Noname manuscript No.
(will be inserted by the editor)

Formalizing Complex Plane Geometry

Filip Marić · Danijela Petrović

Received: date / Accepted: date

Abstract Deep connections between complex numbers and geometry had been well known and carefully studied centuries ago. Fundamental objects that are investigated are the complex plane (usually extended by a single infinite point), its objects (points, lines and circles), and groups of transformations that act on them (e.g., inversions and Möbius transformations). In this paper, we treat the geometry of complex numbers formally and present a fully mechanically verified development within the theorem prover Isabelle/HOL. Apart from applications in formalizing mathematics and in education, this work serves as a ground for formally investigating various non-Euclidean geometries and their intimate connections. We discuss different approaches to formalization and discuss the major advantages of the more algebraically oriented approach.

Keywords Interactive theorem proving · Complex plane geometry · Möbius transformations

1 Introduction

Connections between complex numbers and geometry are deep and intimate. Although complex numbers have been recognized for more than 450 years, their geometric interpretation came only at the end of 18th century in works of Wessel, Argand and Gauss [26]. Their most significant applications in geometry were developed by Cauchy, Riemann, Möbius, Beltrami, Poincaré and others during the 19th-century [26]. Complex numbers present a very suitable apparatus for investigating properties of objects in very different geometries. Geometry has been studied analytically since Descartes, and the Cartesian plane ($\mathbb{R}^2$) is often used as

This work is partially supported by the Serbian Ministry of Education and Science grant ON174021, and Serbian-French Technology Co-Operation grant EGIDE/,,Pavle Savić" 680-00-132/2012-09/12 ("Formalization and automation of geometry").

Faculty of Mathematics
University of Belgrade
Studentski Trg 16
1100 Belgrade, Serbia

a domain for models of geometry (especially in the Euclidean case). However, replacing Cartesian by the complex plane gives simpler and more compact formulas that describe geometric objects, easing the calculations and shedding some new light on the subject. Therefore, the complex plane or some of its parts (e.g., the unit disc or the upper half plane) are often taken as the domain in which models of various geometries (both Euclidean and non-Euclidean ones) are formalized. It is also an important domain for investigations in modern physics (see, for example, Penrose and Rindler [28]). Due to its importance, the geometry of complex numbers has been well described in the literature. There are many textbooks describing the subject in great detail (during our work we have intensively used the textbooks written by Needham [26] and Schwerdtfeger [30]). Also, there is a plethora of course material (handouts, notes, slides) available online. However, we are not aware of any existing formalization of this subject. In this paper we present our fully formal, mechanically-verified exposition of the complex plane geometry which is, up to the best of our knowledge, first of this kind.

The need for rigorous justifications of arguments in geometry have been recognized for more than two millennia — Euclid's ,,Elements" are one of the first cases of mathematical deduction and form one of the most beautiful and influential works of science in the history of humankind. In the last century, the work of Hilbert [13] and Tarski [29] enriched us with much more precise developments of synthetic geometry. In the last several decades, with the advent of theorem provers and interactive proof-assistants, the level of formality and rigor in geometrical reasoning has been raised to the highest level. Within the formal theorem proving community, it is often advocated that, apart from the pure ,,L'art pour l'art" view on formalizing classical mathematical results, there are many practical benefits of this task (e.g., in mathematical education). We hope that more mathematicians will adopt this standpoint. The level of rigor has been constantly rising throughout the history of mathematics, and we feel that mechanical theorem proving helps reaching the ultimate ideal of fully rigorous proofs. Formal, mechanically-checked analysis of the content usually fills many gaps often present in classical textbooks and makes the authors think much deeper about the subject that is investigated. As it is often the case in formalization of mathematics, our experience in this work shows that there are not many wrong statements in the informal textbooks. Still, in textbooks that we have analyzed we have found some non-trivial statements that were erroneous and could not be proved. Even more abundant are the proofs that are imprecise, contain uncovered cases and miss some highly non-trivial justifications.

The final product of our present work is a well-developed theory of the extended complex plane (given both as a complex projective space and as the Riemann sphere), its objects (circles and lines), and its transformations (Möbius transformations). It can serve as a very important building block for further formal investigations of models of various geometries (e.g., our motivation for starting this work was to formalize the properties of Poincaré's disc model of hyperbolic geometry). Most of the concepts that we have formalized have already been described in the literature (although there are many details we had to invent since they were not described in the literature that we have consulted). However, our work required compiling many different sources into a uniform formal presentation and translating everything into a unique language since it was originally described in many different ways. For example, even within the same textbook, without any

formal justification, authors freely switch between different settings (e.g., the ordinary and the extended complex plane), switch between geometric and algebraic exposition, often use many unproved non-trivial facts (regarding them as mathematical ,,folklore"), etc. One of our major contributions was clearing this type of imprecisions and making all the material clear, uniform, and self-contained.

Additionally, we feel that equally (or even more) important to the final result is our experience gained along the way, during our different attempts to reach our final goal. Namely, there are many different ways in which the subject has been exposed in the literature. Comparing, for example, Needham [26] and Schwerdtfeger [30], shows two quite different ways of telling the same story — one more geometrically and the other more algebraically inclined. Our experience shows, that choosing the right approach was the crucial step for making the formalization manageable within the proof assistant — it turned out that more algebraic in its nature the approach was, it was easier to formalize, much nicer, more flexible and more robust.

In the paper, for succinctness, we will present only the basic results of our final formalization — the most important definitions and statements. The present paper contains only a brief recapitulation of the original formal development and many properties that have been formally proved are not going to be shown in the paper. Also, no proofs will be shown nor described, as they are all available in the original Isabelle/HOL proof documents[1]. In the presentation, we will mostly use the original Isabelle/HOL notation, simplifying it a bit in some places to make it more approachable for a wider audience.

Outline of the paper. In Subsection 1.1 we discuss some relevant related work. In Section 2 we describe some features of the theorem prover Isabelle/HOL and describe some background theories used in our formalization. Section 3 is the central section and contains main results of our formalization — in Subsection 3.1 we introduce the extended complex plane, in Subsection 3.2 we introduce Möbius transformations, in Subsection 3.3 we introduce generalized circles, in Subsection 3.4 we discuss circle orientation, and in Subsection 3.5 we discuss some important subgroups of Möbius transformations. In Section 4 we discuss different approaches that we have taken in our formalization, their problems and advantages. Finally, in Section 5 we draw conclusions and discuss some potential further work.

1.1 Related Work

During the last decade, there have been many results in formalizing geometry in proof-assistants. Parts of Hilberts seminal book ,,Foundations of Geometry" [13] have been formalized both in Coq and Isabelle/Isar. Formalization of first two groups of axioms in Coq, in an intuitionistic setting was done by Dehlinger et al. [3]. First formalization in Isabelle/HOL was done by Fleuriot and Meikele [23], and some further developments were made in master thesis of Scott [31]. Large fragments of Tarski's geometry [29] have been formalized in Coq by Narboux et al. [25]. Within Coq, there are also formalizations of von Platos constructive geometry

[1] Isabelle theory files and proof documents are available at `http://argo.matf.bg.ac.rs/formalizations/`

by Kahn [33,17], French high school geometry by Guilhot [8], ruler and compass geometry by Duprat [4], projective geometry by Magaud et al. [19], etc.

In our previous work [22,21], we have already formally investigated a Cartesian model of Euclidean geometry. Timothy Makarios has shown independence of Tarski's Euclidean axiom by formalizing models of Tarski's Euclidean and Tarski's non-Euclidean geometries (the Klein-Beltrami model) [20]. Within that work, the real projective plane has been formalized in Isabelle/HOL.

As a part of the Flyspeck project, Harrison developed a very rich theory (that includes algebra, topology and analysis) of Euclidean n-dimensional space $\mathbb{R}^n$ in theorem prover HOL Light [10,12].

Some automated theorem provers in geometry have also been integrated with proof assistants. For example, Janičić et al. describe a detailed formalization (including implementation details) of the area method [16]. Connecting algebraic methods (Gröbner bases and Wu's methods) with Coq has been done by Grégoire et al. [7] and by Géneveaux et al. [5].

Different results in complex analysis have also been shown in theorem provers. Milewski has proved the fundamental theorem of algebra in Mizar [24], Geuvers et al. have proved the same theorem in Coq [6], Harrison has implemented complex quantifier elimination in HOL and used it in different formalizations, including geometry, etc.

2 Background

In this subsection, we will introduce the theorem prover Isabelle/HOL used for our formalization, its background logic, and notation. We will also briefly describe some results that are part of our formalization, but more general in nature (some lemmas about complex numbers, and the theory of linear algebra of the space $\mathbb{C}^2$).

2.1 Isabelle/HOL

Isabelle [27] is a generic proof assistant, but its most developed application is higher order logic (Isabelle/HOL). Formalizations of mathematical theories are made by defining new notions (types, constants, functions, etc.), and proving statements about them (lemmas, theorems, etc.). This is often done using the declarative proof language Isabelle/Isar [34]. Isar is a very rich language, and we will here describe only the syntax of constructions used in this paper. Definitions are made using the syntax **definition** x **where** "x = ..."", where x is the constant being defined. Lemmas are specified using the syntax **lemma assumes** *assms* **shows** *concl* where *assms* are assumptions and *concl* is the conclusion of the lemma. If there are no assumptions, the keyword **shows** can be omitted. We will also use the syntax **lemma** "$\bigwedge x_1, \ldots x_k.\ [\![asm_1;\ \ldots;\ asm_n]\!] \Longrightarrow concl$" where $asm_1, \ldots, asm_n$ are the assumptions, *concl* is the conclusion, and $x_1, \ldots, x_k$ are universally quantified variables.

Logic formulas are written in the HOL logic using the standard notation (e.g., the connectives $\wedge$, $\vee$, $\longrightarrow$, $\neg$, quantifiers $\forall$ and $\exists$). Terms can use let-bindings (e.g., `let` $x = 3$ `in` $3 * x$) and if-then-else expressions (e.g., `if` $x > 0$ `then` x `else` $-x$), with the standard semantics.

HOL is a typed logic. To express that x is of some type τ we write $x :: \tau$. The predefined type `bool` denotes Booleans, `nat` denotes natural numbers, `int` denotes integers, `real` denotes real numbers, while the type `complex` denotes complex numbers. The imaginary unit is denoted by ii. All these types support ordinary arithmetic operations (e.g., $+$, $-$, $*$, $/$). Conversion from real to complex number will denoted by `cor`, the real and imaginary parts of a complex number by `Re` and `Im`, the complex conjugate by `cnj`, the module of a complex number by $|\_|$, and the argument by `arg` (in Isabelle/HOL it is always in the interval $(-\pi, \pi]$). The complex sign function `sgn` computes the complex number on the unit circle that has the same argument as the given non-zero complex number (i.e., `sgn` $z = z/|z|$). This function is overloaded and it also applies to real numbers (that overloading is mathematically justified as for all real x it holds that `sgn` $(x + ii * 0) =$ `sgn` x). The function `cis` applied to α computes `cos` $\alpha + ii*$`sin` α.

The type of sets containing elements of the type τ is denoted by τ `set`. Isabelle/HOL set-theoretic notation is close to that of standard mathematics, with a few minor exceptions. Set difference is written as $X - Y$, and the image of a function f over a set X is written as $f`X$. The product type is denoted by $\tau_1 \times \tau_2$. Function type is denoted as $\tau_1 \Rightarrow \tau_2$. Functions are usually curried and function applications are written in prefix form, common to functional programming, as `f x` (instead of $f(x)$, that is closer to standard mathematical notation). The predicate `inj` denotes that the function is injective, `bij` that it is a bijection. The predicate `continuous_on X f` denotes that the given function `f` is continuous on the given set `X`. We consider only metric spaces and once we prove that the domain and the co-domain types of `f` are metric spaces for some distance functions (i.e., that they instantiate the `metric_space` type class[2]), all applications of the `continuous_on` predicate implicitly assume those distance functions and their induced topologies.

New types can be introduced in several ways. The simplest way is to use the **type_synonym** command that just introduces a new name for an existing type.

Another way is by using type definitions and then a new type is specified to be isomorphic to some non-empty subset of an existing type. For example, a type can be introduced as **typedef** `three = "{0::nat, 1, 2}"`, generating a proof obligation to show that the type is non-empty. Bijection between the new abstract type and its representation type is given by two functions: `Rep_three :: three` $\Rightarrow$ `nat`, and `Abs_three :: nat` $\Rightarrow$ `three`, satisfying `Rep_three` $x \in \{0, 1, 2\}$, `Rep_three (Abs_three` x`) =` x, and $y \in \{0, 1, 2\} \Longrightarrow$ `Abs_three (Rep_three` y`) =` y. In the rest of the paper, representation functions will be denoted by using $\lfloor\_\rfloor$ brackets, and abstraction functions by using $\lceil\_\rceil$ brackets. The lifting/transfer package [15] can simplify working with types introduced by **typedef**. In that case, users usually need not explicitly use the representation and abstraction functions.

Another way to introduce new types, often used in mathematics, are the quotient types. In Isabelle/HOL, there are several packages that facilitate working with quotients, and our formalization uses the lifting/transfer package [15]. First step in defining quotient type is defining an equivalence relation $\approx$ over some existing (representation) type τ. Quotient type κ is then defined by **quotient_type** κ `=`

[2] Haskell-like type classes [9] are convenient Isabelle/HOL mechanisms for organizing specifications. We say that a type instantiates a type class if there are one on more functions defined on that type that satisfy the assumptions required by that type class. For example, `metric_space` type class requires a distance function (metric) satisfying the standard metric axioms.

τ `/` $\approx$. Functions over the quotient type are defined in two steps. First, a function f_τ `::` ... τ ... is defined over the representation type τ. Then, that function is lifted to the quotient type by using **lift_definition** f_κ `::` ... κ ... **is** f_τ. This generates a proof obligation to show that the definition does not depend on the choice of representative. More details can be found in the literature [18,15].

