|}{|z-q|})|\,,$$

der p og q er endepunktene for $\overleftrightarrow{zw}$. Dette uttrykket er invariant under ombytting av p og q, så vi kan anta at $p * w * z * q$. Vi har da situasjonen illustrert på figur 4.



FIGUR 4

Av figuren ser vi at $\dfrac{|w-q|}{|w-p|} = \tan A$ og $\dfrac{|z-q|}{|z-p|} = \tan B$. Men vi har også $\tan A = \operatorname{Im} w/|\operatorname{Re} w - p|$ og $\tan B = \operatorname{Im} z/|\operatorname{Re} z - p|$, så vi får

$$\frac{|w-q|}{|w-p|}/\frac{|z-q|}{|z-p|} = \frac{\tan A}{\tan B} = \frac{\operatorname{Im} w}{\operatorname{Im} z} \cdot \frac{|\operatorname{Re} z - p|}{|\operatorname{Re} w - p|}\,.$$

Måten vi har valgt p, w, z og q på, sikrer at både $\dfrac{\operatorname{Im} w}{\operatorname{Im} z} > 1$ og $\dfrac{|\operatorname{Re} z - p|}{|\operatorname{Re} w - p|} > 1$. Derfor er $\dfrac{|w-q|}{|w-p|} / \dfrac{|z-q|}{|z-p|} > \dfrac{\operatorname{Im} w}{\operatorname{Im} z} > 1$, og

$$d(z,w) = |\ln(\frac{|w-p|}{|w-q|} / \frac{|z-p|}{|z-q|})| = \ln(\frac{|w-q|}{|w-p|} / \frac{|z-q|}{|z-p|}) > \ln(\frac{\operatorname{Im} w}{\operatorname{Im} z})$$

Vi har altså ekte ulikhet når $\operatorname{Re} z \neq \operatorname{Re} w$ og likhet når $\operatorname{Re} z = \operatorname{Re} w$. □

For å vise trekantulikheten er det nok å betrakte tilfellet $z = i$, $w = ti$, der $t > 1$, siden vi alltid kan flytte oss over i denne situasjonen ved å bruke (d5). Hvis u er at annet punkt, så får vi

$$\begin{aligned} d(z,u) + d(u,w) &\geq |\ln(\operatorname{Im} z/\operatorname{Im} u)| + |\ln(\operatorname{Im} u/\operatorname{Im} w)| \geq \\ &|\ln(\operatorname{Im} z/\operatorname{Im} u) + \ln(\operatorname{Im} u/\operatorname{Im} w)| = |\ln(\operatorname{Im} z/\operatorname{Im} w)| = d(z,w)\,. \end{aligned}$$

Dette er (d3). Likhet har vi hvis og bare hvis $\operatorname{Re} u = 0$ og $\ln(\operatorname{Im} z/\operatorname{Im} u)$ og $\ln(\operatorname{Im} u/\operatorname{Im} w)$ har samme tegn — dvs. hvis og bare hvis u også ligger på den imaginære aksen og enten $1 = \operatorname{Im} z < \operatorname{Im} u < \operatorname{Im} w$ eller $1 = \operatorname{Im} z > \operatorname{Im} u > \operatorname{Im} w$. Men det betyr nettopp at $z * u * w$, så (d5) følger. □

Oppgaver.

5.1 Vis at en funksjon $f : (0,\infty) \to \mathbb{R}$ som er deriverbar i 1 og tilfredsstiller ligningen $f(st) = f(s)+f(t)$ for alle s og t må være lik $C\ln(t)$ for en konstant C.

5.2 Vis at metrikken d definerer den vanlige topologien (underromstopologien fra $\mathbb{R}^2$) i det øvre halvplan. (Dette er ikke helt lett på det nåværende stadium, men vil være enklere etter kapittel 7. Se oppgave 7.2)

6. Vinkelmål. $\mathbb{H}$ som conform modell

Vanlig vinkelmål i $\mathbb{R}^2$ er en funksjon som til enhver vinkel tilordner et tall mellom 0 og $2R$ (for 2 rette), slik at to vinkler er kongruente (i Euklidsk forstand) hvis og bare hvis de tilordnes det samme tallet. Dessuten er tilordningen *additiv,* i den forstand at hvis to vinkler tilordnes tallene U og V, så vil summen av dem tilordnes $U + V$.

Det siste kan presiseres slik: hvis vinkelen $\angle ABC$ tilordnes tallet U og $\angle CBD$ tallet V, der A og D er på motsatt side av linjen $\overleftrightarrow{BC}$ og $U+V < 2R$, så tilordnes $\angle ABD$ tallet $U + V$. (Betingelsene sikrer at vi snakker om *sum* og ikke *differens* av vinkler, og at vi bare snakker om vinkler mindre enn to rette.) Normerer vi vinkelmålet, f. eks. ved å kreve at rette vinkler tilordnes målet R og at alle tall i intervallet $(0, 2R)$ skal oppnås, så bestemmer dette vinkelmålet entydig. Det er altså essensielt *kongruensbegrepet* som bestemmer vinkelmålet.

I $\mathbb{H}$ har vi definert kongruens som ekvivalens ved reelle Möbiustransformasjoner. Men lemma 2.1 sier at lineære brøktransformasjoner bevarer *Euklidsk* vinkelmål, og det er trivielt at kompleks konjugasjon gjør det samme. Derfor vil alle Möbiustransformasjoner, og spesielt de reelle, også bevare Euklidsk vinkelmål. Det følger at kongruens av Euklidske vinkler også kan karakteriseres som ekvivalens ved Möbiustransformasjoner. Derfor

kan vi bruke det samme vinkelmålet i $\mathbb{H}$ som vi har i det Euklidske planet som $\mathbb{H}$ er en delmengde av.

Vi uttrykker dette ved å si at Poincarés øvre halvplanmodell $\mathbb{H}$ er en vinkeltro eller *konform* modell. Dette er en av de egenskapene som gjør den så mye bedre og nyttigere enn f. eks. Beltrami–Kleins modell, som ikke er vinkeltro.

Siden stereografisk projeksjon også bevarer vinkler, følger det at halvsfæremodellen $\mathbb{B}$ er konform, og derfor også alle andre modeller vi får fra den igjen, ved andre stereografiske projeksjoner. Spesielt gjelder dette for Poincarés diskmodell $\mathbb{D}$, som vi skal studere nærmere i neste kapittel.

7. Poincarés diskmodell $\mathbb{D}$

Vi vil nå overføre det vi har gjort med øvre halvplanmodellen $\mathbb{H}$ for det hyperbolske plan til diskmodellen $\mathbb{D}$. Dette kan vi gjøre med homeomorfien $G = \Phi \circ F \circ \Phi^{-1}$, der Φ er stereografisk projeksjon og F er en identifikasjon av halvsfærene $B_1 = \{(x, y, z) | y > 0\}$ og $B_2 = \{(x, y, z) | z < 0\}$. Velger vi $F(x, y, z) = (x, z, -y)$, får vi av formlene (1.1) og (1.2), etter litt regning:

$$G(u, v) = \left(\frac{2u}{u^2 + (v+1)^2}, \frac{u^2 + v^2 - 1}{u^2 + (v+1)^2}\right),$$

eller, hvis vi skriver G med komplekse tall og setter $z = u + iv$:

$$G(z) = \frac{z + \bar{z} + i(z\bar{z} - 1)}{|z + i|^2} = \frac{(iz + 1)(\bar{z} - i)}{(z + i)(\bar{z} - i)} = \frac{iz + 1}{z + i}.$$

Dette er en (kompleks) lineær brøktransformasjon som restrikterer til en homeomorfi $G : \mathbb{H} \approx \mathbb{D}$. Vi ser at $G(0) = -i$, $G(1) = 1$ og $G(-1) = -1$, og dette bestemmer G entydig, etter korollar 2.1. (Vi kunne også ha brukt dette til å definere G.) G bevarer vinkler og sirkler i $\overline{\mathbb{C}}$, så den må avbilde sirkelen gjennom 0, 1 og -1, dvs. $\overline{\mathbb{R}}$, på sirkelen gjennom $-i$, 1 og -1, dvs. $\mathbb{S}^1 = \{z \in \mathbb{C} \,|\, |z| = 1\}$. Derfor vil $\mathbb{H}$–linjene avbildes på enten sirkelbuer som står ortogonalt på $\mathbb{S}^1$ eller diametere. Disse kurvene blir de hyperbolske linjene i disk–modellen, eller $\mathbb{D}$*–linjer.*

Siden G bevarer vinkler, vil denne modellen også bli *konform* — dvs. de hyperbolske vinklene er de samme som de Euklidske.