2.2 Some Background Theories

Complex numbers. Although Isabelle/HOL has some basic support for complex numbers, it was not sufficient for our needs, so we had to make some significant effort and extend it. We have proved many lemmas that are very technical and not interesting for a high-level formalization description so we will not mention them in this paper (e.g., **lemma** "`arg` i `=` $pi/2$" or **lemma** "$|z|^2$ `= Re` $(z *$ `cnj` $z)$"). One of the most useful definitions in this section is the definition of *angle canonization* function $\lfloor \_ \rfloor$, that takes into account 2π periodicity of sine and cosine and maps any angle to its canonical value that lies within the interval $(-\pi, \pi]$. With this function, for example, multiplicative properties of the `arg` function can be easily expressed and proved.

lemma "$z_1 * z_2 \neq 0 \Longrightarrow$ `arg`$(z_1 * z_2)$ `=` $\lfloor$`arg` z_1 + `arg` $z_2 \rfloor$"

Since complex numbers are often treated as vectors, introducing the *scalar product* between two complex numbers (it has been defined as $\langle z_1, z_2 \rangle = (z_1 *$ `cnj` $z_2 + z_2 *$ `cnj` $z_1)/2$) showed out to be useful to succinctly express some conditions.

Linear algebra. Next important theory for further formalization is the theory of linear algebra of $\mathbb{C}^2$. Representing vectors and matrices of arbitrary dimensions pose a challenge in HOL, because of lack of dependent types [10]. There are some available formalizations of n-dimensional matrices and vectors (e.g., the one included in the Isabelle/HOL library or the one available on Archive of Formal Proofs [32]), but none of these includes the notions that we need (e.g., eigenvalues, congruence, diagonalization). In our current formalization and its foreseen extensions we only need to consider finite dimension spaces $\mathbb{C}^2$ and in some situations $\mathbb{R}^3$. Therefore, we have only formalized some linear algebraic properties of these small dimensional spaces. *Complex vectors* (`C2_vec`) are defined as pairs of complex numbers. Similarly, *complex matrices* (`C2_mat`) are defined as 4-tuples of complex numbers (matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is represented by (A, B, C, D)). *Matrix addition* is denoted by $+$, *subtraction* by $-$, *scalar multiplication* of vectors is denoted by $*_{sv}$, and matrices by $*_{sm}$. Both vectors and matrices form vector spaces under these operations. *Scalar product* of two vectors is denoted by $*_{vv}$, the *product of vector and matrix* by $*_{vm}$, the *product of matrix and a vector* by $*_{mv}$, and the *product of two matrices* by $*_{mm}$. Both *zero vector* and *zero matrix* are denoted by `0`, *identity matrix* is denoted by by `eye`, the *determinant* of a matrix is denoted by `mat_det`, its *trace* (the sum of diagonal elements) by `mat_trace`, the *inverse matrix* by `mat_inv`, *transpose* by `mat_transpose`, *conjugation* of every vector element by `vec_cnj`, *conjugation* of every matrix element by `mat_cnj`, etc. Regular matrices form a group under multiplication. Many standard notions of linear algebra have been introduced. For example, *eigenvalues* and *eigenvectors* are defined and characterized in the following way.

definition `eigenval :: "complex ⇒ C2_mat ⇒ bool" where`
`"eigenval k A ⟷ (∃v. v ≠ 0 ∧ A *mv v = k *sv v)"`
lemma `"eigenval k A ⟷ k^2 - mat_trace A * k + mat_det A = 0"`

The *adjoint* of a matrix is its conjugate transpose. *Hermitian* matrices are the ones equal to their adjoint, while *unitary* matrices are the ones whose inverse is equal to their adjoint.

definition `mat_adj where "mat_adj H = mat_cnj (mat_transpose H)"`
definition `hermitian where "hermitian H ⟷ mat_adj H = H"`
definition `unitary where "unitary M ⟷ mat_adj M *mm M = eye"`

Other background notions needed in this paper are going to be introduced along the way, and we refer the reader to our original proof documents for more details.

3 Main Results

3.1 Extended Complex Plane

A very important step in developing the geometry of the complex plane is extending the plane $\mathbb{C}$ with an additional element (treated as the infinite point). The extended plane will be denoted by $\overline{\mathbb{C}}$. There are several different approaches [26, 30] to define $\overline{\mathbb{C}}$. The most appealing approach computationally is the based on homogeneous coordinates, and the most appealing approach visually is based on the stereographic projection of the Riemann sphere.

3.1.1 CP^1 — Homogeneous Coordinates

The extended complex plane $\overline{\mathbb{C}}$ is identified with a complex projective line (the one-dimensional projective space over the complex field, sometimes denoted by CP^1). Each point of $\overline{\mathbb{C}}$ is represented by a pair of complex homogeneous coordinates (not both equal to zero). Two pairs of homogeneous coordinates represent the same point in $\overline{\mathbb{C}}$ iff they are proportional by a non-zero complex factor. Isabelle/HOL formalization of this concept relies on the lifting/transfer package for quotients [15] and is done in three stages[3].

First, the type of non-zero pairs of complex numbers (also treated as non-zero complex vectors) is introduced.

typedef $\texttt{C2\_vec}_{\neq 0}$ `= "{v::C2_vec. v ≠ 0}"`

This gives the representation function $\texttt{Rep\_C2\_vec}_{\neq 0}$ (that we will denote by $\lfloor\_\rfloor_{C2}$) returning a (non-zero) pair of complex numbers for each given element of the auxiliary type $\texttt{C2\_vec}_{\neq 0}$ and the abstraction function $\texttt{Abs\_C2\_vec}_{\neq 0}$ (that we will denote by $\lceil\_\rceil^{C2}$) returning an element of $\texttt{C2\_vec}_{\neq 0}$ for each given non-zero pair of complex numbers.

Second, two elements of the type $\texttt{C2\_vec}_{\neq 0}$ are said to be equivalent iff their representations are proportional.

[3] One stage could be avoided by using partial quotients offered by the lifting/transfer package. This feature has not been used in our formalization due to some problems in the early versions of the quotient package. All problems have been fixed in the meantime, but our formalization was quite developed, and it would be quite tedious to change it.

definition $\approx_{C2}$ :: "C2_vec$_{\neq 0}$ ⇒ C2_vec$_{\neq 0}$ ⇒ bool" **where**
"$z_1 \approx_{C2} z_2 \longleftrightarrow$ (∃ (k::complex). $k \neq 0 \wedge \lfloor z_2 \rfloor_{C2} = k *_{sv} \lfloor z_1 \rfloor_{C2}$)"

It is quite easy to show that $\approx_{C2}$ is an equivalence relation.

Finally, the type of extended complex numbers given by homogeneous coordinates are defined as equivalence classes of $\approx_{C2}$ and are introduced as the following quotient type.

quotient_type complex$_{hc}$ = C2_vec$_{\neq 0}$ / $\approx_{C2}$

To summarize, on the lowest representation level there is the type of pairs of complex numbers, on the next level there is the type of non-zero complex 2×2 vectors (represented by the previous type) and on the highest level there is the quotient type inhabited by equivalence classes — dealing with this quotient type (its representation and abstraction) is done behind the scenes, by the lifting and transfer package [15]. These three layers of abstraction can be confusing for an ordinary mathematician who is used to identify them, but they are necessary in a formal setting where each object must have a unique type (for example, it is usual to consider that $(1, i)$ is both a pair of complex numbers, and a non-zero complex vector, but in our formalization $(1, i)$ is a pair of complex numbers, while $\lceil (1, i) \rceil^{C2}$ is a non-zero complex vector). In the paper we will always use a non-aggressive notation ($\lfloor \_ \rfloor$ and $\lceil \_ \rceil$) for representation and abstraction functions. Just ignoring these brackets can make the text more approachable and more like the ordinary mathematical texts.

Ordinary and infinite numbers. Each ordinary complex number can be converted to an extended complex number.

definition of_complex_rep :: "complex ⇒ C2_vec$_{\neq 0}$" **where**
of_complex_rep z = $\lceil (z, 1) \rceil^{C2}$
lift_definition of_complex :: "complex ⇒ complex$_{hc}$" **is** of_complex_rep

The single point at infinity is defined the following way

definition inf_hc_rep :: C2_vec$_{\neq 0}$ **where** inf_hc_rep = $\lceil (1, 0) \rceil^{C2}$
lift_definition ∞_{hc} :: "complex$_{hc}$" **is** inf_hc_rep

It is easily shown that all extended complex numbers are either ∞_{hc} (iff their second homogeneous coordinate is zero) or can be obtained by converting from an ordinary complex number (iff their second homogeneous coordinate is not zero).

lemma "z = ∞_{hc} ∨ (∃ x. z = of_complex x)"

Notation 0_{hc}, 1_{hc} and i_{hc} is used to denote extended complex counterparts of 0, 1, and i.

Arithmetic operations. Arithmetic operations on ordinary complex numbers can be extended to the extended complex plane.

On the lowest, representation level, the *addition* of (z_1, z_2) and (w_1, w_2) is defined as $(z_1 * w_2 + w_1 * z_2, z_2 * w_2)$, i.e.,

definition plus_hc_rep :: "C2_vec$_{\neq 0}$ ⇒ C2_vec$_{\neq 0}$ ⇒ C2_vec$_{\neq 0}$"
where "plus_hc_rep z w = (**let** (z_1, z_2) = $\lfloor z \rfloor_{C2}$; (w_1, w_2) = $\lfloor w \rfloor_{C2}$
in $\lceil (z_1 * w_2 + w_1 * z_2, z_2 * w_2) \rceil^{C2}$)"

This gives a non-zero pair of homogeneous coordinates unless both z_2 and w_2 are zero (corresponding to the sum of two infinite values), otherwise, it gives an ill-defined element $\lceil(0,0)\rceil^{C2}$.[4] The definition is lifted to the quotient type.

lift_definition $+_{hc}$ :: "complex$_{hc}$ ⇒ complex$_{hc}$ ⇒ complex$_{hc}$" **is** plus_hc_rep

This generates the proof obligation $[\![z \approx_{C2} z'; w \approx_{C2} w']\!] \Longrightarrow z +_{hc} w \approx_{C2} z' +_{hc} w'$, that is easily proved by case analysis on whether both z_2 and w_2 are zero. Note that, due to the requirement of HOL that all functions are total, we could not define the function only for the well-defined cases, and in the lifting proofs we also had to deal with the ill-defined cases.

Next, it is shown that this operation extends the ordinary addition of complex numbers (the operation + on $\mathbb{C}$).

lemma "of_complex z $+_{hc}$ of_complex w = of_complex $(z + w)$"

The sum of an ordinary complex number and ∞_{hc} is ∞_{hc} (however, $\infty_{hc} +_{hc} \infty_{hc}$ is ill-defined).

lemma "of_complex z $+_{hc}$ ∞_{hc} = ∞_{hc}"
lemma "∞_{hc} $+_{hc}$ of_complex z = ∞_{hc}"

The operation $+_{hc}$ is associative and commutative, but ∞_{hc} does not have an inverse, so $+_{hc}$ on $\overline{\mathbb{C}}$ does not have the nice algebraic properties of + on $\mathbb{C}$.

Other arithmetic operations are also extended to $\overline{\mathbb{C}}$. On the lowest, representation type, the *unary minus* of (z_1, z_2) is $(-z_1, z_2)$, the *multiple* of (z_1, z_2) and (w_1, w_2) is $(z_1 * z_2, w_1 * w_2)$, and the *reciprocal* of (z_1, z_2) is (z_2, z_1) – these operations are then lifted to the abstract quotient type yielding the operations denoted by uminus$_{hc}$, $*_{hc}$, and recip$_{hc}$. *Subtraction* (denoted by $-_{hc}$) is defined by using $+_{hc}$ and uminus$_{hc}$, and *division* (denoted by $:_{hc}$) by using $*_{hc}$ and recip$_{hc}$. As in the case of addition, it is shown that all these operations match the ordinary operations on the finite part of the extended complex plane (e.g. **lemma** uminus$_{hc}$ (of_complex z) = of_complex $(-z)$). Next lemmas show the behavior of these operation when the infinite point is involved (note that the expressions $0_{hc} *_{hc} \infty_{hc}$, $\infty_{hc} *_{hc} 0_{hc}$, $0_{hc} :_{hc} 0_{hc}$, and $\infty_{hc} :_{hc} \infty_{hc}$ are ill-defined).

lemma "uminus$_{hc}$ ∞_{hc} = ∞_{hc}"
lemma "recip$_{hc}$ ∞_{hc} = 0_{hc}" "recip$_{hc}$ 0_{hc} = ∞_{hc}"
lemma "$z \neq 0_{hc} \Longrightarrow z *_{hc} \infty_{hc} = \infty_{hc} \quad \wedge \quad \infty_{hc} *_{hc} z = \infty_{hc}$"
lemma "$z \neq 0_{hc} \Longrightarrow z :_{hc} \infty_{hc} = 0_{hc}$"
lemma "$z \neq \infty_{hc} \Longrightarrow \infty_{hc} :_{hc} z = \infty_{hc}$"

Complex *conjugation* is also extended to $\overline{\mathbb{C}}$ (on the representation type (z_1, z_2) is mapped to $(\overline{z_1}, \overline{z_2})$), giving the operation cnj$_{hc}$. A very important operation in complex geometry is the *inversion over the unit circle*:

[4] All the functions (including the abstraction function $\lceil\_\rceil^{C2}$) in HOL are total. However, all the provided lemmas about that function include the precondition that its argument is not $(0,0)$. Therefore, there is no way to reason about the value $\lceil(0,0)\rceil^{C2}$ and it should be considered to be ill-defined. The sum $\infty_{hc} +_{hc} \infty_{hc}$ cannot be defined so that $\overline{\mathbb{C}}$ becomes a group under addition — the law $-a + a = 0$ requires that $\infty_{hc} +_{hc} \infty_{hc} = 0_{hc}$ (since the opposite element of ∞_{hc} must be ∞_{hc}), but that would break the associativity since then it holds that $(\infty_{hc} +_{hc} \infty_{hc}) +_{hc} 1_{hc} = 1_{hc} \neq 0_{hc} = \infty_{hc} +_{hc} (\infty_{hc} +_{hc} 1_{hc})$.

definition inversion$_{hc}$ :: "complex$_{hc}$ ⇒ complex$_{hc}$" **where**
"inversion$_{hc}$ = cnj$_{hc}$ ∘ recip$_{hc}$"

The most basic properties of inversion are then easily proved.

lemma "inversion$_{hc}$ ∘ inversion$_{hc}$ = id"
lemma "inversion$_{hc}$ 0_{hc} = ∞_{hc}" "inversion$_{hc}$ ∞_{hc} = 0_{hc}"

Ratio and cross ratio. The (simple) *ratio* and the *cross-ratio* are very important concepts in projective geometry and the extended complex plane (cross-ratio is a characterizing invariant of Möbius transformations – the fundamental transformations of $\overline{\mathbb{C}}$, and it is possible to define lines using ratio and circles using cross-ratio of points).

Ratio of points z, v and w is usually defined as $\frac{z-v}{z-w}$. Our definition introduces it in homogeneous coordinates.

definition ratio_rep **where** "ratio_rep z v w =
(let (z_1, z_2) = $\lfloor z \rfloor_{C2}$; (v_1, v_2) = $\lfloor v \rfloor_{C2}$; (w_1, w_2) = $\lfloor w \rfloor_{C2}$
in $\lceil((z_1 * v_2 - v_1 * z_2) * w_2, (z_1 * w_2 - w_1 * z_2) * v_2)\rceil^{C2}$)"
lift_definition ratio :: "complex$_{hc}$ ⇒ complex$_{hc}$ ⇒ complex$_{hc}$ ⇒ complex$_{hc}$"
is ratio_rep

Note that this is well-defined in all cases except when $z = w = v$ or $z = v = \infty_{hc}$ or $z = w = \infty_{hc}$ or $v = w = \infty_{hc}$ (however, in the lifting proofs these ill-defined cases must also be covered). The original ratio of differences is defined in all cases except when $z = w = v$ or $z = \infty_{hc}$ or $v = w = \infty_{hc}$, so our definition in homogeneous coordinates naturally extends the original definition. Following lemmas show the behavior of the ratio in all well-defined cases (it matches the original ratio of differences whenever it is defined).

lemma "$[\![z \neq v \vee z \neq w;\ z \neq \infty_{hc};\ v \neq \infty_{hc} \vee w \neq \infty_{hc}]\!] \Longrightarrow$
ratio z v w = (z $-_{hc}$ v) $:_{hc}$ (z $-_{hc}$ w)"
lemma $[\![v \neq \infty_{hc};\ w \neq \infty_{hc}]\!] \Longrightarrow$ ratio ∞_{hc} v w = 1_{hc}
lemma $[\![z \neq \infty_{hc};\ w \neq \infty_{hc}]\!] \Longrightarrow$ ratio z ∞_{hc} w = ∞_{hc}
lemma $[\![z \neq \infty_{hc};\ v \neq \infty_{hc}]\!] \Longrightarrow$ ratio z v ∞_{hc} = 0_{hc}

The last two lemmas are consequences of the first one. Also, note that the ratio cannot be defined for the case when at least two points are infinite in a natural way (so that the ratio function remains continuous in all of its parameters).

The cross-ratio is defined over 4 points (z, u, v, w), usually as $\frac{(z-u)(v-w)}{(z-w)(v-u)}$. Again, we define it using homogeneous coordinates.

definition cross_ratio_rep **where** "cross_ratio_rep z u v w =
(let (z_1, z_2) = $\lfloor z \rfloor_{C2}$; (u_1, u_2) = $\lfloor u \rfloor_{C2}$;
(v_1, v_2) = $\lfloor v \rfloor_{C2}$; (w_1, w_2) = $\lfloor w \rfloor_{C2}$ in
$\lceil(z_1 * u_2 - u_1 * z_2) * (v_1 * w_2 - w_1 * v_2), (z_1 * w_2 - w_1 * z_2) * (v_1 * u_2 - u_1 * v_2))\rceil^{C2}$"
lift_definition cross_ratio :: "complex$_{hc}$ ⇒ complex$_{hc}$ ⇒
complex$_{hc}$ ⇒ complex$_{hc}$ ⇒ complex$_{hc}$" **is** cross_ratio_rep

This is well-defined in all cases except when $z = u = w$ or $z = v = w$ or $z = u = v$ or $u = v = w$ (note that infinite values for z, u, v or w are allowed, which is not the case in the original fractional formulation). Some basic properties of the cross-ratio are given by the following lemmas.

lemma "$[\![(z \neq u \wedge v \neq w) \vee (z \neq w \wedge u \neq v); z \neq \infty_{hc}; u \neq \infty_{hc}; v \neq \infty_{hc}; w \neq \infty_{hc}]\!]$
$\Longrightarrow$ `cross_ratio` $z\ u\ v\ w$ = $((z -_{hc} u) *_{hc} (v -_{hc}) :_{hc} ((z -_{hc} w) *_{hc} (v -_{hc} u))$"
lemma "`cross_ratio` $z\ 0_{hc}\ 1_{hc}\ \infty_{hc}$ = z"
lemma "$[\![\ z_1 \neq z_2; z_1 \neq z_3\ ]\!] \Longrightarrow$ `cross_ratio` $z_1\ z_1\ z_2\ z_3$ = 0_{hc}"
lemma "$[\![\ z_2 \neq z_1; z_2 \neq z_3\ ]\!] \Longrightarrow$ `cross_ratio` $z_2\ z_1\ z_2\ z_3$ = 1_{hc}"
lemma "$[\![\ z_3 \neq z_1; z_3 \neq z_2\ ]\!] \Longrightarrow$ `cross_ratio` $z_3\ z_1\ z_2\ z_3$ = ∞_{hc}"

3.1.2 Riemann Sphere and Stereographic Projection

The extended complex plane can be identified with a *Riemann (unit) sphere* Σ by means of *stereographic projection* [26,30]. The sphere is projected from its north pole N to the xOy plane (identified with $\mathbb{C}$). This projection establishes a bijective map sp between $\Sigma \backslash N$ and the finite complex plane $\mathbb{C}$. The infinite point is defined as the image of N.

In Isabelle/HOL, the sphere Σ is defined as a new type.

typedef `riemann_sphere` = "$\{(x, y, z)$::`R3_vec`. $x^2 + y^2 + z^2 = 1\}$"

Again, this defines functions `Rep_riemann_sphere` (that will be denoted by $\lfloor\_\rfloor_{R3}$ and `Abs_riemann_sphere` (that will be denoted by $\lceil\_\rceil^{R3}$) that connect the points of the abstract type (`riemann_sphere`) and the representation type (triples of real numbers). Stereographic projection is introduced in the following way:

definition `stereographic_rep` :: "`riemann_sphere` $\Rightarrow$ `C2_vec`$_{\neq 0}$" **where**
"`stereographic_rep` M =
(`let` (x, y, z) = $\lfloor M \rfloor_{R3}$
`in if` $(x, y, z) \neq (0, 0, 1)$ `then` $\lceil (x + i * y, 1 - z) \rceil^{C2}$ `else` $\lceil (1, 0) \rceil^{C2}$)"
lift_definition `stereographic` :: "`riemann_sphere` $\Rightarrow$ `complex`$_{hc}$" **is**
`stereographic_rep`

For all points, this is well-defined (the vector $(x + i * y, 1 - z)$ is non-zero as $(x, y, z) \neq (0, 0, 1)$, and $(1, 0)$ is clearly non-zero).

Inverse stereographic projection is defined in the following way.

definition `inv_stereographic_rep` :: "`C2_vec`$_{\neq 0}$ $\Rightarrow$ `riemann_sphere`" **where**
"`inv_stereographic_rep` z =
(`let` (z_1, z_2) = $\lfloor z \rfloor_{C2}$
`in if` z_2 = 0 `then` $\lceil (0, 0, 1) \rceil^{R3}$
`else let` z = z_1/z_2; XY = $2 * z$ / `cor` $(1 + |z|^2)$; Z = $(|z|^2 - 1)/(1 + |z|^2)$
`in` $\lceil (Re\ XY,\ Im\ XY,\ Z) \rceil^{R3}$)"
lift_definition `inv_stereographic` :: "`complex`$_{hc}$ $\Rightarrow$ `riemann_sphere`" **is**
`inv_stereographic_rep`

For all points this is well-defined (the sum of squares of three coordinates is 1 in both cases so the `Abs_riemann_sphere` function can safely be applied).

The connection between the two functions is given by the following lemmas.

lemma "`stereographic` $\circ$ `inv_stereographic` = `id`"
lemma "`inv_stereographic` $\circ$ `stereographic` = `id`"
lemma "`bij stereographic`" "`bij inv_stereographic`"

The proofs are not difficult but require formalizing some tedious calculations.

Chordal distance. Riemann sphere can be made a metric space. One of the most common ways to introduce metric is *chordal metric* – distance between two points on the sphere is the length of the chord that joins them.

definition $\texttt{dist}_{rs}$ :: "riemann_sphere ⇒ riemann_sphere ⇒ real" **where**
"$\texttt{dist}_{rs}\ M_1\ M_2$ = (let (x_1, y_1, z_1) = $\lfloor M_1 \rfloor_{R3}$; (x_2, y_2, z_2) = $\lfloor M_2 \rfloor_{R3}$
in norm $(x_1 - x_2, y_1 - y_2, z_1 - z_2)$)"

The function `norm` is a Isabelle/HOL library function and in this case it computes the Euclidean vector norm in $\mathbb{R}^3$. Using the (already available) fact that $\mathbb{R}^3$ is a metric space (under the distance function $\lambda\ x\ y.\ \texttt{norm}(x - y)$), it was not difficult to show that the type `riemann_sphere` equipped with $\texttt{dist}_{rs}$ is a metric space, i.e., an instantiation of the `metric_space` type class.

Although it is defined on the sphere, the chordal metric has its representation in the plane.

lemma assumes
"stereographic M_1 = of_complex m_1" "stereographic M_2 = of_complex m_2"
shows "$\texttt{dist}_{rs}\ M_1\ M_2$ = $2 * |m_1 - m_2|$ / (sqrt $(1 + |m_1|^2)$ * sqrt $(1 + |m_2|^2)$)"
lemma assumes "stereographic M_1 = ∞_{hs}" "stereographic M_2 = of_complex m"
shows "$\texttt{dist}_{rs}\ M_1\ M_2$ = 2 / sqrt $(1 + |m|^2)$"
lemma assumes "stereographic M_1 = of_complex m" "stereographic M_2 = ∞_{hs}"
shows "$\texttt{dist}_{rs}\ M_1\ M_2$ = 2 / sqrt $(1 + |m|^2)$"
lemma assumes "stereographic M_1 = ∞_{hs}" "stereographic M_2 = ∞_{hs}"
shows "$\texttt{dist}_{rs}\ M_1\ M_2$ = 0"

These lemmas make a distinction between finite and infinite points, but this case analysis can be avoided if homogeneous coordinates are used.

definition "$\langle\!\langle z, w \rangle\!\rangle$ = (vec_cnj $\lfloor z \rfloor_{C2}$) $*_{vv}$ ($\lfloor w \rfloor_{C2}$)"
definition "$\langle\!\langle z \rangle\!\rangle$ = sqrt (Re $\langle\!\langle z, z \rangle\!\rangle$)"
definition "dist_hc_rep = 2 * sqrt $(1 - |\langle\!\langle z, w \rangle\!\rangle|^2 / (\langle\!\langle z \rangle\!\rangle^2 * \langle\!\langle w \rangle\!\rangle^2))$"
lift_definition $\texttt{dist}_{hc}$ :: "$\texttt{complex}_{hc}$ ⇒ $\texttt{complex}_{hc}$ ⇒ real **is** dist_hc_rep
lemma "$\texttt{dist}_{rs}\ M_1\ M_2$ = $\texttt{dist}_{hc}$ (stereographic M_1) (stereographic M_2)"

This form is sometimes called *Fubini-Study metric.*

The type $\texttt{complex}_{hc}$ equipped with the $\texttt{dist}_{hc}$ metric is also an instantiation of the `metric_space` type class. This trivially follows from the last lemma that connects it to the metric space on the Riemann sphere. There are also direct proofs of this (e.g., Hille [14] gives a direct proof due to Shizuo Kakutani, however the proof is incomplete as the possibility of one point being infinite is not considered) and we have formalized them[5]. It turned out that some properties (e.g., the triangle inequality) are easier to prove on the Riemann sphere using the function $\texttt{dist}_{rs}$, but some properties (e.g., that the metric space is perfect, i.e., that it does not have isolated points) are easier to prove in the projection using the function $\texttt{dist}_{hc}$, indicating the significance of having different models of the same concept.

Using the chordal metric in the extended plane, and the Euclidean metric on the sphere in $\mathbb{R}^3$, the stereographic and inverse stereographic projections are proved to be continuous.

[5] Our formalization started without considering the Riemann sphere and so we could only use a direct proof in the beginning, but at one point we introduced the Riemann sphere and using it explicitly simplified many proofs, including this one.

```
lemma "continuous_on UNIV stereographic"
      "continuous_on UNIV inv_stereographic"
```

Note that in the previous lemma, metrics are implicit (as described in Section 2).

3.2 Möbius Transformations

Möbius transformations (also called homographic, linear fractional, or bilinear transformations) are the fundamental transformations of the extended complex plane. In our formalization they are introduced algebraically. Each transformation is represented by a regular (non-singular, non-degenerate) 2×2 matrix that acts linearly on homogeneous coordinates. As proportional homogeneous coordinates represent same points of $\overline{\mathbb{C}}$, proportional matrices will represent the same Möbius transformation. Again, the formalization proceeds in three steps using the lifting/transfer package. First, the type of regular matrices is introduced.

```
typedef C2_mat_reg = "{M :: C2_mat. mat_det M ≠ 0}"
```

The representation function `Rep_C2_mat_reg` will be denoted by $\lfloor \_ \rfloor_M$ and the abstraction function `Abs_C2_mat_reg` will be denoted by $\lceil \_ \rceil^M$. Regular matrices form a group under multiplication that is usually called *general linear group* and denoted by $GL(2, \mathbb{C})$. In some cases its subgroup, *special linear group*, denoted by $SL(2, \mathbb{C})$, and containing only the matrices with the determinant 1 is considered.

Möbius group. Two regular matrices are considered to be equivalent iff their representations are proportional.

```
definition ≈M :: "C2_mat_reg ⇒ C2_mat_reg ⇒ bool" where
  "M1 ≈M M2 ⟷ (∃ (k::complex). k ≠ 0 ∧ ⌊M2⌋M = k *sm ⌊M1⌋M)"
```

It is easy to show that this is an equivalence relation. *Möbius elements* are introduced as equivalence classes over this relation.

```
quotient_type mobius = C2_mat_reg / ≈M
```

We will sometimes use the auxiliary constructor `mk_mobius` that returns a Möbius element (an equivalence class) for the given 4 complex parameters (it makes sense only when the corresponding matrix is regular).