De reelle Möbiustransformasjonene *Möb*($\mathbb{H}$) svarer til en gruppe *Möb*($\mathbb{D}$) av transformasjoner av $\mathbb{D}$ som vil bevare vinkler og hyperbolske linjer, og ellers har samme egenskaper i forhold til $\mathbb{D}$ som *Möb*($\mathbb{H}$) til $\mathbb{H}$. *Möb*($\mathbb{D}$) er definert ved at $f \in$ *Möb*($\mathbb{D}$) $\iff$ $G^{-1} f G \in$ *Möb*($\mathbb{H}$). Med andre ord: enhver $f \in$ *Möb*($\mathbb{D}$) kan skrives som GgG^{-1}, for en $g \in$ *Möb*($\mathbb{H}$)

Når vi nå skal undersøke hvordan en slik f vil se ut, er det enklest å bruke matriserepresentasjon av brøktransformasjoner. G svarer da til matrisen $\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}$, og for G^{-1} vil vi bruke formelen $G^{-1}(z) = \dfrac{iz - 1}{-z + i}$, som svarer til matrisen $\begin{bmatrix} i & -1 \\ -1 & i \end{bmatrix}$. Hvis $g(z) = \dfrac{az + b}{cz + d}$, a, b, c, d reelle, får vi da

at GgG^{-1} svarer til

$$\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} i & -1 \\ -1 & i \end{bmatrix} = \begin{bmatrix} ia+c & ib+d \\ a+ic & b+id \end{bmatrix}\begin{bmatrix} i & -1 \\ -1 & i \end{bmatrix} =$$
$$\begin{bmatrix} -a+ci-ib-d & -ia-c-b+id \\ ia-c-b-id & -a-ic+ib-d \end{bmatrix} =$$
$$\begin{bmatrix} -a-d+(c-b)i & -b-c-(a-d)i \\ -b-c+(a-d)i & -a-d-(c-b)i \end{bmatrix}$$

Den siste matrisen kan skrives $\begin{bmatrix} \alpha & \beta \\ \bar\beta & \bar\alpha \end{bmatrix}$, med $\alpha = -a-d+(c-b)i$ og $\beta = -b-c-(a-d)i$. På den annen side er det lett å se at ethvert par av komplekse tall (α, β) kan skrives entydig slik. Det betyr at $M\ddot{o}b^+(\mathbb{H})$ gir opphav til komplekse lineære brøktransformasjoner på formen

$$g(z) = \frac{\alpha z + \beta}{\bar\beta z + \bar\alpha}\,. \tag{7.1}$$

En enkel utregning gir $\alpha\bar\alpha - \beta\bar\beta = 4(ad-bc)$, så det er naturlig her å normere α og β slik at $\alpha\bar\alpha - \beta\bar\beta = 1$. (Merk at $(k\alpha, k\beta)$ bare definerer samme funksjon som (α, β) hvis k er *reell*, så vi kan bare normere ved multiplikasjon med reelle tall. Enhver slik normering vil bevare tegnet til $\alpha\bar\alpha - \beta\bar\beta$.) Analogt med $M\ddot{o}b^+(\mathbb{H})$ vil vi bruke notasjonen $M\ddot{o}b^+(\mathbb{D})$ for denne undergruppen av $M\ddot{o}b(\mathbb{D})$

Det tilsvarende regnestykket for $M\ddot{o}b^-(\mathbb{H})$ blir litt annerledes på grunn av den komplekse konjugasjonen som inngår. Hvis $g(z) = \dfrac{a\bar z + b}{c\bar z + d}$, med a, b, c, d reelle, så er

$$GgG^{-1}(z) = G\left(\frac{a\overline{G^{-1}(z)} + b}{c\overline{G^{-1}(z)} + d}\right),$$

der $\overline{G^{-1}}$ nå svarer til den *konjugerte* av matrisen til G^{-1}, dvs. $\begin{bmatrix} -i & -1 \\ -1 & -i \end{bmatrix} = -\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}$. Denne matrisen definerer samme funksjon som $\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}$ — dvs. G. Derfor må vi regne ut

$$\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix} = \cdots$$
$$\cdots = \begin{bmatrix} -a+d+(c+b)i & c-b-(a+d)i \\ -(c-b)+(a+d)i & a-d+(c+b)i \end{bmatrix} = \begin{bmatrix} \alpha' & \beta' \\ -\bar\beta' & -\bar\alpha' \end{bmatrix}$$

med $\alpha' = -a+d+(c+b)i$ og $\beta' = c-b-(a+d)i$. Men $\begin{bmatrix} \alpha' & \beta' \\ -\bar\beta' & -\bar\alpha' \end{bmatrix}$ bestemmer samme funksjon som

$$\begin{bmatrix} i\alpha' & i\beta' \\ -i\bar\beta' & -i\bar\alpha' \end{bmatrix} = \begin{bmatrix} i\alpha' & i\beta' \\ \overline{i\beta'} & \overline{i\alpha'} \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \bar\beta & \bar\alpha \end{bmatrix}$$

Her er $\alpha\bar\alpha - \beta\bar\beta = \alpha'\bar\alpha' - \beta'\bar\beta' = -4(ad-bc)$, som er positiv hvis $g \in M\ddot{o}b^-(\mathbb{H})$, så igjen kan vi normere slik at $\alpha\bar\alpha - \beta\bar\beta = 1$

Altså svarer $M\ddot{o}b^-(\mathbb{H})$ til avbildninger på formen

$$f(z) = \frac{\alpha\bar{z} + \beta}{\bar{\beta}\bar{z} + \bar{\alpha}}\,, \tag{7.2}$$

med $\alpha\bar{\alpha} - \beta\bar{\beta} = 1$. Vi bruker notasjonen $M\ddot{o}b^-(\mathbb{D})$ for disse.

Eksempler. (1) Kompleks konjugasjon i $\mathbb{D}$. Her kan vi sette $\alpha = -1$, $\alpha' = i$ og $\beta = \beta' = 0$. Ligningene over gir $-a + d = 0,\ c + b = 1,\ a + d = 0$ og $c - d = 0$. Derfor er $a = d = 0$ og $c = d = 1/2$. Derfor svarer kompleks konjugasjon i $\mathbb{D}$ til $f(z) = \frac{1}{2}/\frac{1}{2}\bar{z} = 1/\bar{z} = z/|z|^2$ i $\mathbb{H}$, altså inversjon i sirkelen ($\mathbb{H}$–linjen) $|z| = 1$.

(2) På samme måte kan vi se at inversjon (refleksjon) i den imaginære aksen i $\mathbb{H}$ svarer til refleksjon i den imaginære aksen i $\mathbb{D}$ også.

(3) La oss nå bestemme hvilke elementer i $M\ddot{o}b(\mathbb{D})$ som har 0 som fikspunkt. For $M\ddot{o}b^+(\mathbb{D})$ får vi ligningen $\dfrac{\alpha\cdot 0 + \beta}{\bar{\beta}\cdot 0 + \alpha} = 0$ — dvs. $\beta = 0$. Betingelsen $\alpha\bar{\alpha} - \beta\bar{\beta} = 1$ gir $|\alpha| = 1$, så vi kan skrive $\alpha = \mathrm{e}^{i\theta/2}$ for en $\theta \in \mathbb{R}$. Alle elementer i $M\ddot{o}b^+(\mathbb{D})$ som har 0 som fikspunkt kan altså skrives $f(z) = e^{i\theta/2}z/e^{-i\theta/2} = e^{i\theta}z$. Men dette er formelen for en rotasjon med vinkelen θ, skrevet som kompleks multiplikasjon. Omvendt er enhver slik rotasjon et element i $M\ddot{o}b^+(\mathbb{D})$.

Det samme regnestykket gir at et element i $M\ddot{o}b^-(\mathbb{D})$ som har 0 som fikspunkt har formen $f(z) = e^{i\theta}\bar{z}$, som er en refleksjon om linjen som danner vinkelen $\theta/2$ med x-aksen. (Vi kan f. eks. skrive $e^{i\theta}z = e^{i\theta/2}\overline{e^{-i\theta/2}\bar{z}}$, som betyr at vi først roterer vinkelen $-\theta/2$, så reflekter vi i x-aksen, og tilslutt roterer vi tilbake med vinkelen $\theta/2$.) Disse avbildningene — rotasjoner og refleksjoner — utgjør gruppen av ortogonale avbildninger i dimensjon 2, som svarer til gruppen $O(2)$ av ortogonale 2×2–matriser.

Vi kan også overføre metrikken d fra $\mathbb{H}$ til $\mathbb{D}$. La oss skrive $d_{\mathbb{H}}$ for d. Da vil formelen

$$d_{\mathbb{D}}(z_1, z_2) = d_{\mathbb{H}}(G^{-1}(z_1), G^{-1}(z_2))$$

definere en metrikk på $\mathbb{D}$ slik at G og G^{-1} bli inverse *isometrier*. La oss utlede et uttykk for $d_{\mathbb{D}}$.

Vi husker at $d_{\mathbb{H}}(w_1, w_2) = |\ln(|g(w_2)|)|$, der g er en lineær brøktransformasjon som tar w_1 på i og w_2 også til et punkt på den imaginære aksen. En slik g er entydig bestemt av kravene $g(p) = 0,\ g(w_1) = i$ og $g(q) = \infty$, der p og q er endepunktene til den hyperbolske linjen l som går gjennom w_1 og w_2. Dermed blir $d_{\mathbb{D}}(z_1, z_2) = |\ln(|gG^{-1}(z_2)|)|$ der gG^{-1} er en brøktransformasjon som tar $G(p)$ på 0, z_1 på i og $G(q)$ på ∞. Men $G(p)$ og $G(q)$ er endepunktene til $G(l)$, som er den entydig bestemte $\mathbb{D}$–linjen som går gjennom z_1 og z_2. Altså har vi:

$$d_{\mathbb{D}}(z_1, z_2) = |\ln(|h(z_2)|)|\,,$$

der h er er karakterisert som en brøktransformasjon som tar z_1 på i og endepunktene til den hyperbolske linjen gjennom z_1 og z_2, på 0 og ∞.

Legg merke til at selv om dette tilsynelatende er den samme formelen som for $d_{\mathbb{H}}$, så er tolkningen litt forskjellig. z_1 og z_2 er nå punkter i $\mathbb{D}$, og h blir en annen funksjon, siden 'den hyperbolske linjen' som bestemmer den ikke betyr det samme i de to situasjonene. Men vi får den samme eksplisitte

formelen som for $d_{\mathbb{H}}$; bare litt enklere, siden ingen linjer i $\mathbb{D}$ har endepunkt i ∞. Derfor får vi her alltid

$$h(z) = \frac{z-p}{z-q}\,\frac{z_1-q}{z_1-p}\,i\,,$$

der p og q er de to endepunktene for $\mathbb{D}$–linjen gjennom z_1 og z_2. ($h(z)$ er entydig bestemt, men byttes ut med $-1/h(z)$ hvis vi bytter om p og q. Men $|\ln(|h(z)|)| = |\ln(|-1/h(z)|)|$, så dette endrer ikke formelen for $d_{\mathbb{H}}$.) Altså har vi:

$$d_{\mathbb{D}}(z_1,z_2) = |\ln(|\frac{z_2-p}{z_2-q}\,\frac{z_1-q}{z_1-p}|)|\,. \tag{7.3}$$

Vi skal nå bruke den ekstra symmetrien vi har i $\mathbb{D}$ til å studere metrikken litt nøyere, og spesielt utlede formler for $d_{\mathbb{D}}(z_1,z_2)$ som ikke involverer endepunktene for D–linjen gjennom z_1 og z_2. Først observerer vi at siden rotasjonene om origo er isometrier, så er avstanden fra 0 til et punkt z lik avstanden fra 0 til r, der $r = |z| \in [0,1) \in \mathbb{D}$. Men endepunktene for $\mathbb{D}$–linjen gjennom 0 og r er 1 og -1, så avstandsformelen gir

$$d_{\mathbb{D}}(0,z) = d_{\mathbb{D}}(0,r) = |\ln(|\frac{0-(-1)}{0-1}\,\frac{r-1}{r-(-1)}|)| = \ln(\frac{1+r}{1-r})\,.$$

Denne ligningen kan vi også løse med hensyn på r, og vi får

$$r = \tanh(\frac{d_{\mathbb{D}}(0,z)}{2})\,. \tag{7.4}$$

For å finne avstanden mellom to vilkårlige punkter z_1 og z_2, kan vi først flytte z_1 til 0 ved en isometri $f \in \mathit{Möb}^+(\mathbb{D})$, og så bruke formelen ovenfor for avstanden mellom 0 og $f(z_2)$.

Hvis $f(z) = \dfrac{az+b}{\bar{b}z+\bar{a}}$ tilfredsstiller $f(z_1) = 0$, så må $b = -az_1$, Derfor kan vi skrive

$$f(z) = \frac{a(z-z_1)}{-\overline{az_1}z+\bar{a}} = \frac{a}{\bar{a}}\left(\frac{z-z_1}{-z_1z+1}\right)\,.$$

Setter vi derfor $\rho = |f(z_2)| = \dfrac{|z_2-z_1|}{|1-\bar{z}_1z_2|}$, får vi

$$d_{\mathbb{D}}(z_1,z_2) = \ln(\frac{1+\rho}{1-\rho})\,,\quad \text{eller}\quad \rho = \tanh(\frac{d_{\mathbb{D}}(z_1,z_2)}{2})\,.$$

En annen formel får vi ved å ta utgangspunkt i den andre av disse formlene og bruke (se oppgave 7.6b)

$$\sinh^2(\frac{d_{\mathbb{D}}}{2}) = \frac{\tanh^2(d_{\mathbb{D}}/2)}{1-\tanh^2(d_{\mathbb{D}}/2)} = \frac{\rho^2}{1-\rho^2}\,.$$

Setter vi inn for ρ, får vi

$$\frac{\rho^2}{1-\rho^2} = \frac{\dfrac{|z_2-z_1|^2}{|1-\bar{z}_1z_2|^2}}{1-\dfrac{|z_2-z_1|^2}{|1-\bar{z}_1z_2|^2}} = \frac{|z_2-z_1|^2}{|1-\bar{z}_1z_2|^2-|z_2-z_1|^2}\,.$$

Nevneren her er

$$(1-\bar{z}_1z_2)(1-z_1\bar{z}_2)-(z_1-z_2)(\bar{z}_1-\bar{z}_2)=$$
$$1-\bar{z}_1z_2-z_1\bar{z}_2+|z_1|^2|z_2|^2-(|z_1|^2-z_1\bar{z}_2-\bar{z}_1z_2+|z_2|^2)=$$
$$1-|z_1|^2-|z_2|^2+|z_1|^2|z_2|^2=(1-|z_1|^2)(1-|z_2|^2)\,.$$

Derfor er altså

$$\sinh^2(\frac{d_{\mathbb{D}}}{2})=\frac{|z_2-z_1|^2}{(1-|z_1|^2)(1-|z_2|^2)}\,.$$

Nå bruker vi identiteten $\cosh(2w)=2\sinh^2(w)+1$ med $w=d_{\mathbb{D}}/2$ og får tilslutt

$$\cosh(d_{\mathbb{D}}(z_1,z_2))=1+\frac{2|z_2-z_1|^2}{(1-|z_1|^2)(1-|z_2|^2)}\,. \tag{7.5}$$