Möbius elements form a group under operations that will now define. This group is called the *projective general linear group* and denoted by $PGL(2, \mathbb{C})$. Again, $SGL(2, \mathbb{C})$ containing elements with the determinant 1 can be considered. *Composition of Möbius elements* is obtained by multiplying their representing matrices.

```
definition mobius_comp_rep :: "C2_mat_reg ⇒ C2_mat_reg ⇒ C2_mat_reg"
  where "moebius_comp_rep M1 M2 = ⌈⌊M1⌋M *mm ⌊M2⌋M⌉M"
lift_definition mobius_comp :: "mobius ⇒ mobius ⇒ mobius" is
  mobius_comp_rep
```

Similarly, the *inverse Möbius element* is obtained by taking the inverse representative matrix.

definition mobius_inv_rep :: "C2_mat_reg ⇒ C2_mat_reg" **where**
"mobius_inv_rep M = $\lceil$mat_inv $\lfloor M \rfloor_M \rceil^M$"
lift_definition mobius_inv :: "mobius ⇒ mobius" **is** "mobius_inv_rep"

Finally, *identity Möbius element* is represented by the identity matrix.

definition mobius_id_rep :: "C2_mat_reg" **where** "mobius_id_rep = $\lceil$eye$\rceil^M$"
lift_definition mobius_id :: "mobius" **is** mobius_id_rep

All these definitions always introduce well-defined objects (as the product of regular matrices is regular and the inverse of a regular matrix is regular). Proof obligations necessary to lift the definitions (e.g., $M_1 \approx_M M_2 \Longrightarrow$ mobius_inv_rep $M_1 \approx_M$ mobius_inv_rep M_2) are easily discharged. Composition, inverse and identity establish the group structure on the set of Möbius elements. This is shown by showing that the type mobius along with these operations is an instantiation of the group_add type class built-in Isabelle/HOL. Therefore, we will sometimes denote mobius_comp by $+$, mobius_inv by unary $-$, and mobius_id by 0, make sums of Möbius elements $f + g$, differences of elements $f - g$, and so on.

Möbius group action. Action of every Möbius group element on the points of the extended complex plane $\overline{\mathbb{C}}$ induces a mapping from $\overline{\mathbb{C}}$ to $\overline{\mathbb{C}}$ that is a *Möbius transformation*. The action is given by the function mobius_pt.

definition mobius_pt_rep :: "C2_mat_reg ⇒ C2_vec$_{\neq 0}$ ⇒ C2_vec$_{\neq 0}$"
where "moebius_pt_rep M z = $\lceil \lfloor M \rfloor_M$ $*_{mv}$ $\lfloor z \rfloor_{C2} \rceil^{C2}$"
lift_definition mobius_pt :: "mobius ⇒ complex$_{hc}$ ⇒ complex$_{hc}$" **is**
mobius_pt_rep

Since the product of a regular matrix and a non-zero vector is a non-zero vector, the result is always well-defined. Lifting the definition generates the obligation $[\![M \approx_M M';\ z \approx_{C2} z']\!] \Longrightarrow$ mobius_pt_rep M $z \approx_{C2}$ mobius_pt_rep M' z', that is quite easily discharged.

Group operations on Möbius elements correspond to operations on their induced Möbius transformations (composition of mappings, inverse mapping and the identity mapping).

lemma "mobius_pt (mobius_comp M_1 M_2) = (mobius_pt M_1) ∘ (mobius_pt M_2)"
lemma "mobius_pt (mobius_inv M) = inv (mobius_pt M)"
lemma "mobius_pt (mobius_id) = id"

The action is transitive (as it is always a bijective map).

lemma "bij (mobius_pt M)"

In the classic literature Möbius transformations are often expressed in the form $\frac{az+b}{cz+d}$, and the following lemma justifies this (but with a special case for the infinite argument z).

lemma assumes "mat_det $(a, b, c, d) \neq 0$"
shows "moebius_pt (mk_mobius a b c d) z =
(if $z \neq \infty_{hc}$ then
((of_complex a) $*_{hc}$ z $+_{hc}$ (of_complex b)) $:_{hc}$
((of_complex c) $*_{hc}$ z $+_{hc}$ (of_complex d))
else (of_complex a) $:_{hc}$ (of_complex c))"

An arbitrary transformation of $\overline{\mathbb{C}}$ is a Möbius transformation iff it is an action of some Möbius group element.

definition `is_mobius :: "(complex`$_{hc}$ `⇒ complex`$_{hc}$`) ⇒ bool" where`
`"is_mobius` f `⟷ (∃` M`. ` f `= mobius_pt` M`)"`

Note that most results listed so far depend on the fact that the representation matrix of the Möbius transformation is regular — otherwise, the action would be degenerate and crush the whole plane $\overline{\mathbb{C}}$ into a single point.

Some special Möbius transformations. Many transformations encountered in geometry are special kinds of Möbius transformations. Very important subgroup is the group of *Euclidean similarities* (also called *integral transformations*). They are determined by using two complex parameters (and represent Möbius transformations when the first one is not zero).

definition `similarity :: "complex ⇒ complex ⇒ mobius" where`
`"similarity` a b `= mk_mobius` a b 0 1`"`

Similarities form a group (that is sometimes called the *parabolic group*).

lemma `"⟦`$a \neq 0$`; ` $c \neq 0$`⟧ ⟹ mobius_comp (similarity` a b`) (similarity` c d`) =`
`similarity` $(a * c)$ $(a * d + b)$`"`
lemma `"`$a \neq 0$ `⟹ mobius_inv (similarity` a b`) = similarity` $(1/a)$ $(-b/a)$`"`
lemma `"id_mobius = similarity` 1 0`"`

Their action is a linear transformation of $\mathbb{C}$, and each non-constant linear transformation of $\mathbb{C}$ is the action of an element of the similarity group.

lemma `"`$a \neq 0$ `⟹ mobius_pt (similarity` a b`) =`
`(λ` z`. (of_complex` a`) *`$_{hc}$ z `+`$_{hc}$ `(of_complex` b`))"`

Euclidean similarities are the only Möbius group elements such that their action leaves the ∞_{hc} fixed.

lemma `"mobius_pt` M ∞_{hc} `= `∞_{hc} `⟷ (∃` a b`. `$a \neq 0$ `∧` M `= similarity` a b`)"`

If both ∞_{hc} and 0_{hc} are fixed, then its a similarity with coefficients a and $b = 0$, and the action is of the form `λ` z`. (of_complex` a`) *`$_{hc}$ z.

lemma `"mobius_pt` M ∞_{hc} `= `∞_{hc} `∧ mobius_pt` M 0_{hc} `= `0_{hc} `⟷`
`(∃` a`. `$a \neq 0$ `∧` M `= similarity` a 0`)"`

Euclidean similarities include translations, rotations, and dilatations, and every Euclidean similarity can be decomposed using these.

definition `"translation` v `= similarity` 1 v`"`
definition `"rotation` ϕ `= similarity (cis` ϕ`)` 0`"`
definition `"dilatation` k `= similarity (cor` k`)` 0`"`
lemma `"`$a \neq 0$ `⟹ similarity` a b `=`
`(translation` b`) + (rotation (arg` a`)) + (dilatation` $|a|$`)"`

Reciprocal $(1_{hc} :_{hc} z)$ is also a Möbius transformation.

definition `"reciprocation = mk_mobius` $(1, 0, 0, 1)$`"`
lemma `"recip`$_{hc}$ `= mobius_pt reciprocation"`

On the other hand, inversion is not a Möbius transformation (it is a canonical example of so-called anti-Möbius transformations, or antihomographies).

A very important fact is that every Möbius transformation can be composed of Euclidean similarities and a reciprocation. One possible way to achieve this is given by the following lemma (the case when $c = 0$ is the case of Euclidean similarities, and it has already been analyzed).

lemma assumes "$c \neq 0$" **and** "$a * d - b * c \neq 0$"
shows "`mk_mobius` $a\ b\ c\ d$ =
`translation` (a/c) + `rotation_dilatation` $((b * c - a * d)/(c * c))$ +
`reciprocal` + `translation` (d/c)"

Decomposition is used in many proofs. Namely, to show that every Möbius transformation has some property, it suffices to show that reciprocation and all Euclidean similarities have that property, and that the property is preserved under compositions (usually, most of the effort goes to proving the reciprocation case, while the rest is much simpler).

lemma assumes "$\bigwedge v.\ P$ (`translation` v)" "$\bigwedge \alpha.\ P$ (`rotation` α)"
"$\bigwedge k.\ P$ (`dilatation` k)" "P (`reciprocation`)"
"$\bigwedge M_1\ M_2.\ [\![P\ M_1;\ P\ M_2]\!] \Longrightarrow P\ (M_1 + M_2)$"
shows "$P\ M$"

Cross-ratio as a Möbius transformation For any fixed three points z_1, z_2 and z_3, `cross_ratio` $z\ z_1\ z_2\ z_3$ can be seen as a function of a single variable z. The following lemma guarantees that this function is a Möbius transformation, and by the properties of the cross-ratio it maps z_1 to 0_{hc}, z_2 to 1_{hc} and z_3 to ∞_{hc}.

lemma "$[\![z_1 \neq z_2;\ z_1 \neq z_3;\ z_2 \neq z_3]\!] \Longrightarrow$
`is_mobius` ($\lambda\ z.$ `cross_ratio` $z\ z_1\ z_2\ z_3$)"

Then, the cross-ratio can be used to show that there is a Möbius transformation mapping any three different points to 0_{hc}, 1_{hc} and ∞_{hc}, respectively. Since Möbius transformations form a group, a simple consequence of this is that there is a Möbius transformation mapping any three different points to any three different points.

lemma "$[\![z_1 \neq z_2;\ z_1 \neq z_3;\ z_2 \neq z_3]\!] \Longrightarrow (\exists\ M.$ `mobius_pt` $M\ z_1 = 0_{hc} \wedge$
`mobius_pt` $M\ z_2 = 1_{hc} \wedge$ `mobius_pt` $M\ z_3 = \infty_{hc}$)"

The next lemma turns out to have very important applications in further proof development, as it enables so-called ,,without-loss-of-generality (wlog)" reasoning [11]. Namely, if the property is preserved under Möbius transformations, then instead of showing that the property holds for any three different points one can only show that the property holds for points 0_{hc}, 1_{hc}, and ∞_{hc}.

lemma assumes "$P\ 0_{hc}\ 1_{hc}\ \infty_{hc}$" "$z_1 \neq z_2$" "$z_1 \neq z_3$" "$z_2 \neq z_3$"
"$\bigwedge M\ u\ v\ w.\ P\ u\ v\ w \Longrightarrow$
P (`mobius_pt` $M\ u$) (`mobius_pt` $M\ v$) (`mobius_pt` $M\ w$)"
shows "$P\ z_1\ z_2\ z_3$"

One of the first applications of ,,wlog" reasoning for Möbius is in analyzing fixed points of Möbius transformations. It is easy to show that only the identity transformation has the fixed points 0_{hc}, 1_{hc}, and ∞_{hc}. It also holds that if a Möbius transformation M has three different fixed points, it is the identity transformation. The direct proof of this relies on the fact that a 2×2 matrix has at most two independent eigenvectors, and that can be easily avoided using ,,wlog" reasoning (as any three different points can be mapped to 0_{hc}, 1_{hc}, and ∞_{hc} by some M' and then $M' + M - M'$ has these three points fixed so it must be 0).

lemma "⟦ `mobius_pt` M 0_{hs} = 0_{hs}; `mobius_pt` M 1_{hs} = 1_{hs};
`mobius_pt` M ∞_{hs} = ∞_{hs} ⟧ ⟹ `M = id_mobius`"
lemma "⟦ `mobius_pt` M z_1 = z_1; `mobius_pt` M z_2 = z_2;
`mobius_pt` M z_3 = z_3; $z_1 \neq z_2$; $z_1 \neq z_3$; $z_2 \neq z_3$ ⟧ ⟹ `M = id_mobius`"

A consequence of this is that there is a unique Möbius transformation mapping three different points to other three different points (it has already been shown that there exists such transformation and if there were two, then their difference would have three different fixed points so it would be identity).

lemma "⟦$z_1 \neq z_2$; $z_1 \neq z_3$; $z_2 \neq z_3$; $w_1 \neq w_2$; $w_1 \neq w_3$; $w_2 \neq w_3$⟧ ⟹ ∃! M.
`mobius_pt` M z_1 = w_1 ∧ `mobius_pt` M z_2 = w_2 ∧ `mobius_pt` M z_3 = w_3"

Möbius transformations preserve cross-ratio. Again, a direct proof would be complicated, so an elegant indirect proof has been formalized (basically, the difference of $\lambda z.$ `cross_ratio` z z_1 z_2 z_3 and M maps $(M\ z_1)$ to 0_{hc}, $(M\ z_2)$ to 1_{hc}, and $(M\ z_3)$ to ∞_{hc}, therefore it must be equal to $\lambda z.$ `cross_ratio` z `(`M z_1`)` `(`M z_2`)` `(`M z_3`)`, and the statement follows by substituting $(M\ z)$ for z).

lemma "⟦$z_1 \neq z_2$; $z_1 \neq z_3$; $z_2 \neq z_3$⟧ ⟹
`cross_ratio` z z_1 z_2 z_3 = `cross_ratio (mobius_pt` M z`) (mobius_pt` M z_1`)`
`(mobius_pt` M z_2`) (mobius_pt` M z_3`)`"

3.3 Circlines

A very important property of the extended complex plane is that it is possible to treat circles and lines in a uniform way. The basic object is *generalized circle*, or *circline* for short. In our formalization, we follow the approach described by Schwerdtfeger [30] and represent circlines by Hermitian, non-zero 2×2 matrices. In the original formulation, a matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ corresponds to the equation $A * z * \texttt{cnj}\, z + B * \texttt{cnj}\, z + C * z + D = 0$, where $C = \texttt{cnj}\, B$ and A and D are real (as the matrix is Hermitian). The key insight is that this equation represents a line when $A = 0$ or a circle, otherwise.

Again, our formalization proceeds in three stages. First, the type of Hermitian, non-zero matrices is introduced.

definition `is_C2_mat_herm :: "C2_mat ⇒ bool" where`
"`is_C2_mat_herm H` ⟷ `hermitian` H ∧ $H \neq 0$"
typedef `C2_mat_herm = "`{H `:: C2_mat. is_C2_mat_herm` H}"

The representation function `Rep_C2_mat_herm` will be denoted by $\lfloor\_\rfloor_H$, and the abstraction function `Abs_C2_mat_herm` will be denoted by $\lceil\_\rceil^H$. Considering the interpretation in the form of an equation, it is clear that proportional matrices should be considered equivalent. This time matrices are proportional by a real non-zero factor.

definition $\approx_{cm}$ `:: "C2_mat_herm ⇒ C2_mat_herm ⇒ bool" where`
`"`$H_1 \approx_{cm} H_2 \longleftrightarrow$ `(∃ (k::real). k ≠ 0 ∧ ⌊H2⌋_H = cor k *_sm ⌊H1⌋_H)"`

It is easily shown that this is an equivalence relation, and circlines are defined by a quotient construction as its equivalence classes.

quotient_type `circline = C2_mat_herm /` $\approx_{cm}$

An auxiliary constructor `mk_circline` returns a circline (an equivalence class) for given four complex numbers A, B, C and D (provided that they form a Hermitian, non-zero matrix).

Each circline determines a corresponding set of points. Again, a description given in homogeneous coordinates is a bit better than the original description defined only for ordinary complex numbers. The point with homogeneous coordinates (z_1, z_2) will belong to the set of circline points iff $A * z_1 * \texttt{cnj}\, z_1 + B * \texttt{cnj}\, z_1 * z_2 + C * z_1 * \texttt{cnj}\, z_2 + D * z_2 * \texttt{cnj}\, z_2 = 0$. Since this is a quadratic form determined by a vector of homogeneous coordinates and the Hermitian matrix, the set of points on a given circline is formalized as follows (we also here print the definitions of bilinear and quadratic forms, that are introduced in our background theory of linear algebra).

definition `"bilinear_form` $H\ z_1\ z_2$ `= (vec_cnj` z_1`) *_vm H *_vv` z_2`"`
definition `"quad_form` $H\ z$ `= bilinear_form` $H\ z\ z$`"`
definition `on_circline_rep :: "C2_mat_herm ⇒ C2_vec_≠0 ⇒ bool" where`
`"on_circline_rep` $H\ z \longleftrightarrow$ `quad_form ⌊H⌋_H ⌊z⌋_C2 = 0"`
lift_definition `on_circline :: "circline ⇒ complex_hc ⇒ bool" is`
`on_circline_rep`
definition `circline_set :: "complex_hc set" where`
`"circline_set` H `= {`z`. on_circline` $H\ z$`}"`

Lifting the definition of `on_circline` generates the proof obligation $[\![H_1 \approx_{cm} H_2;\ z_1 \approx_{C2} z_2]\!] \Longrightarrow$ `on_circline_rep` $H_1\ z_1 \longleftrightarrow$ `on_circline_rep` $H_2\ z_2$ that is easily discharged.

Some special circlines. Among all circlines most prominent ones are the unit circle, the x-axis, and the imaginary unit circle.

definition `"unit_circle_rep =` $\lceil(1,0,0,-1)\rceil^H$`"`
lift_definition `unit_circle :: "circline" is unit_circle_rep`
definition `"x_axis_rep =` $\lceil(0,i,-i,0)\rceil^H$`"`
lift_definition `x_axis :: "circline" is x_axis_rep`
definition `"imag_unit_circle_rep =` $\lceil(1,0,0,1)\rceil^H$`"`
lift_definition `imag_unit_circle :: "circline" is imag_unit_circle_rep`

It is easy to show some basic properties of these circlines. For example:

lemma `"`$0_{hc} \in$ `circline_set x_axis" "`$1_{hc} \in$ `circline_set x_axis"`
`"`$\infty_{hc} \in$ `circline_set x_axis"`

Connection with lines and circles in ordinary Euclidean plane. In the extended complex plane, there is no difference between the notion of line and circle. However, lines can be defined as those circlines whose matrices have coefficient $A = 0$, or, equivalently as those circlines that contain the point ∞_{hc}.

definition `is_line_rep` **where**
`"is_line_rep` $H \longleftrightarrow$ `(let` (A, B, C, D) `=` $\lfloor H \rfloor_H$ `in` $A = 0$`)"`
lift_definition `is_line :: "circline ⇒ bool"` **is** `is_line_rep`
definition `is_circle_rep` **where**
`"is_circle_rep` $H \longleftrightarrow$ `(let` (A, B, C, D) `=` $\lfloor H \rfloor_H$ `in` $A \neq 0$`)"`
lift_definition `is_circle :: "circline ⇒ bool"` **is** `is_circle_rep`
lemma `"is_line` $H \longleftrightarrow \neg$ `is_circle` H`" "is_line` $H \vee$ `is_circle` H`"`
lemma `"is_line` $H \longleftrightarrow \infty_{hc} \in$ `circline_set` H`"`
`"is_circle` $H \longleftrightarrow \infty_{hc} \notin$ `circline_set` H`"`

Every Euclidean circle and Euclidean line (in the ordinary complex plane, using the standard, Euclidean metric) can be represented by a circline.

definition `mk_circle_rep` μ r `=` $\lceil (1,\ -\mu,\ -$`cnj` $\mu,\ |\mu|^2 - ($`cor` $r)^2) \rceil^H$
lift_definition `mk_circle :: "complex ⇒ real ⇒ circline"` **is** `mk_circle_rep`
lemma `"`$r \geq 0 \Longrightarrow$ `circline_set (mk_circle` μ r`) = of_complex ‘` $\{z.\ |z - \mu| = r\}$`"`
definition `mk_line_rep` **where** `"mk_line_rep` z_1 z_2 `=`
`(let` $B = i * (z_2 - z_1)$ `in` $\lceil (0,\ B,$ `cnj` $B,\ -(B *$ `cnj` $z_1 +$ `cnj` $B * z_1) \rceil^H$`)"`
lift_definition `mk_line :: "complex ⇒ complex ⇒ circline"` **is** `mk_line_rep`
lemma `"`$z_1 \neq z_2 \Longrightarrow$
`circline_set (mk_line` z_1 z_2`) - {`∞_{hc}`} = of_complex ‘` $\{z.$ `collinear` z_1 z_2 $z\}$`"`

The opposite also holds, and the set of points determined by a circline is always either a Euclidean circle or a Euclidean line. For a given circline, the following functions determine the corresponding circle or line parameters (the center and the radius in case of circle or some two different points in case of line).

definition `euclidean_circle_rep` **where** `"euclidean_circle_rep` H `=`
`(let` (A, B, C, D) `=` $\lfloor H \rfloor_H$ `in (`$-B/A$`, sqrt(Re` $((B * C - A * D)/(A * A)))))$`"`
lift_definition `euclidean_circle :: "circline ⇒ complex × real"` **is**
`euclidean_circle_rep`
definition `euclidean_line_rep` **where** `"euclidean_line_rep` H `=`
`(let` (A, B, C, D) `=` $\lfloor H \rfloor_H$`;`
z_1 `=` $-(D * B)/(2 * B * C)$`;`
z_2 `=` z_1 `+` i `*  sgn (if arg` B `> 0 then` $-B$ `else` B`)`
`in (`z_1`,` z_2`))"`
lift_definition `euclidean_line :: "circline ⇒ complex × complex"` **is**
`euclidean_line_rep`

The normal vector of the line is the vector orthogonal to the coefficient B — in order to be able to lift the definition (so that returned points are the same for every circline representative matrix), in the definition of the second point the vector B had to be normalized, giving slightly larger expression than $z_2 = z_1 + i * B$.

Since the cardinality of set of points on the circline depends on the sign of the expression `Re`$((B*C-A*D)/(A*A))$, circlines can be classified into three categories, depending on the sign of the determinant (which is always a real number, since the matrix is Hermitian).

definition circline_type_rep **where**
"circline_type_rep H = sgn (Re (mat_det ($\lfloor H \rfloor_H$)))"
lift_definition circline_type :: "circline ⇒ real" **is** circline_type_rep

The proof obligation $H \approx_{cm} H' \Longrightarrow$ circline_type_rep H = circline_type_rep H' is easy discharged, as Re (mat_det (k $*_{sm}$ H)) = (Re k)^2 $*$ Re (mat_det H) holds for all Hermitian matrices H and all k with imaginary part 0.

Now, it becomes clear that the set of points on the given circline is empty iff the circline type is positive (these are called *imaginary circlines*) , that consists of a single point iff the type is zero (these are called *point circlines*), and that it is infinite iff type type is negative (these are called *real circlines*). Surprisingly, this fact turned out to be very hard to prove formally, and was proved only when Möbius action on circlines was formalized to allow ,,wlog" reasoning. Note that there are no imaginary lines since when $A = 0$, then mat_det $H \geq 0$.

Finally, the connection between real circlines and Euclidean lines and circles can be established.

lemma
assumes "is_circle H" "(μ, r) = euclidean_circle H"
shows "circline_set H = of_complex ` $\{z.\ |z - \mu| = r\}$"
lemma
assumes "is_line H" "(z_1, z_2) = euclidean_line H" "circline_type $H < 0$"
shows "circline_set H - $\{\infty_{hc}\}$ = of_complex ` $\{z.$ collinear z_1 z_2 $z\}$"

Note that the first lemma also holds for point circles and imaginary circles as both sets are empty. However, the second lemma only holds for real lines as in the case of a point line it holds that $z_1 = z_2$, so the left set is empty, but the right is the universal set.

Circlines on the Riemann sphere. Real circlines in the plane correspond to circles on the Riemann sphere, and we have formally established this connection. Every circle in three-dimensional space can be obtained as the intersection of a sphere and a plane. We establish a one-to-one correspondence between circles on the Riemann sphere and planes in space. Note that the plane need not intersect the sphere, but we will still say that it defines some imaginary circle. The correspondence between planes in space and circlines in the extended complex plane has been described by Schwerdtfeger [30]. However, the author failed to note that for one special circline (the one with the identity representative matrix), there does not exist a plane in $\mathbb{R}^3$ that would correspond to it — in order to have this, instead of considering planes in $\mathbb{R}^3$, we must consider three-dimensional projective space and consider the infinite (hyper)plane. Therefore, we define the planes in the following way (again in three stages).

typedef R4_vec$_{\neq 0}$ = "$\{(a, b, c, d)$ :: R4_vec. $(a, b, c, d) \neq 0\}$"

Note that in $\mathbb{R}^3$, one of the numbers a, b, or c would have to be different from 0. However, our definition allows to have the plane $(0, 0, 0, d)$ lying at infinity. The representation function will be denoted by $\lfloor \_ \rfloor_{R4}$, and the abstraction function will be denoted by $\lceil \_ \rceil^{R4}$. Again, two planes are equivalent iff they are proportional (this time by a non-zero real factor).

definition $\approx_{R4}$ :: "R4_vec$_{\neq 0}$ ⇒ R4_vec$_{\neq 0}$ ⇒ bool" **where**
"$\alpha_1 \approx_{R4} \alpha_2 \longleftrightarrow (\exists k.\ k \neq 0 \wedge \lfloor \alpha_2 \rfloor_{R4} = k * \lfloor \alpha_1 \rfloor_{R4})$"

Finally, planes (and circles inside them obtained as intersections with the Riemann sphere) are defined as equivalence classes of this relation.

quotient_type plane = R4_vec$_{\neq 0}$ / $\approx_{R4}$

Plane coefficients give a linear equation and the point on the Riemann sphere lies on the circle determined by the plane iff its representation satisfies that linear equation.

definition on_sphere_circle_rep **where**
"on_sphere_circle_rep α M $\longleftrightarrow$
(let (a, b, c, d) = $\lfloor \alpha \rfloor_{R4}$; ($X$, Y, Z) = $\lfloor M \rfloor_{R3}$
in $a * X + b * Y + c * Z + d = 0$)"
lift_definition on_sphere_circle :: "plane ⇒ riemann_sphere ⇒ bool **is**
on_sphere_circle_rep
definition sphere_circle_set :: "riemann_sphere set" **where**
"sphere_circle_set α = {A. on_sphere_circle α A}"

Note that we did not need to introduce the points in three-dimensional projective space (and their homogeneous coordinates) as we are only interested in the points on the Riemann sphere that are not infinite.

Next, we introduce stereographic and inverse stereographic projection between circles on the Riemann sphere (i.e., the corresponding planes) and circlines in the extended complex plane.

definition stereographic_circline_rep **where**
"stereographic_circline_rep α =
(let (a,b,c,d) = $\lfloor \alpha \rfloor_{R4}$; $A = \mathtt{cor}\,((c+d)/2)$; $B = (\mathtt{cor}\,a + i * \mathtt{cor}\,b)/2)$;
$C = (\mathtt{cor}\,a - i * \mathtt{cor}\,b)/2$; $D = \mathtt{cor}\,((d-c)/2))$
in $\lceil (A,B,C,D) \rceil^{H}$"
lift_definition stereographic_circline :: "plane ⇒ circline" **is**
stereographic_circline_rep
definition inv_stereographic_circline_rep **where**
"inv_stereographic_circline_rep H =
(let (A,B,C,D) = $\lfloor H \rfloor_{H}$
in $\lceil (\mathtt{Re}(B+C), \mathtt{Re}(i*(C-B)), \mathtt{Re}(A-D), \mathtt{Re}(D+A)) \rceil^{R4}$"
lift_definition inv_stereographic_circline :: "circline ⇒ plane" **is**
inv_stereographic_circline_rep

These two mappings are bijective and mutually inverse. The projection of the set of points on a circle on the Riemann sphere is exactly the set of points on the circline obtained by the stereographic projection that we have just defined.

lemma "stereographic_circline ∘ inv_stereographic_circline = id"
lemma "inv_stereographic_circline ∘ stereographic_circline = id"
lemma "bij stereographic_circline" "bij inv_stereographic_circline"
lemma "stereographic ‘ sphere_circle_set α =
circline_set (stereographic_circline α)"

Chordal circlines. Another interesting fact is that real circlines are sets of points that are equidistant from some given points (there are always exactly two of them), but in the chordal metric. On the Riemann sphere these two points (we will call them chordal centers) are obtained as intersections of the sphere and the line that goes through the center of the circle and is normal to the plane that contains the circle.