Dette er kanskje den vanligste formen avstandsformelen gis på. Vi kan nå gå tilbake til $\mathbb{H}$ med isometrien G, og få en tilsvarende formel for $d_{\mathbb{H}}$:

$$d_{\mathbb{H}}(w_1,w_2)=d_{\mathbb{D}}(G(w_1),G(w_2))=d_{\mathbb{D}}(\frac{iw_1+1}{w_1+i},\frac{iw_2+1}{w_2+i})\,.$$

Enkel utregning gir

$$\frac{iw_2+1}{w_2+i}-\frac{iw_1+1}{w_1+i}=\cdots=\frac{w_2-w_1}{(w_2+1)(w_1+1)}\,,\quad\text{og}$$

$$1-|\frac{iw+1}{w+i}|^2=\frac{(w+i)(\bar{w}-i)-(iw+1)(-i\bar{w}+1)}{|w+i|^2}=$$
$$\cdots=\frac{2\,i\bar{w}-2\,iw}{|w+i|^2}=\frac{4\,\mathrm{Im}\,w}{|w+i|^2}\,.$$

Setter vi disse uttrykkene inn i formelen for $d_{\mathbb{D}}$ ovenfor, får vi

$$\cosh(d_{\mathbb{H}}(w_1,w_2))=1+\frac{|w_2-w_1|^2}{2\,(\mathrm{Im}\,w_1)(\mathrm{Im}\,w_2)}\,. \tag{7.6}$$

Oppgaver.

7.1. (a) Vis at *hyperbolske sirkler*, dvs. mengder som kan skrives på formen $\{z\in\mathbb{D}\,|\,d_{\mathbb{D}}(z,z_0)=r\}$, der z_0 er et fiksert punkt og $r>0$, også er Euklidske sirkler.

(b) Samme spørsmål med $\mathbb{H}$ i stedet for $\mathbb{D}$

7.2. Vis at $d_{\mathbb{D}}$ definerer underromstopologien i $\mathbb{D}\subset\mathbb{C}$.

(Hint: Vis først (1) $d_{\mathbb{D}}:\mathbb{D}\times\mathbb{D}\to[0,\infty)$ er kontinuerlig i den Euklidske topologien på $\mathbb{D}$, og (2) $d_{\mathbb{D}}(z_1,z_2)\geq|z_1-z_2|$ for alle z_1 og z_2 i D.)

Hvorfor har vi nå også løst opgave 5.2?

7.3. Avstanden mellom to mengder U og V i et metrisk rom er definert som $d(U,V)=\inf\{d(u,v)\,|\,u\in U,\ v\in V\}$.

La l_1 og l_2 være to hyperbolske linjer med et felles endepunkt. Vis at $d(l_1,l_2)=0$.

Vis omvendt at hvis l_1 og l_2 er to linjer som ikke skjærer hverandre og $d(l_1, l_2) = 0$, så har de et felles endepunkt. (Du kan her med fordel bruke hint (2) i oppgave 7.2.)

7.4. Vis at hvis l er en hyperbolsk linje og z er et punkt som ikke ligger på l, så finnes en entydig linje l' som går gjennom z og som står ortogonalt på l.

7.5. Vis at hvis l_1 og l_2 er to linjer som ikke har noe felles endepunkt og som heller ikke skjærer hverandre, så har de en felles normal.

7.6. I dette kapitlet og i det neste bruker vi flere relasjoner mellom de hyperbolske funksjonene. Verifiser disse formlene:

(a) $\cosh^2 x - \sinh^2 x = 1$

(b) $\sinh^2 x = \dfrac{\tanh^2 x}{1 - \tanh^2 x}$

(c) $\sinh 2x = 2 \sinh x \cosh x$

(d) $\cosh 2x = \cosh^2 x + \sinh^2 x$

(e) $\dfrac{1 + \tanh^2 x}{1 - \tanh^2 x} = \cosh 2x$

(f) $\dfrac{2 \tanh x}{1 - \tanh^2 x} = \sinh 2x$

8. Trekanter og trigonometriske formler i det hyperbolske plan

I dette kapitlet skal vi bruke Poincarés diskmodell $\mathbb{D}$ for det hyperbolske plan, men de formlene vi er ute etter, vil være uavhengige av hvilken modell vi bruker. Derfor er de også gyldige i $\mathbb{H}$.

Med en *trekant* skal vi mene et trippel av punkter (hjørnene) som ikke ligger på en linje. Hvert par av hjørner utspenner segmenter, stråler og linjer. Segmentene kaller vi da *sidene* i trekanten, og *vinkelen* i et hjørne er det paret av stråler som har dette hjørnet felles. En vilkårlig trekant er da alltid kongruent med en som har ett hjørne i 0, et annet langs den positive reelle aksen, og der det tredje har positiv imaginærverdi: hvis u, z og w er hjørnene, kan vi flytte u til 0, rotere slik at (bildet av) z havner på den positive reelle aksen, og eventuelt konjugere for å oppnå $\operatorname{Im} w > 0$. La oss si at trekanten nå er i "standard posisjon". (Men merk at det for en gitt trekant er seks forskjellige måter å gjøre dette på.) Den nye trekanten vil ha like store vinkler og sider som den opprinnelige, så når vi skal studere relasjoner mellom vinkler og sider, kan vi like godt anta at trekanten er i standard posisjon.

La nå vinklene være α, β og γ, målt i radianer, og la de motstående sidene ha lengde a, b og c, målt i den hyperbolske metrikken. Vi ordner symbolene slik at vinkelen i 0 er α og den andre vinkelen langs den reelle aksen er β. Hjørnet på den reelle aksen identifiserer vi med et reelt tall $r \in (0, 1)$ og det tredje hjørnet kan vi skrive $w = se^{i\alpha}$, for en $s \in (0, 1)$. Da er, etter avstandsformelen (7.5):

$$\begin{aligned}\cosh a &= 1 + 2\frac{|r - se^{i\alpha}|^2}{(1-r^2)(1-s^2)} = 1 + 2\frac{r^2 + s^2 - 2rs\cos\alpha}{(1-r^2)(1-s^2)} \\ &= \frac{1 - r^2 - s^2 + r^2s^2 + 2r^2 + 2s^2 - 4rs\cos\alpha}{(1-r^2)(1-s^2)} \\ &= \frac{1+r^2}{1-r^2}\,\frac{1+s^2}{1-s^2} - \frac{2r}{1-r^2}\,\frac{2s}{1-s^2}\,\cos\alpha\end{aligned}$$

Men $r = \tanh(c/2)$ (7.4), og derfor

$$\frac{1+r^2}{1-r^2} = \frac{1+\tanh^2(c/2)}{1-\tanh^2(c/2)} = \cosh c\,, \text{og}$$
$$\frac{2r}{1-r^2} = \frac{2\tanh(c/2)}{1-\tanh^2(c/2)} = \sinh c\,.$$

Tilsvarende gjelder for s og b, og vi får

Proposisjon 8.1. *(Den første cosinussetningen.)*

$$\cosh a = \cosh b \cosh c - \sinh b \sinh c \cos\alpha$$

(Vi får naturligvis to slike relasjoner til ved å permutere hjørnene.)

Korollar 8.1. *(Pythagoras i det hyperbolske tilfellet.)*
Hvis $\alpha = \pi/2$, så er $\cosh a = \cosh b \cosh c$.

Setter vi inn rekkeutviklinger for cosh, får vi

$$1 + \frac{a^2}{2} + \frac{a^4}{4!} + \cdots = (1 + \frac{b^2}{2} + \frac{b^4}{4!} + \cdots)(1 + \frac{c^2}{2} + \frac{c^4}{4!} + \cdots)\,.$$

Multipliserer vi ut og løser med hensyn på a^2–leddet, får vi $a^2 = b^2 + c^2 +$ ledd av orden minst 4. Derfor blir dette tilnærmet den Euklidske Pythagoras' setning for små trekanter.

Vi setter nå $A = \cosh a$, $B = \cosh b$ og $C = \cosh c$. Da blir $\sinh a = \sqrt{A^2 - 1}$, etc., og den første cosinussetningen kan skrives

$$\sqrt{B^2-1}\sqrt{C^2-1}\cos\alpha = BC - A$$

eller, hvis vi kvadrerer:

$$(B^2-1)(C^2-1)(1 - \sin^2\alpha) = (BC - A)^2\,.$$

Denne ligningen kan vi løse med hensyn på $\sin^2\alpha$:

$$\sin^2\alpha = \frac{(B^2-1)(C^2-1) - (BC-A)^2}{(B^2-1)(C^2-1)} = \cdots = \frac{2\,ABC - A^2 - B^2 - C^2 + 1}{(B^2-1)(C^2-1)}\,.$$

Det følger at kvosienten

$$\frac{\sin^2\alpha}{\sinh^2 a} = \frac{\sin^2\alpha}{A^2-1} = \frac{2\,ABC - A^2 - B^2 - C^2 + 1}{(A^2-1)(B^2-1)(C^2-1)}\,,$$

og dette uttrykket er helt symmetrisk i A, B, C. Derfor får vi det samme om vi bytter ut (α, a) med (β, b) eller (γ, c), og vi har:

$$\frac{\sin^2\alpha}{\sinh^2 a} = \frac{\sin^2\beta}{\sinh^2 b} = \frac{\sin^2\gamma}{\sinh^2 c}\,.$$

Siden både sin og sinh er positive i det aktuelle området, har vi bevist

Proposisjon 8.2. *(Den hyperbolske sinussetningen.)*

$$\frac{\sin\alpha}{\sinh a} = \frac{\sin\beta}{\sinh b} = \frac{\sin\gamma}{\sinh c}$$

For små x er $\sinh x \approx x$, så igjen ser vi at for små trekanter har vi tilnærmet den Euklidske sinussetningen.

For hyperbolske trekanter finnes også en annen type cosinussetning:

Proposisjon 8.3. *(Den andre cosinussetningen.)*

$$\cos\alpha = -\cos\beta\cos\gamma + \sin\beta\sin\gamma\cosh a$$

(Her får vi naturligvis også to formler til, ved å permutere $\alpha,\ \beta$ og γ.)

Bevis. Etter den første cosinussetningen kan vi skrive

$$\cos\alpha = \frac{BC - A}{\sqrt{(B^2-1)(C^2-1)}}\,,\ \text{etc.}$$

Setter vi inn disse tre uttrykkene, får vi

$$\cos\alpha + \cos\beta\cos\gamma =$$
$$\frac{BC - A}{\sqrt{(B^2-1)(C^2-1)}} + \frac{(AC-B)(AB-C)}{(A^2-1)\sqrt{(B^2-1)(C^2-1)}} =$$
$$\cdots = A\frac{1 - A^2 - B^2 - C^2 + 2\,ABC}{(A^2-1)\sqrt{(B^2-1)(C^2-1)}}$$

Bruker vi nå uttrykket for $\sin^2$ fra diskusjonen som ledet til sinussetningen ovenfor, får vi

$$\sin\beta\sin\gamma = \frac{1 - A^2 - B^2 - C^2 + 2\,ABC}{(A^2-1)\sqrt{(B^2-1)(C^2-1)}}\,.$$

Derfor har vi

$$\cos\alpha + \cos\beta\cos\gamma = A\sin\beta\sin\gamma = \sin\beta\sin\gamma\cosh a\,.$$

□

Fra den første cosinussetningen følger de klassiske kongruenssetningene. F. eks. bestemmer lengdene av de tre sidene også de tre vinklene, og dermed hele trekanten opp til kongruens (den såkalte SSS-setningen — for "side–side–side"). Dette er som i Euklidsk geometri.

Den andre cosinussetningen, derimot, har ingen analogi i Euklidsk geometri. Den sier at gitt vinklene i en trekant, så er også *sidene* fullstendig bestemt (og derfor også hele trekanten, opp til kongruens). Som i sfærisk geometri er altså også her likedannede trekanter kongruente!

I det hyperbolske plan kan vi også snakke om trekanter som har et eller flere hjørner i *det uendelig fjerne*. Vi vil kalle slike trekanter *enkelt, dobbelt* eller *trippelt asymptotiske,* avhengig av om de har ett, to eller tre slike hjørner. Det er naturlig å tenke på slike trekanter som grenser for endelige trekanter der et eller flere av hjørnene går mot uendelig.

La oss se litt nærmere på de tre typene asymptotiske trekanter:

(i) Trippelt asymptotisk. Alle sidene blir her hele hyperbolske linjer. Vi velger å betrakte trekantene i $\mathbb{H}$. Da er de bestemt av tre punkter i $\overline{\mathbb{R}}$. Hvis vi har to slike trekanter bestemt av punktene p_1, q_1, r_1 og p_2, q_2, r_2, så finnes etter korollar 2.1 en lineær brøktransfomasjon f som avbilder det ene triplet av punkter på det andre, og siden punktene er i $\overline{\mathbb{R}}$, kan f kan skrives med reelle koeffisienter. Det er også lett å se av formlene i beviset at om vi bytter om rekkefølgen på to av punktene, så skifter determinanten ($'ad - bc'$) tegn. Derfor kan vi anta at $f \in M\ddot{o}b^+(\mathbb{H})$. Men da må hele trekantene avbildes på hverandre, så de er kongruente.

Altså har vi: *Alle trippelt asymptotiske trekanter er kongruente.*

(ii) Dobbelt asymptotisk. To av sidene blir stråler og den siste blir en linje. Denne gangen betrakter vi trekanten som liggende i $\mathbb{D}$. Hvis vi flytter det endelige hjørnet til 0 ved en Möbiustransformasjon, så blir de to sidene som har dette hjørnet felles, radier i modellen. Vi ser da at vinkelen i 0 vil bestemme trekanten opp til kongruens.

(iii) Enkelt asymtotisk. To av sidene blir stråler og den siste et endelig segment. Da får vi to endelige hjørner med vinkler α og β, og en endelig, mellomliggende side c. Den tredje vinkelen blir $\gamma = 0$.

Den andre cosinussetningen vil fremdeles gjelde (gå til grensen igjen), og den gir:

$$1 = \cos\gamma = -\cos\alpha\cos\beta + \sin\alpha\sin\beta\cosh c\,.$$

Denne ligningen bestemmer den tredje av parameterne α, β og c, gitt de to andre. F. eks er

$$\cosh c = \frac{1 + \cos\alpha\cos\beta}{\sin\alpha\sin\beta}\,.$$

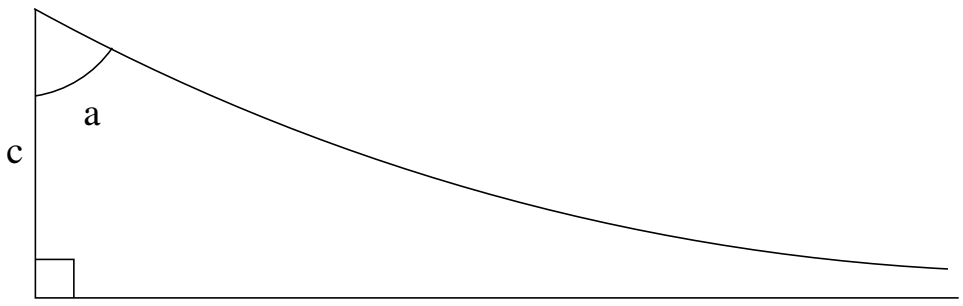
(Se også oppgave 8.3.) Et viktig tilfelle er når den ene vinkelen, f. eks. β, er $\pi/2$. (Se figur 5.) Da blir

$$\cosh c = \frac{1}{\sin\alpha}$$

Det er vanlig å gi denne relasjonen en annen form: løs ligningen med hensyn på e^{-c} og få

$$\mathrm{e}^{-c} = \frac{1 - \cos\alpha}{\sin\alpha} = \tan\frac{\alpha}{2}$$

Disse relasjonene viser hvordan avstand er bestemt av vinkler (og omvendt). Spesielt definerer denne ligningen en entydig lengdeenhet i det hyperbolske planet, når vinkelmålet er gitt!