A chordal circline determined by the given point a and radius r is determined in the following way.

```
definition chordal_circle_rep where "chordal_circle_rep μc rc =
  (let (μ1,μ2) = ⌊μc⌋C2;
       A = 4 * |μ2|² - (cor rc)² * (|μ1|² + |μ2|²); B = -4 * μ1*cnj μ2;
       C = -4*cnj μ1 * μ2; D = 4 * |μ1|² - (cor rc)² * (|μ1|² + |μ2|²)
    in mk_circline_rep A B C D)"
lift_definition chordal_circle :: "complex_hc ⇒ real ⇒ circline" is
  chordal_circle_rep
lemma "z ∈ circline_set (chordal_circle μc rc) ⟷
       rc ≥ 0 ∧ dist_hc z μc = rc"
```

If a circline is given, then its chordal centers and radii can be determined relying on the following lemmas (depending on whether coefficients B and C in the representation matrix are zero).

```
lemma
 assumes "is_C2_mat_herm (A,B,C,D)" "Re (A * D) < 0" "B = 0"
 shows
   "mk_circline A B C D = chordal_circle ∞hc sqrt(Re ((4 * A)/(A - D)))"
   "mk_circline A B C D = chordal_circle 0hc sqrt(Re ((4 * D)/(D - A)))"
lemma assumes
 "is_C2_mat_herm (A,B,C,D)" "Re (mat_det (A,B,C,D)) < 0" "B ≠ 0"
 "C * μc² + (D - A) * μc - B = 0" "rc = sqrt((4 + Re((4 * μc/B) * A))/(1 + Re(|μc|²)))"
 shows "mk_circline A B C D = chordal_circle (of_complex μc) rc"
```

As in the previous cases, the function that returns chordal parameters could be introduced (it would need to distinguish between the cases of $B = 0$ and $B \neq 0$ and in the other case to solve the quadratic equation describing the chordal center).

Symmetry. Since ancient Greeks, the circle inversion was seen as a counterpart of line reflection. In the extended complex plane there are no substantial differences between circles and lines. Therefore, we will consider only one kind of relation and call two points *circline symmetric* if they are mapped to one another using either reflection or inversion over arbitrary line or circle. When, seeking the algebraic characterization of this relation we were a bit surprised how simple and elegant it was – points are symmetric iff the bilinear form of their representation vectors and matrix is zero.

```
definition circline_symmetric_rep where
  "circline_symmetric_rep z1 z2 H ⟷ bilinear_form ⌊z1⌋C2 ⌊z2⌋C2 ⌊H⌋H = 0"
lift_definition circline_symmetric :: "complex_hc ⇒ complex_hc ⇒
  circline ⇒ bool" is circline_symmetric_rep
```

Returning to the set of points on the circline and comparing our two definitions, it becomes clear that points on the circline are exactly those that are invariant under the symmetry of the circline.

lemma "on_circline H z $\longleftrightarrow$ circline_symmetric H z z"

Möbius action on circlines. We have already seen that Möbius transformation act on the points of $\overline{\mathbb{C}}$. They can also act on circlines (and the definition is chosen so that the two actions are compatible). We also print the definition of congruence operation of two matrices (defined in our background theory of linear-algebra).

definition "congruence M H = mat_adj M $*_{mm}$ H $*_{mm}$ M"
definition mobius_circline_rep
  :: "C2_mat_reg ⇒ C2_mat_herm ⇒ C2_mat_herm" **where**
  "mobius_circline_rep M H = ⌈congruence (mat_inv $\lfloor M \rfloor_M$) $\lfloor H \rfloor_H$⌉^{H}"
lift_definition mobius_circline :: "mobius ⇒ circline ⇒ circline" **is**
  mobius_circline_rep

Möbius actions on circlines have similar properties as Möbius actions on points. For example,

lemma "mobius_circline (mobius_comp M_1 M_2) =
    mobius_circline M_1 ∘ mobius_circline M_2"
lemma "mobius_circline (mobius_inv M) = inv (mobius_circline M)"
lemma "mobius_circline (mobius_id) = id"
lemma "inj mobius_circline"

The central lemma in this section connects the action of Möbius transformations on points and on circlines (and shows that the Möbius transformations map circlines to circlines).

lemma "mobius_pt M ‘ circline_set H =
    circline_set (mobius_circline M H)"

Circline type is also preserved (implying, for example, that real circlines are mapped to real circlines).

lemma "circline_type (mobius_circline M H) = circline_type H"

Another important property (a bit more general than the previous one) is that the symmetry of points is preserved by Möbius transformations (the so-called symmetry principle).

lemma assumes "circline_symmetric z_1 z_2 H"
  shows "circline_symmetric (mobius_pt M z_1) (mobius_pt M z_2)
                (mobius_circline M H)"

The last two lemmas are quite prominent geometrical results, and, due to the convenient, algebraic representation they were relatively easy to prove in our formalization. Both proofs rely on the following simple fact of linear algebra.

lemma "mat_det M $\neq$ 0 $\Longrightarrow$ "bilinear_form z_1 z_2 H =
    bilinear_form (M $*_{mv}$ z_1) (M $*_{mv}$ z_2) (congruence (mat_inv M) H)"

Circline uniqueness. In Euclidean geometry, it is a well-known fact that there is a unique line through any two different points and a unique circle through any three different points. Similar results hold in $\overline{\mathbb{C}}$. However, a case-analysis over the type of circlines must be performed. Positive type circlines contain no points, so there are no uniqueness results for them. Zero type circlines consist of a single point and for each point there is a unique zero type circline containing it. There is a unique circline through any three different points (and it must be of a negative type).

lemma "$\exists!\ H.$ `circline_type` H = 0 $\wedge$ $z \in$ `circline_set` H"
lemma "$[\![z_1 \neq z_2;\ z_1 \neq z_3;\ z_2 \neq z_3]\!] \Longrightarrow$
$\exists!\ H.\ z_1 \in$ `circline_set` $H \wedge z_2 \in$ `circline_set` $H \wedge z_3 \in$ `circline_set` H"

Very surprisingly, we did not manage to prove these lemmas directly. However, employing ,,wlog" reasoning and mapping the points to canonical position (0_{hc}, 1_{hc}, and ∞_{hc}) gave us very short and elegant proofs (as it was easy to show computationally that the `x_axis` is the only circline through these three points). As lines are characterized as exactly those circlines that contain ∞_{hc}, so it is clear that there is a unique line through any two finite points.

Circline set cardinality. Another thing usually taken for granted is the cardinality of circlines of different type. We have already said that these proofs required ,,wlog" reasoning, but this time we have used ,,wlog" reasoning of a different kind. In many cases it turns out that it is simpler to reason about circles if their center is in the origin — in those cases, their matrix is diagonal. We have formalized the special case of the famous result of linear algebra claiming that each Hermitian 2x2 matrix is congruent to a real diagonal matrix. Moreover, the elements on the diagonal are the real eigenvalues of the matrix and the congruence is established by a unitary matrix — a congruence could be also established by a simpler, translation matrix, but then it would not have so nice properties.

lemma assumes "`hermitian` H"
shows "$\exists\, k_1\, k_2\, M.$ `mat_det` $M \neq 0 \wedge$ `unitary` $M \wedge$
`congruence` $M\ H$ = (`cor` $k_1, 0, 0,$`cor` k_2)"

The consequence is that for every circline there is a unitary Möbius transformation that transforms it to a position such that its center is in the origin (in fact, there are two such transformations if eigenvalues are different). We shall see that unitary transformations correspond to rotations of the Riemann sphere, so the last fact has a simple geometrical explanation. Circlines could be diagonalized by using translations only, but unitary transformations often have nicer properties.

lemma "$\exists\ M\ H'.$ `unitary_mobius` $M \wedge$
`mobius_circline` $M\ H$ = $H' \wedge$ `circline_diag` H'"
lemma assumes "$\bigwedge H'.$ `circline_diag` $H' \Longrightarrow P\ H$"
"$\bigwedge M\ H.\ P\ H \Longrightarrow P$ (`mobius_circline` $M\ H$)"
shows "$P\ H$"

The predicate `unitary_mobius` lifts the `unitary` condition from $\mathbb{C}^2$ matrices to the `mobius` type. Similarly, `circline_diag` lifts the diagonal matrix condition to the `circline` type.

Using this kind of ,,wlog" reasoning it becomes fairly easy to show the following characterizations of circline set cardinality.

lemma "circline_type H > 0 $\longleftrightarrow$ circline_set H = {}"
lemma "circline_type H = 0 $\longleftrightarrow$ $\exists z$. circline_set H = $\{z\}$"
lemma "circline_type H < 0 $\longleftrightarrow$
$\exists z_1\, z_2\, z_3$. $z_1 \neq z_2 \wedge z_1 \neq z_3 \wedge z_2 \neq z_3 \wedge$ circline_set $H \supseteq \{z_1, z_2, z_3\}$"

An important, non-trivial, consequence of the circline uniqueness and the circline set cardinality is that the function circline_set is injective, i.e., for each non-empty set of points of a circline, there is a unique class of proportional matrices determining it.

lemma "$[\![$ circline_set H_1 = circline_set H_2; circline_set $H_1 \neq \{\}$ $]\!] \Longrightarrow$ $H_1 = H_2$"

The lemma does not hold for the empty set of points as there are many non-equivalent matrices determining it (each imaginary circline has the empty set of points). Although we could have made the definition of circlines that declare all imaginary circlines to be equivalent, our current definition distinguishes different imaginary circlines and gives their finer classification. Looking at the Riemann sphere shows that it is very natural to distinguish different imaginary circlines since they are identified with different planes that do not intersect the sphere.

3.4 Oriented Circlines

In this section we describe how the orientation is introduced for the circlines. Many important concepts depend on the orientation. One of the most important is the concept of *disc* — inside area of a circline. Similarly as the set of circline points, the set of disc points is introduced using the quadratic form induced by the circline matrix — the set of points of the circline disc is the set of points such that satisfy that $A * z * \mathtt{cnj}\, z + B * \mathtt{cnj}\, z + C * z + D < 0$, where $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is a circline matrix representative. Since the set of disc points must be invariant to the choice of representative, it is clear that oriented circlines matrices are equivalent only if they are proportional by a positive real factor (recall that unoriented circline allowed arbitrary non-zero real factors).

definition $\approx_{ocm}$ :: "C2_mat_herm $\Rightarrow$ C2_mat_herm $\Rightarrow$ bool" **where**
"$H_1 \approx_{ocm} H_2 \longleftrightarrow$ ($\exists$ (k::real). $k > 0 \wedge \lfloor H_2 \rfloor_H$ = cor k $*_{sm}$ $\lfloor H_1 \rfloor_H$)"

It is easily shown that this is an equivalence relation, so circlines are defined by a quotient construction as its equivalence classes.

quotient_type o_circline = C2_mat_herm / $\approx_{ocm}$

Now we can use the quadratic forms to define the interior, boundary and the exterior of an oriented circline.

definition on_o_circline_rep :: "C2_mat_herm $\Rightarrow$ C2_vec$_{\neq 0}$ $\Rightarrow$ bool" **where**
"on_o_circline_rep H $z \longleftrightarrow$ quad_form $\lfloor H \rfloor_H$ $\lfloor z \rfloor_{C2}$ = 0"
definition in_o_circline_rep :: "C2_mat_herm $\Rightarrow$ C2_vec$_{\neq 0}$ $\Rightarrow$ bool" **where**
"in_o_circline_rep H $z \longleftrightarrow$ quad_form $\lfloor H \rfloor_H$ $\lfloor z \rfloor_{C2}$ < 0"
definition out_o_circline_rep :: "C2_mat_herm $\Rightarrow$ C2_vec$_{\neq 0}$ $\Rightarrow$ bool" **where**
"out_o_circline_rep H $z \longleftrightarrow$ quad_form $\lfloor H \rfloor_H$ $\lfloor z \rfloor_{C2}$ > 0"

These definitions are then lifted to `on_o_circline`, `in_o_circline`, and `out_o_circline` (proving the necessary obligations), and, finally, the next three definitions are introduced.

definition `o_circline_set :: "complex`$_{hc}$ `set" where`
`"o_circline_set` H `= {`z`. on_o_circline` H z`}"`
definition `disc :: "complex`$_{hc}$ `set" where`
`"disc` H `= {`z`. in_o_circline` H z`}"`
definition `disc_compl :: "complex`$_{hc}$ `set" where`
`"disc_compl` H `= {`z`. out_o_circline` H z`}"`

These three sets are mutually disjoint, and they fill up the entire plane.

lemma `"disc` H $\cap$ `disc_compl` H `= {}"`
`"disc` H $\cap$ `o_circline_set` H `= {}"`
`"disc_compl` H $\cap$ `o_circline_set` H `= {}"`
`"disc` H $\cup$ `disc_compl` H $\cup$ `o_circline_set` H `= UNIV"`

Given an oriented circline, one can trivially obtain its unoriented counterpart, and these two share the same set of points.

lift_definition `of_o_circline (_`$^{\bigcirc}$`) :: "o_circline` $\Rightarrow$ `circline" is` id
lemma `"circline_set (`$H^{\bigcirc}$`) = o_circline_set` H`"`

Note that in the previous **lift_definition** we have introduced the superscript notation for the function `of_o_circline`, so, for example, $H^{\bigcirc}$ in the lemma is a shorthand for `of_o_circline` H.