FIGUR 5

Oppgaver.

8.1. Vis at i en trekant er to vinkler like store hvis og bare hvis de motstående sidene er like lange.

8.2. Vis at i en likesidet trekant er alle vinklene like store. Vis også at hvis sidene har lengde a og vinklene er α, så er $2\cosh(a/2)\sin(\alpha/2) = 1$. (Hint: del trekanten i to.)

8.3. Vis at i en enkelt asymptotisk trekant med endelige vinkler α og β og endelig side c, så er β bestemt av α og c.

9. Buelengde og areal

Anta at en kurve $\mathcal{C}$ i et metrisk rom er gitt ved en parameterfremstillig $z(t),\ t \in [a,b]$, og anta for enkelthets skyld at z er injektiv. *Buelengden* til $\mathcal{C}$ er da definert som

$$\sup_{a=t_0<t_1<\cdots<t_n=b} \sum_i d(z(t_{i+1})z(t_i))\,, \tag{9.1}$$

hvis dette tallet eksisterer. (Vi tar sup over alle partisjoner $a = t_0 < t_1 < \cdots < t_n = b$ av intervallet $[a,b]$.)

Anta nå at buelengden til restriksjonen av z til $[a,t]$ eksisterer for alle $t \in [a,b]$, og kall den $s(t)$. s blir da en voksende funksjon fra $[a,b]$ til $[0,\infty)$. Vi skal ikke her gå inn på den generelle teorien, men man kan vise at hvis s er deriverbar, så er

$$s'(t) = \lim_{h\to 0} \frac{s(t+h)-s(t)}{h} = \lim_{h\to 0+} \frac{d(z(t+h),z(t))}{h}\,. \tag{9.2}$$

Omvendt, hvis denne grensen eksisterer og er kontinuerlig, kan vi finne $s(t)$ ved integrasjon.

En viktig observasjon er at hvis $g : X \to Y$ er en isometri, der X og Y gjerne kan være to forskjellig metriske rom, og $z(t),\ t \in [a,b]$ er en kurve i X, så vil $gz(t)$ være en kurve i Y, og hvis z har buelengde s, så har også gz buelengde s. Dette ser vi direkte av definisjonen (9.1), siden vi alltid vil ha $d(gz(t_{i+1}), gz(t_i)) = d(z(t_{i+1})z(t_i))$.

Eksempel 9.1. Hvis $z(t)$ parametriserer et segment av en hyperbolsk linje, betyr betingelsen (d4) til metrikken i kapittel 5 at buelengden er lik den hyperbolske avstanden mellom endepunktene.

Eksempel 9.2. I $\mathbb{R}^2$ har vi formelen

$$s'(t) = \sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}\,, \tag{9.3}$$

der $z(t) = (x(t), y(t))$, forutsatt at $z(t)$ er deriverbar. Vi kan også skrive dette som $(\frac{ds}{dt})^2 = (\frac{dx}{dt})^2 + (\frac{dy}{dt})^2$. Denne formelen skal gjelde for enhver parametrisering, og dette uttrykker vi med relasjonen

$$ds^2 = dx^2 + dy^2\,. \tag{9.4}$$

(Denne relasjonen kan gis en helt presis tolkning som en ligning i et passende valgt vektorrom, men her behøver vi ikke tenke på den som noe annet enn en invariant notasjon for (9.3)).

Vi skal nå finne tilsvarende uttrykk for buelengden i de to modellene $\mathbb{H}$ og $\mathbb{D}$ for det hyperbolske plan.

Vi begynner med $\mathbb{H}$ — dvs. vi setter $d = d_{\mathbb{H}}$. Avstandsformelen (7.6) gir:

$$\cosh(d_{\mathbb{H}}(z(t+h), z(t))) = 1 + \frac{|z(t+h) - z(t)|^2}{2\,(\operatorname{Im} z(t+h))(\operatorname{Im} z(t))}\,. \tag{9.5}$$

For å forenkle notasjonen, skriver vi $d(h) = d_{\mathbb{H}}(z(t+h), z(t))$. Taylors formel for cosh gir at vi kan skrive $\cosh w = 1 + \frac{w^2}{2} + \eta(w)w^2$, der $\lim_{w\to 0} \eta(w) = 0$. (9.5) gir da

$$1 + \frac{(d(h))^2}{2} + \eta(d(h))(d(h))^2 = 1 + \frac{|z(t+h) - z(t)|^2}{2\,(\operatorname{Im} z(t+h))(\operatorname{Im} z(t))}\,,$$

og derfor

$$(\frac{d(h)}{h})^2(1 + 2\,\eta(d(h))) = \left|\frac{z(t+h) - z(t)}{h}\right|^2 \frac{1}{(\operatorname{Im} z(t+h))(\operatorname{Im} z(t))}\,. \tag{9.6}$$

Vi ser at hvis $z(t)$ er deriverbar, så vil $\lim_{h\to 0}\left(\dfrac{d(z(t+h), z(t))}{h}\right)^2$ eksistere og være lik $\left((\frac{dx}{dt})^2 + (\frac{dy}{dt})^2\right)/(\operatorname{Im} z(t))^2$. Siden de involverte uttrykkene blir positive hvis $h > 0$, får vi derfor også

$$s'(t) = \lim_{h\to 0+} \frac{d(z(t+h), z(t))}{h} = \frac{\sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}}{\operatorname{Im} z(t)}$$

Analogt med eksemplet $\mathbb{R}^2$ ovenfor kan vi også skrive (husk at $y = \operatorname{Im} z$)

$$ds^2 = \frac{dx^2 + dy^2}{y^2}\,. \tag{9.7}$$

I Poincarédisken kan vi gjøre en tilsvarende analyse. Den eneste forskjellen blir at faktoren $\dfrac{1}{(\operatorname{Im} z(t+h))(\operatorname{Im} z(t))}$ på høyre side i formelen (9.6) erstattes med $\dfrac{4}{(1 - |z(t+h)|^2)(1 - |z(t)|^2)}$. Vi ender da opp med

$$ds^2 = 4\frac{dx^2 + dy^2}{(1 - x^2 - y^2)^2}\,. \tag{9.8}$$

Disse formlene kan vi, som med (9.4), tenke på som en måte å fortelle hvordan vi skal finne ds/dt, gitt en parameterfremstilling $z(t) = (x(t), y(t))$ for kurven. F. eks. sier (9.8) at i $\mathbb{D}$ får vi

$$(\frac{ds}{dt})^2 = 4\frac{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}{(1 - x(t)^2 - y(t)^2)^2}\,.$$

Eksempel 9.3. La $\mathcal{C}$ være en hyperbolsk sirkel med hyperbolsk radius ρ. Siden buelengde bevares under isometrier, kan vi anta at $\mathcal{C} \subset \mathbb{D}$ og at den har sentrum i 0. $\mathcal{C}$ er da også en Euklidsk sirkel med sentrum i 0, og med

Euklidsk radius gitt ved formelen (7.4) — dvs. $r = \tanh(\rho/2)$. Som parameterfremstilling av $\mathcal{C}$ kan vi da velge $z(t) = re^{it} = (r\cos t, r\sin t),\ t \in [0, 2\pi]$, og vi får $dx/dt = -r\sin t,\ dy/dt = r\cos t$ og $x^2 + y^2 = r^2$. Det gir

$$(\frac{ds}{dt})^2 = 4\frac{(-r\sin t)^2 + (r\cos t)^2}{(1-r^2)^2} = \frac{4r^2}{(1-r^2)^2}\,.$$

Buelengden av $\mathcal{C}$ blir derfor

$$s(\mathcal{C}) = \int_0^{2\pi} \frac{2r}{1-r^2}\,dt = \frac{4\pi r}{1-r^2} = \pi\frac{4\,\tanh(\frac{\rho}{2})}{1-\tanh^2(\frac{\rho}{2})} = 2\pi\sinh(\rho)\,.$$

Siden $\sinh\rho > \rho$ og $\sinh'\rho > 1$ for $\rho > 0$, ser vi at omkretsen av en sirkel er større og vokser fortere som funksjon av radien i det hyperbolske planet enn i det Euklidske planet. Mer eksplisitt, så ser vi at for små ρ er omkretsen tilnærmet lik $2\pi\rho$, men når ρ er stor, vokser omkretsen som πe^ρ.