For each circline, there is exactly one opposite oriented circline

definition `"opp_o_circline_rep` H `=` $\lceil -1 *_{sm} \lfloor H \rfloor_H \rceil^H$`"`
lift_definition `opp_o_circline (_`$^{\leftrightarrow}$`) :: "o_circline` $\Rightarrow$ `o_circline" is`
`opp_o_circline_rep`

Finding opposite circline is idempotent, and opposite circlines share the same set of points, but exchange disc and its complement.

lemma `"(`$H^{\leftrightarrow}$`)`$^{\leftrightarrow}$ $= H$`"`
lemma `"o_circline_set (`$H^{\leftrightarrow}$`) = o_circline_set` H`"`
`"disc (`$H^{\leftrightarrow}$`) = disc_compl` H`" "disc_compl (`$H^{\leftrightarrow}$`) = disc` H`"`

The functions `_`$^{\bigcirc}$ and `o_circline_set` are injective in some sense.

lemma `"`$H_1{}^{\bigcirc} = H_2{}^{\bigcirc} \Longrightarrow H_1$ `= ` $H_2 \vee H_1$ `= `$H_2{}^{\leftrightarrow}$`"`
lemma `"`$[\![$`o_circline_set` H_1 `= o_circline_set` H_2`;` `o_circline_set` $H_1 \neq \{\}]\!] \Longrightarrow$
$H_1 = H_2 \vee H_1 = H_2{}^{\leftrightarrow}$`"`

Given a representative Hermitian matrix of a circline, it represents exactly one of the two possible oriented circlines. The choice of what should be called a positive orientation is arbitrary. We follow Schwerdtfeger [30], use the leading coefficient A as the first criterion, and say that circline matrices with $A > 0$ are called positively oriented, and with $A < 0$ negatively oriented. However, Schwerdtfeger did not discuss the possible case of $A = 0$ (the case of lines), so we had to extend his definition to achieve a total characterization.

```
definition "pos_o_circline_rep where "pos_o_circline_rep H ⟷
   (let (A, B, C, D) = ⌊H⌋_H
     in Re A > 0 ∨
        (Re A = 0 ∧ ((B ≠ 0 ∧ arg B > 0) ∨ (B = 0 ∧ Re D > 0))))"
lift_definition pos_o_circline :: "o_circline ⇒ bool" is pos_o_circline_rep
```

Now, exactly one of the two oppositely oriented circlines is positively oriented.

```
lemma "pos_o_circline H ∨ pos_o_circline (H^↔)"
      "pos_o_circline (H^↔) ⟷ ¬ pos_o_circline H"
```

The orientation of circles is both algebraically simple (the sign of the coefficient A) and geometrically natural, due to the following simple characterization.

```
lemma "∞_h ∉ o_circline_set H ⟹ pos_o_circline H ⟷ ∞_h ∉ disc H"
```

Another nice geometric characterization of positive orientation is that the positively oriented Euclidean circles contain their Euclidean centers in the disc.

```
lemma assumes "is_circle (H^○)" "circline_type (H^○) < 0"
              "(μ, r) = euclidean_circle (H^○)"
      shows "pos_oriented H ⟷ of_complex μ ∈ disc H"
```

Note that the orientation of lines and point circles is artificially introduced (only to have a total positive orientation characterization), and it does not have a natural geometric interpretation. This breaks the continuity of orientation, and we think that it is not possible to introduce the orientation of lines, so that the orientation function becomes everywhere continuous. Therefore, in most lemmas that tell something about the orientation we will explicitly exclude the case of lines.

Having a total characterization for the positive orientation allows to create a coercion from an unoriented to an oriented circline (returning always the positively oriented circline).

```
definition of_circline_rep :: "C2_mat_herm ⇒ C2_mat_herm" where
   "of_circline_rep H = (if pos_o_circline_rep H then H
                         else opp_o_circline_rep H)"
lift_definition of_circline (_^◌) :: "circline ⇒ o_circline" is of_circline_rep
```

There are many elementary properties of the function `of_circline` proved, and here we list some most important.

```
lemma "o_circline_set (H^◌) = circline_set H"
lemma "pos_o_circline (H^◌)"
lemma "(H^◌)^○ = H" "pos_o_circline H ⟹ (H^○)^◌ = H"
lemma "H1^◌ = H2^◌ ⟹ H1 = H2"
```

Möbius action on oriented circlines. On the representation level, the Möbius action on an oriented circline is the same as on an unoriented circline.

```
lift_definition mobius_o_circline :: "mobius ⇒ o_circline ⇒ o_circline" is
   mobius_circline_rep
```

Möbius action on (unoriented) circlines could have been defined using the action on oriented circlines, but not the other way around.

lemma "mobius_circline M H = (mobius_o_circline M $(H^{\circlearrowleft})$)$^{\circlearrowright}$"
lemma "let H_1 = mobius_o_circline M H; H_2 = (mobius_circline M $(H^{\circlearrowright})$)$^{\circlearrowleft}$
in H_1 = H_2 ∨ H_1 = $H_2{}^{\leftrightarrow}$"

Möbius actions on oriented circlines have similar properties as Möbius actions on unoriented ones. For example, they agree with inverse (**lemma** "mobius_o_circline (mobius_inv M) = inv (mobius_o_circline M)"), with composition, identity transformation, they are injective (inj mobius_circline), and so on. The central lemmas in this section connects the action of Möbius transformations on points, on oriented circlines, and discs.

lemma "mobius_pt M ‘ o_circline_set H =
o_circline_set (mobius_o_circline M H)"
lemma "mobius_pt M ‘ disc H = disc (mobius_o_circline M H)"
lemma "mobius_pt M ‘ disc_compl H = disc_compl (mobius_o_circline M H)"

All Euclidean similarities preserve circline orientation.

lemma assumes "$a \neq 0$" "M = similarity a b" "$\infty_{hc} \notin$ o_circline_set H"
shows "pos_o_circline H ⟷ pos_o_circline (mobius_o_circline M H)"

Orientation of the image of a given oriented circline H under a given Möbius transformation M depends on whether the pole of M (the point that M maps to ∞_{hc}) lies in the disc or in the disc complement of H (if the pole lies on the circline H, then the circline maps onto a line and we do not discuss the orientation).

lemma
"0_{hc} ∈ disc_compl H ⟹ pos_o_circline (mobius_o_circline reciprocation H)"
"0_{hc} ∈ disc H ⟹ ¬ pos_o_circline (mobius_o_circline reciprocation H)"
lemma
assumes "M = mk_mobius a b c d" "$c \neq 0$" "$a*d - b*c \neq 0$"
shows "pole M ∈ disc H ⟶ ¬ pos_o_circline (mobius_o_circline M H)"
"pole M ∈ disc_compl H ⟶ pos_o_circline (mobius_o_circline M H)"

Note that this is different to what is claimed by Schwerdtfeger [30]: ,,Reciprocation preserves the orientation of a circle which does not contain 0, but inverts the orientation of any circle containing 0 as an interior point. Every Möbius transformation preserves the orientation of any circle that does not contain its pole. If circle contains its pole, then the image circle has its orientation opposite.”. Our formalization shows that the orientation of the image circle does not depend on the orientation of the initial one. For example, in the case of reciprocation, the orientation of the initial circle depends only on the sign of the coefficient A in a representation matrix (i.e., on the relationship between the circle disc and the infinite point). On the other hand, since reciprocation exchanges coefficients A and D, the orientation of the image circle depends only on the sign of the coefficient D (i.e., on the relationship between the initial circle disc and the point zero — the pole of reciprocation). The coefficients A and D are totally independent, so the orientation of the image does not depend on the orientation of the initial circle.

Angle preservation. Möbius transformations are conformal, meaning that they preserve oriented angle between oriented circlines. If angle is defined in purely algebraic terms (following Schwerdtfeger [30]), then this property is a very easy to prove. We also print the definition of a mixed determinant defined in our background theory of linear algebra.

```
fun mat_det_mix :: "C2_mat ⇒ C2_mat ⇒ complex" where
  "mat_det_mix (A1, B1, C1, D1) (A2, B2, C2, D2) =
     A1 * D2 − B1 * C2 + A2 * D1 − B2 * C1"
definition cos_angle_rep where
  "cos_angle_rep H1 H2 = − Re (mat_det_mix ⌊H1⌋H ⌊H2⌋H) /
                           2 * (sqrt (Re (mat_det ⌊H1⌋H * mat_det ⌊H2⌋H)))"
lift_definition cos_angle :: "o_circline ⇒ o_circline ⇒ complex" is
  cos_angle_rep
lemma "cos_angle H1 H2 =
  cos_angle (moebius_o_circline M H1) (moebius_o_circline M H2)"
```

However, this definition is not intuitive, and for pedagogical reasons we want to connect it to the more common definition. First, we define the angle between two complex vectors ($\lfloor \_ \rfloor$ denotes the angle canonization function described earlier).

```
definition ang_vec ("∡") where "∡ z1 z2 = ⌊arg z2 - arg z1⌋"
```

Given a center μ of an ordinary Euclidean circle and a point z on it, we define the tangent vector in z as the radius vector $\overrightarrow{\mu z}$, rotated by $\pi/2$, clockwise or counterclockwise, depending on the circle orientation.

```
definition tang_vec :: "complex ⇒ complex ⇒ bool ⇒ complex" where
   "tang_vec μ z p = sgn_bool p * i * (z − μ)"
```

The Boolean p encodes the orientation of the circle, and the function `sgn_bool` p returns 1 when p is true, and -1 for when p is false. Finally, angle between two oriented circles at their common point z is defined as the angle between tangent vectors at z.

```
definition ang_circ where
   "ang_circ z μ1 μ2 p1 p2 = ∡ (tang_vec μ1 z p1) (tang_vec μ2 z p2)"
```

Finally, the connection between algebraic and geometric definition of angle cosine is given by the following lemma.

```
lemma assumes "is_circle (H1○)" "is_circle (H2○)"
      "circline_type (H1○) < 0" "circline_type (H2○) < 0"
      "(μ1, r1) = euclidean_circle (H1○)"
      "(μ2, r2) = euclidean_circle (H2○)"
      "of_complex z ∈ o_circline_set H1 ∩ o_circline_set H2"
   shows "cos_angle H1 H2 =
          cos (ang_circ z μ1 μ2 (pos_o_circline H1) (pos_o_circline H2))"
```

To prove this lemma we needed to show the law of cosines in Isabelle/HOL, but it turned out to be a rather simple task.

3.5 Some Important Subgroups of Möbius Transformations

We have already described the parabolic group (the group of Euclidean similarities), crucial for the Euclidean plane geometry. Now we will describe characterizations of two very important subgroups of the Möbius group — the group of sphere rotations, important for the elliptic plane geometry, and the group of disc automorphisms, important for the hyperbolic plane geometry.

Sphere rotations. General unitary group, denoted by $GU_2(\mathbb{C})$ is the group that contains all Möbius transformations represented by generalized unitary matrices.

definition `unitary_gen` **where**
`"unitary_gen` $M \longleftrightarrow (\exists\ k$`::complex.` $k \neq 0 \wedge$ `mat_adj` $M *_{mm} M$ `=` $k *_{sm}$ `eye)"`

Although the definition allows any complex factor k, it turns out that k can only be real. Generalized unitary matrices can be factored into ordinary unitary matrices and positive multiples of the identity matrix.

definition `unitary` **where** `"unitary` $M \longleftrightarrow$ `mat_adj` $M *_{mm} M$ `= eye"`
lemma `"unitary_gen` $M \longleftrightarrow$
$(\exists\ k\ M'.\ k > 0 \wedge$ `unitary` $M' \wedge M =$ `(cor` $k *_{sm}$ `eye)` $*_{mm}\ M')$`"`

The group of unitary matrices is very important as it describes all rotations of the Riemann sphere (it is isomorphic to the real special orthogonal group $SO_3(\mathbb{R})$). One characterization of $GU_2(\mathbb{C})$ in $\overline{\mathbb{C}}$ is that it is a group of transformations that leave the imaginary unit circle fixed (this is the circle with the identity representation matrix, contained in the plane at infinity).

lemma `"mat_det` $(A, B, C, D) \neq 0 \Longrightarrow$ `unitary_gen` $(A,\ B,\ C,\ D) \longleftrightarrow$
`moebius_circline (mk_moebius` $A\ B\ C\ D$`) imag_unit_circle =`
`imag_unit_circle"`

The characterization of generalized unitary matrices in coordinates is given by the following lemma.

lemma `"unitary_gen M` $\longleftrightarrow$ $(\exists\ a\ b\ k.$ `let` $M' = (a,\ b,\ -$`cnj` $b,$ `cnj` $a)$ `in`
$k \neq 0 \wedge$ `mat_det` $M' \neq 0 \wedge M = k *_{sm} M')$`"`

Along the way we have also defined the special unitary group $SU_2(\mathbb{C})$, containing generalized unitary matrices with unit determinant. They are recognized by the form $(a,\ b,\ -$`cnj` $b,$ `cnj` $a)$, without the multiple k, and we used this to derive the coordinate form of generalized unitary matrices.

Disc automorphisms. A dual group to the previous one is the group of generalized unitary matrices with the $1-1$ signature ($GU_{1,1}(\mathbb{C})$).

definition `unitary11` **where**
`"unitary11` $M \longleftrightarrow$ `mat_adj` $M *_{mm} (1, 0, 0, -1) *_{mm} M = (1, 0, 0, -1)$`"`
definition `unitary11_gen` **where**
`"unitary11_gen` $M \longleftrightarrow (\exists\ k$`::complex.` $k \neq 0 \wedge$
`mat_adj` $M *_{mm} (1, 0, 0, -1) *_{mm} M =$ `k` $*_{sm}\ (1, 0, 0, -1))$`"`

Again, the definition allows a complex factor k, but it is shown that only real factors are plausible.

A characterization of the $GU_{1,1}(\mathbb{C})$ is that it contains all Möbius transformations that leave the unit circle fixed.

lemma `"mat_det` $(A, B, C, D) \neq 0 \Longrightarrow$ `unitary11_gen` $(A, B, C, D) \longleftrightarrow$
`moebius_circline (mk_moebius` $A\ B\ C\ D$`) unit_circle = unit_circle"`

The characterization of generalized unitary 1-1 matrices in coordinates is given by the following lemmas.

lemma "unitary11_gen $M \longleftrightarrow$ ($\exists$ a b k. let $M' = (a, b,$ cnj $b,$ cnj $a)$ in
$k \neq 0 \wedge$ mat_det $M' \neq 0 \wedge (M = k *_{sm} M' \vee M = k *_{sm}$ (cis pi, 0, 0, 1) $*_{sm} M'))$
lemma "unitary11_gen M $\longleftrightarrow$ ($\exists$ a b k. let $M' = (a, b,$ cnj $b,$ cnj $a)$ in
$k \neq 0 \wedge$ mat_det $M' \neq 0 \wedge M = k *_{sm} M'$)"

Note that the first lemma is subsumed by the second one. However, the first lemma was simpler to prove, and gives matrices of another shape $k *_{sm} (a, b, -$cnj $b, -$cnj $a)$ — geometrically, the second kind of transformation combines the first kind with an additional central symmetry.