Vi skal nå gå over til *areal*, og spesielt finne en formel for arealet av en trekant. Igjen skal vi avholde oss fra å gå inn i noen videre diskusjon av selve arealbegrepet, men begrense oss til å observere at en "fornuftig" arealfunksjon A i alle fall må være definert på kompakte (lukkede og begrensede) mengder hvor randen består av endelig mange glatte kurvestykker, at kongruente mengder har samme areal, og at A må være additiv, i den forstand at $A(U \cup V) + A(A \cap V) = A(U) + A(V)$. Dessuten bør selvfølgelig arealet av glatte kurvestykker og punkter være 0.

Det er ikke vanskelig å se at en slik funksjon vil være entydig bestemt opp til en konstant faktor (som tilfellet var med avstandbegrepet), og siden formlene vi skal komme frem til definerer slike funksjoner, kunne vi nøye oss med å skrive dem opp og verifisere at kravene er oppfylt. Men forhåpentligvis vil det følgende bidra til å forklare den geometriske bakgrunnen for formlene, og hvorfor de blir normert som de er (valg av den konstante faktoren).

Betrakt først $\mathbb{H}$. La $\Omega \subset \mathbb{H}$ være en mengde som beskrevet ovenfor, og tenk på identitetsavbildningen som en parametrisering. På vanlig måte kan vi da dekke *parameterområdet* Ω med et rutenett med hjørner (x_i, y_j), der $\{x_i\}_i$ og $\{y_j\}_j$ er voksende følger i $\mathbb{R}$. La $R(i,j)$ være ruten $[x_i, x_{i+1}] \times [y_i, y_{i+1}]$. Da vil $\sum_{R(i,j)\cap\Omega\neq\emptyset} A(R(i,j))$ approksimere $A(\Omega)$, og bedre jo finere rutenettet er. Derfor bør arealet være gitt ved et integral på formen

$$A(\Omega) = \iint_\Omega K(x,y)\,dx\,dy\,,$$

der

$$K(x,y) = \lim_{\Delta x,\Delta y\to 0} \frac{A(R(\Delta x,\Delta y))}{\Delta x\Delta y}\,,$$

hvor $R(\Delta x, \Delta y)$ er rektanglet $[x, x+\Delta x] \times [y, y+\Delta y]$. De to kantene $[x, x+\Delta x] \times \{y\}$ og $\{x\} \times [y, y+\Delta y]$ er kurver som møtes ortogonalt i $\mathbb{H}$, så $A(R(\Delta x, \Delta y))$ bør kunne approksimeres med produktet av lengdene av disse sidene. En naturlig normalisering er da at

$$\lim_{\Delta x,\Delta y\to 0} \frac{A(R(\Delta x,\Delta y))}{d_{\mathbb{H}}(x, x+\Delta x)\,d_{\mathbb{H}}(y, y+\Delta y)} = 1\,.$$

Men av (9.7) får vi $\lim_{\Delta x\to 0}\dfrac{d_{\mathbb{H}}(x,x+\Delta x)}{|\Delta x|}=\lim_{\Delta y\to 0}\dfrac{d_{\mathbb{H}}(y,y+\Delta y)}{|\Delta y|}=\dfrac{1}{y}$. Setter vi alt dette sammen, får vi $K(x,y)=\dfrac{1}{y^2}$, og dermed

$$A(\Omega)=\iint_\Omega \frac{dx\,dy}{y^2} \tag{9.9}$$

Dette skal være vår definisjon av arealfunksjonen i $\mathbb{H}$, og arealet er dermed definert for alle mengder hvor integralet er definert.

Denne diskusjonen kan også brukes til å vise at arealet er invariant under kongruens — dvs. $A(g\Omega)=A(\Omega)$ hvis g er en Möbiustransformasjon, men det kan være instruktivt å se hvordan dette også kan verifiseres ut fra formelen.

La da $g(z)=\dfrac{az+b}{cz+d}$, med $a,b,c,d\in\mathbb{R}$ og $ad-bc=1$, og anta $\Omega'=g\Omega$. For å skille mellom Ω og Ω' bruker vi notasjonen $z=x+iy$ for punktene i Ω og $w=u+iv$ for punktene i Ω'. Formelen for skifte av variable i dobbeltintegraler gir

$$A(\Omega')=\iint_{\Omega'}\frac{du\,dv}{v^2}=\iint_\Omega |J(g)(z)|\,\frac{dx\,dy}{(\operatorname{Im} g(z))^2} \tag{9.10}$$

der $|J(g)|$ er Jacobideterminanten til g, betraktet som en funksjon mellom delmengder av $\mathbb{R}^2$. Men det følger av Cauchy–Riemanns ligninger at for en kompleks analytisk avbildning g er $|J(g)(z)|=|g'(z)|^2$. (Oppgave 9.1.) Når $g(z)=\dfrac{az+b}{cz+d}$, er $g'(z)=(ad-bc)/(cz+d)^2$, så i vårt tilfelle får vi $|J(g)(z)|=1/|cz+d|^4$.

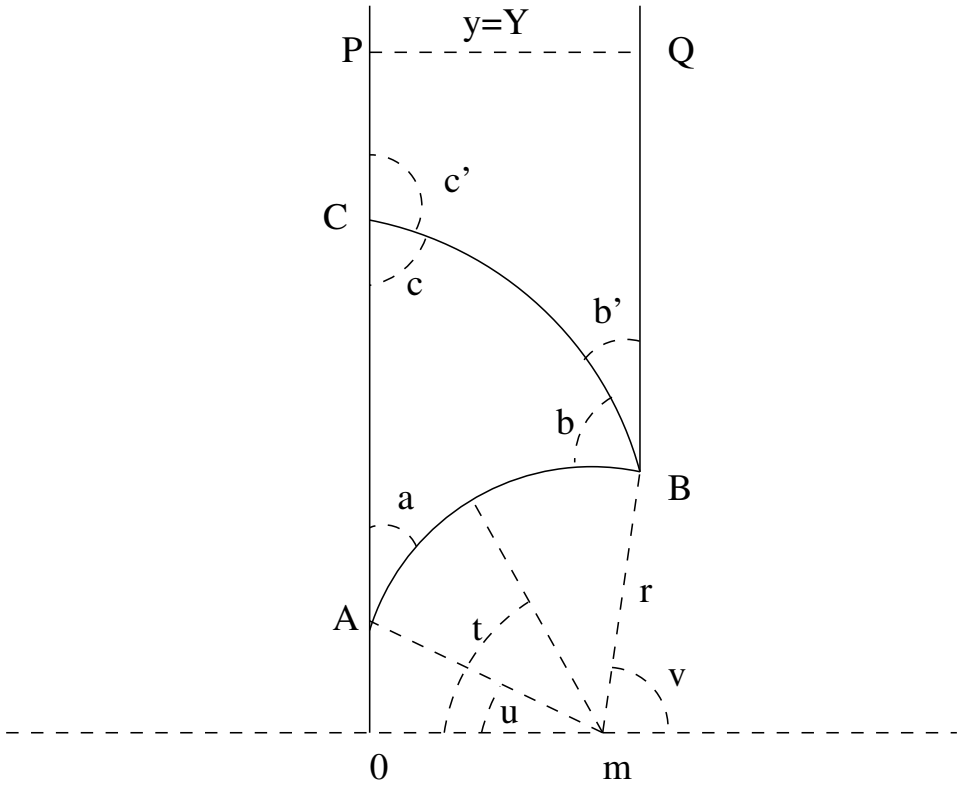
Formelen (2.1) gir at $\operatorname{Im}(g(z))=\operatorname{Im} z/|cz+d|^2=y/|cz+d|^2$. Setter vi alt dette inn i (9.10), får vi

$$A(\Omega')=\iint_\Omega \frac{1}{|cz+d|^4}\,\frac{dx\,dy}{\left(\dfrac{y}{|cz+d|^2}\right)^2}=\iint_\Omega\frac{dx\,dy}{y^2}=A(\Omega)$$

For å vise at alle Möbiustransformasjoner bevarer areal, er det nå nok å observere at speiling i den imaginære aksen, $g(z)=-\bar z$, også gjør det, siden $M\ddot{o}b(\mathbb{H})$ er generert av g og $M\ddot{o}b^+(\mathbb{H})$. Men g bevarer y og har Jacobideterminant -1, så (9.10) gir igjen $A(\Omega')=A(\Omega)$.

Eksempel 9.4. Arealet av trekanter. Enhver trekant i $\mathbb{H}$ er kongruent med en som har én side langs den imaginære aksen og et hjørne med positiv realdel. På figur 6 er tegnet en slik trekant, med hjørnene A, B og C. Her er m og r sentrum og radius i den sirkelen (hyperbolske linjen) som A og B utspenner.

Vi skal først beregne arealet til figuren $ABQP$ begrenset av det horisontale Euklidske segmentet $y=Y$ og de hyperbolske segmentene $[A,P]$, $[A,B]$ og $[B,Q]$. Av figuren ser vi da at $x=m-r\cos t,\ t\in[u,\pi-v]$. Da er



FIGUR 6

$dx = r\sin t\,dt$, og for hver t varierer y fra $r\sin t$ til Y. Vi får

$$\begin{aligned}\iint_{ABQP}\frac{dx\,dy}{y^2} &= \int_u^{\pi-v}\left[\int_{r\sin t}^{Y}\frac{dy}{y^2}\right] r\sin t\,dt\\ &= \int_u^{\pi-v}\left[\frac{1}{r\sin t}-\frac{1}{Y}\right] r\sin t\,dt\\ &= \pi-u-v-\frac{m+r\cos v}{Y}\end{aligned}$$

Det første vi merker oss nå, er at når Y går mot ∞, går dette uttrykket mot $\pi-u-v$. Det betyr at den asymptotiske trekanten med hjørnene A, B og ∞ har endelig areal, lik $\pi-(u+v)$. Observer nå at vinkelbena til u står ortogonalt på $\mathbb{H}$–linjene (AP) og (AB), så vi har $u=a$. På samme måte har vi $v=b+b'$, så $u+v$ er lik summen av vinklene i trekanten, siden den tredje vinkelen er 0. Ved å la et eller begge hjørnene A og B gå mot den reelle aksen, ser vi at denne formelen fortsatt gjelder om vi har en dobbelt eller trippelt asymptotisk trekant.

Den endelige trekanten ABC må ha areal lik differensen mellom arealene til to slike asymptotiske trekanter, den ene med areal $\pi-(a+b+b')$ og den andre med areal $\pi-(b'+c')$ — altså $\pi-a-b-b'-\pi+b'+c'=\pi-a-b-c$, siden $c+c'=\pi$. (Se figuren.) Vi formulerer dette som

Proposisjon 9.1. *Arealet av en trekant med vinkler a, b og c er lik*

$$\pi - (a + b + c)\,.$$

Denne formelen gjelder også for asymptotiske trekanter, dvs. hvis en eller flere av vinklene er 0.

Dette er et slående og særdeles viktig resultat. Det sier at arealet bare er avhengig av summen av vinklene i trekanten, og siden arealet alltid er positivt, får vi at vinkelsummen alltid er mindre enn π. Vi får også at arealet av alle trekanter er mindre enn π, med maksimum bare oppnådd av de trippelt asymptotiske trekantene (som alle er kongruente).

[Det er også interessant å sammenligne denne arealformelen med arealformelen i *sfærisk* geometri. Der var arealet av en trekant med vinklene a, b og c lik $(a + b + c - \pi)R^2$, der R er kulens radius. Formelt kan det se ut som hyperbolsk geometri svarer til "sfærisk geometri på en kule med radius i". Dette er en analogi som bli enda mer slående når vi sammenligner sinus- og cosinussetningene i de to geometriene, og bruker relasjonene $\sinh(ia) = i \sin a$ og $\cosh(ia) = \cos a$.]

Vi kan naturligvis også betrakte arealfunksjonen i diskmodellen $\mathbb{D}$, der en tilsvarende analyse vil gi formelen

$$A(\Omega) = \iint_\Omega \frac{4\,dx\,dy}{(1 - x^2 - y^2)^2}\,. \tag{9.11}$$

På grunn av symmetrien i $\mathbb{D}$ er det ofte naturlig å bruke polarkoordinater $x = r\cos\theta,\ y = r\sin\theta$ her, og vi får

$$A(\Omega) = \iint_T \frac{4r\,dr\,d\theta}{(1 - r^2)^2}\,, \tag{9.12}$$

der T er det aktuelle parameterområdet i (r, θ)–planet.

Eksempel 9.5. La oss regne ut arealet av en hyperbolsk sirkelskive $\mathcal{D}$ med hyperbolsk radius ρ. Som i eksempel 9.3 kan vi anta at sirkelen er en Euklidsk sirkel med sentrum i 0. Den Euklidske radien er da $R = \tanh(\rho/2)$ (7.4), og vi kan parametrisere $\mathcal{D}$ med polarkoordinater $x = r\cos\theta,\ y = r\sin\theta$, der $r \in [0, R]$ og $\theta \in [0, 2\,\pi]$. Etter formel (9.12) har vi da

$$\begin{aligned} A(\mathcal{D}) = \int_0^R \left[\int_0^{2\pi} \frac{4r\,d\theta}{(1 - r^2)^2}\right] dr &= 2\pi \int_0^R \frac{4r\,dr}{(1 - r^2)^2} \\ &= 2\pi \left[\frac{2}{(1 - r^2)}\right]_0^R = 4\pi \frac{R^2}{1 - R^2} \end{aligned}$$

Men $R = \tanh(\rho/2)$ og formelen i oppgave 7.6 b gir

$$A(\mathcal{D}) = 4\pi \frac{R^2}{1 - R^2} = 4\pi \sinh^2(\frac{\rho}{2})\,.$$

Siden $\sinh x > x$ og $\sinh' x > 1$ betyr dette at arealet av en sirkelskive er større og vokser raskere som funksjon av radien i det hyperbolske planet enn i det Euklidske. (Det samme viste vi om om omkretsen i eksempel 9.3). I differensialgeometrien betyr dette at det hyperbolske plan har *negativ krumning*.

[Det er også her interessant å sammenligne med geometri på en sfære av radius 1. Der er omkretsen av en sirkel med (sfærisk) radius ρ lik $2\pi \sin \rho$ og arealet $2\pi \sin^2(\rho/2)$, som begge er mindre og vokser *langsommere* enn i det Euklidske tilfellet. Sfæren har da *positiv* krumning, mens det Euklidske planet har krumning lik 0.]

Oppgaver. 9.1. For å vise invarians av areal under Möbiustransformasjoner brukte vi at $|J(g)(z)| = |g'(z)|^2$ for en kompleks analytisk funksjon g. Verifiser denne påstanden.

9.2. Finn et uttrykk for arealet til en hyperbolsk n–kant.

9.3. Vis at for små ρ er arealet av en hyperbolsk sirkel med (hyperbolsk) radius ρ tilnærmet lik $\pi\rho^2$.

Vis også at forholdet mellom omkretsen og arealet går mot 1 når radien vokser mot ∞.

[Et tips: Her kan du eksperimentere med Poincareś diskmodell:

http://www.math.umn.edu/ garrett/a02/H2.html

http://www.geom.uiuc.edu/java/triangle-area/]