Another important characterization of these transformations is via so-called *Blaschke factors*. Each transformation is a composition of a Blaschke factor (a reflection that brings some point that is not on the unit circle to zero), and a rotation.

lemma assumes "$k \neq 0$" "$M' = (a, b,$ cnj $b,$ cnj $a)$"
"$M = k *_{sm} M'$" "mat_det $M' \neq 0$" "$a \neq 0$"
shows "$\exists$ k' ϕ a'. $k' \neq 0 \wedge a' *$ cnj $a' \neq 1 \wedge$
$M = k' *_{sm}$ (cis ϕ, 0, 0, 1) $*_{mm}$ (1, $-a',$ $-$cnj a', 1)"

The exceptions come when $a = 0$ and then instead of the Blaschke factor, a reciprocation is used (the infinity plays the role of a' in the previous lemma).

lemma assumes "$k \neq 0$" "$M' = (0, b,$ cnj $b, 0)$" "$b \neq 0$" "$M = k *_{sm} M'$"
shows "$\exists$ k' ϕ. $k' \neq 0 \wedge M = k' *_{sm}$ (cis ϕ, 0, 0, 1) $*_{mm}$ (0, 1, 1, 0)"

Matrices of $GU_{1,1}(\mathbb{C})$ naturally split into two subgroups. All transformations fix the unit circle, but the first subgroup consists of transformations that map the unit disc to itself (so-called *disc automorphisms*), while the second subgroup consists of transformations that exchange the unit disc and its complement. Given a matrix, its subgroup can be determined only on by looking at the sign of the determinant of $M' = (a, b,$ cnj $b,$ cnj $a)$. If only $M = (a_1, b_1, c_1, d_1)$, and not M' nor k is given, a criterion to determine the subgroup is the value of sgn(Re $((a_1 * d_1)/(b_1 * c_1)) - 1)$.

Note that all the important subgroups are here described only in pure algebraic terms. We have also formalized some more geometric proofs resulting in equivalent characterization to these we have just described. Additionally, it holds that all analytic disc automorphisms are compositions of Blaschke factors and rotations (however, the proofs relies on mathematical analysis, maximum modulus principle, and the Swartz lemma — techniques that we did not consider). Even the weaker statement claiming that all Möbius disc automorphisms are of this form has not yet been formally proved. The crucial step is showing that disc automorphisms fix the unit circle, and that is something that we did not manage to do without deep topological investigations that we are currently working on.

4 Discussion

When developing a formal library, an important question is how to define the notions and formulate their properties so that the proofs become shorter and so that their large parts can be automated. In this section, we present an example that demonstrates how complicated it is to formalize a proof when notions are defined naively, by following classical mathematical material. Although proofs that

are only semi-formal and that fail to discuss some corner cases are inherently problematic, the problems often become evident only when one tries to formalize them within a proof assistant. Namely, readers sometimes do not care much about corner cases and are usually satisfied with such proofs because they are simple and intuitive.

In the current example, we will consider one classic definition of angle between circles and then analyze one classic proof of the angle preservation property of Möbius transformations that is often encountered in textbooks on the subject (in the rest of this section we will follow Needham [26] which does not aim to be a very formal book, but, still, that kind of reasoning is common for many other authors). Comparison to our purely algebraic definition of angle and the corresponding proof of the angle preservation property given in Section 3.4 reveals that the classic proof lacks formality and, therefore, it is very hard to formalize it within a proof assistant. On the other hand, the classic proof offers more intuition and better understanding (as it can be visualized).

Angles can be defined between oriented, or unoriented curves and angles themselves can be oriented or unoriented. Needham defines angles between two curves in the following way: *,,Let S_1 and S_2 be curves intersecting at z. As illustrated, we may draw their tangent lines T_1 and T_2 at z. The angle between curves S_1 and S_2 at their common point z is the acute angle α from T_1 to T_2. Thus this angle α has a sign attached to it: the angle between S_2 and S_1 is minus the illustrated angle between S_1 and S_2."* So, the angle is defined only between unoriented curves (and that is different from our definition given in Section 3.4), but the angle itself is oriented (and that is the same as in our definition). We first define the unoriented convex and the acute angle between two vectors.

definition "$\measuredangle_c$" **where** "$\measuredangle_c\ z_1\ z_2 \equiv$ `abs` $(\measuredangle\ z_1\ z_2)$"
definition `acutize` **where** "`acutize` α `=` (`if` $\alpha > \frac{\pi}{2}$ `then` π `-` α `else` α)"
definition "$\measuredangle_a$" **where** "$\measuredangle_a\ z_1\ z_2 \equiv$ `acutize` $(\measuredangle_c\ z_1\ z_2)$"

The function `ang_circ_a` is defined as the acute angle between the two tangent vectors of two intersecting circles (it is similar to `ang_circ`, but the returned angle must always be acute). As our circles are oriented, we have shown that the acute angle between the two circles is not affected by the orientation and can only be expressed in terms of three points (the intersection point and the two centers).

lemma "$[\![z \neq \mu_1; z \neq \mu_2]\!] \Longrightarrow$ `ang_circ_a` $z\ \mu_1\ \mu_2\ p_1\ p_2$ `=` $\measuredangle_a\ (z-\mu_1)\ (z-\mu_2)$"

The angle preservation proof for Möbius transformations in the textbook [26] relies on the fact that each Möbius transformation can be decomposed to translations, rotation, dilatation, and inversion. The fact that translations, rotations, and dilatations preserve angles is taken for granted (and, to be honest, formalizing this was rather simple, once the underlying notions were defined appropriately). Therefore, the central challenge is to show that inversion preserves angles, i.e., that *,,inversion in a circle is an anticonformal mapping"*. The proof relies on the *,,fact that given any point z not on the inversion circle K, there is precisely one circle orthogonal to K that passes through z in any given direction"*. Then the proof proceeds *,,Suppose that two curves S_1 and S_2 intersect at z, and that their tangents there are T_1 and T_2, the angle between them being α. To find out what happens to this angle under inversion in K, let us replace S_1 and S_2 with the unique circles R_1 and R_2 orthogonal*

to K that pass through z in the same directions as directions S_1 and S_2, i.e., circles whose tangents at z are T_1 and T_2. Since inversion in K maps each of these circles to themselves, the new angle at $\tilde{z}$ is $-\alpha$. Done."

We have formalized this ,,proof", but this required tremendous amount of effort, compared to the sleek algebraic proof described in Section 3.4. First, the textbook is often imprecise in whether it deals with ,,complex inversion" or ,,geometric inversion" (i.e., between the reciprocation and the inversion put in our terms). In the textbook proof, the author uses inversion over any circle K, but it is sufficient to consider only the reciprocation (always given over the unit circle). Formalizing the textbook reasoning only for the reciprocation already gave quite large formulas, and it would be even more complicated and tedious (if not impossible) to finish the proof using inversion over arbitrary circle. For example, a simple reciprocation of a circle with a center μ and radius r gives a circle with the center $\tilde{\mu} = \mu/\texttt{cor}\ (|\mu|^2 - r^2)$, and radius $\tilde{r} = r/||\mu|^2 - r^2|$, and this relationship would be much more complex for an arbitrary Möbius transformation, if it was written in coordinates, without using matrix notation.

The formal angle preservation statement is the following (`circle` μ r denotes the set $\{z.\ |z - \mu| = r\}$, μ_1, μ_2, $\tilde{\mu_1}$, $\tilde{\mu_2}$, z, $\tilde{z}$ are complex numbers, and r_1, r_2, $\tilde{r_1}$, and $\tilde{r_2}$ are real numbers).

lemma
assumes "$z \in$ `circle` $\mu_1\ r_1$" "$z \in$ `circle` $\mu_2\ r_2$"
"`reciprocation ‘ circle` $\mu_1\ r_1$ `= circle` $\tilde{\mu_1}\ \tilde{r_1}$"
"`reciprocation ‘ circle` $\mu_2\ r_2$ `= circle` $\tilde{\mu_2}\ \tilde{r_2}$"
"$\tilde{z} \in$ `circle` $\tilde{\mu_1}\ \tilde{r_1}$" "$\tilde{z} \in$ `circle` $\tilde{\mu_2}\ \tilde{r_2}$"
shows "`ang_circ_a` $z\ \mu_1\ \mu_2$ `=` `ang_circ_a` $\tilde{z}\ \tilde{\mu_1}\ \tilde{\mu_2}$"

Apart from missing discussion of many special cases, the informal proof misses one key ingredient. Namely, it is easy to prove that the intersection of R_1 and R_2 is $\tilde{z}$ (the intersection of $\tilde{S_1}$ and $\tilde{S_2}$, the images of S_1 and S_2 under inversion), but showing that R_1 and $\tilde{S_1}$ and that R_2 and $\tilde{S_2}$ share tangents at $\tilde{z}$ required not so trivial calculations (that proof relies on the fact that center μ_i' of R_i, the center $\tilde{\mu_i}$ of $\tilde{S_i}$, and $\tilde{z}$ are collinear).

Simple symmetry argument showing that the angles between the circles in their two different intersection points are the same was again not so simple to formalize.

lemma assumes "$\mu_1 \neq \mu_2$" "$r_1 > 0$" "$r_2 > 0$"
"$\{z_1, z_2\} \subseteq$ `circle` $\mu_1\ r_1 \cap$ `circle` $\mu_2\ r_2$" "$z_1 \neq z_2$"
shows "`ang_circ_a` $z_1\ \mu_1\ \mu_2$ `=` `ang_circ_a` $z_2\ \mu_1\ \mu_2$"

We have shown this lemma only after employing ,,wlog" reasoning and moving the configuration so that the centers of the two circles are on the x-axis.

In the proof, we have found many degenerate cases that had to be analyzed separately. First, we had to prove that intersecting circles can share the same center (i.e., that $\mu_1 \neq \mu_2$) only if they are the same, and then the acute angle between tangents is 0. If the two centers are collinear with the intersection point z (i.e., if `collinear` $\mu_1\ \mu_2\ z$ holds), the two circles touch (either from inside or from the outside), and again the acute angle is 0.

Existence of the circle R_i orthogonal to the unit circle, sharing the same tangent in the given point z with the given circle centered in the given point μ_i is given

by the following lemma (`ortho_unit_circ` denotes the set of points on the circle centered in μ_i', orthogonal to unit circle).

lemma
  assumes "$\langle \mu_i$ - z, $z\rangle \neq 0$"
    "μ_i' = $z + (1 - z*\texttt{cnj}\ z) * (\mu_i - z)/(2 * \langle \mu_i - z, z\rangle)$"
  shows "collinear z μ_i μ_i'" "$z \in$ ortho_unit_circ μ_i'"

The analytic expressions reveal some other degenerate cases. The numerator of the fraction can be zero only when the circles intersect on the unit circles (i.e., when $z * \texttt{cnj}\ z = 1$). In that case, the textbook proof cannot be adapted, as $\mu_1' = \mu_2' = z$, and the circles R_1 and R_2 cannot not be constructed (they are the empty circles). The case when denominator is zero (either for μ_1' or μ_2') is also degenerate. That happens when vectors $\mu_i - z$ and z are orthogonal. Geometrically, in that case the circle R_i degenerates into a line (what is not a problem in the extended complex plane, but is a problem in the original proof set in the ordinary complex plane). Therefore, this special case had to be handled separately. So, our formal analysis quickly shows that the simple statement in Needham that ,,given any point z not on the inversion circle K, there is precisely one circle orthogonal to K that passes through z in any given direction" is not true in many cases.

Problems demonstrated in this example occur in many other proofs in the classic literature on the subject, showing that our choice of purely algebraic definitions in our formalization was an extremely important step to keep the formalization simple and the proofs short.

5 Conclusions and Further Work

In this paper, we have described some elements of our formalization of the geometry of the complex plane $\overline{\mathbb{C}}$ both as complex projective line and the Riemann sphere, arithmetic operations in $\overline{\mathbb{C}}$, ratio and cross-ratio, chordal metric in $\overline{\mathbb{C}}$, the group of Möbius transformations and their action on $\overline{\mathbb{C}}$, some of its special subgroups (Euclidean similarities, sphere rotations, disk automorphisms), circlines and their connection with circles and lines, the chordal metric, and the Riemann sphere, Möbius action of circlines, circline uniqueness, circline types and set cardinality, oriented circlines, relations between Möbius transformations and the orientation, angle preservation properties of Möbius transformations, etc. Our current development counts around 12,000 lines of Isabelle/HOL code (all proofs are structured and written in the proof language Isabelle/Isar, and our early attempts that are subsumed by shorter algebraic proofs are not included), around 125 definitions and around 800 lemmas.

The crucial step in our formalization was our decision to use the algebraic representation of all relevant objects (e.g., vectors of homogeneous coordinates, matrices for Möbius transformations, Hermitian matrices for circlines). Although this is not a new approach (for example, Schwerdtfeger's classic book [30] follows this approach quite consistently), it is not so common in the literature (and in the course material available online). Instead, other, more geometrically oriented approaches prevail. We have tried to follow that kind of geometric reasoning in our early work on this subject, but we have encountered many difficulties and did not

have so much success. Based on this experience, we conclude that introducing the powerful techniques of linear algebra makes the work on formalization an order of magnitude simpler than when using just plain geometric reasoning.

It can be argued that sometimes geometrical arguments give better explanations of some theorems, but when only justification is concerned, the algebraic approach is clearly superior. However, to keep the connection with the standard, geometric intuition, several definitions must be introduced (more geometric, and more algebraic ones), and they must be proved equivalent. For example, when the definition of angles is given only through algebraic operations on matrices and their determinants, the angle preservation property is very easy to prove. However, for educational purposes this becomes relevant only when that definition is connected with the standard definition of angle between curves (i.e., their tangent vectors), or, otherwise, the formalization becomes a game with meaningless symbols.

Another important conclusion that we make is that in formal documents, case analysis should be avoided and extensions that help avoiding it should be pursued whenever possible (e.g., it was much better to use the homogeneous coordinates instead of a single distinguished infinity point, it was much simpler to work with circlines than to distinguish between circles and lines, etc.). Keeping different models of the same concept (for example, homogeneous coordinates and the Riemann sphere) also helps, as some proofs are easier in one, and some proofs are easier in other models.

In principle, our proofs are not long (15-20 lines in average with each Isar statement in a separate line). However, some tedious reasoning was sometimes required, especially when switching between real and complex numbers (by the conversion functions `Re` and `cor`). These conversions are usually not present in informal texts, and some better automation of reasoning about them would be welcome. Isabelle's automation was quite powerful in equational reasoning about ordinary complex numbers using `(simp add: field_simps)` (with some minor exceptions). However, the automation was not so good in the presence of inequalities and we had to manually prove many things that would be considered trivial in informal texts.

Since in our formalization quotients are intensively used, porting it to some other prover would require that the prover has good support for them. Introducing the concepts using quotients is a natural operation in Isabelle/HOL and other HOL provers, but might pose challenges to other provers. For example, in an intensional type theory it is not always possible to form the quotient of a type by an equivalence relation. In Coq, the problem with quotient types is that there is no general way of forming them, without axioms. One approach considers quotients of setoids [1] — setoid is a type with an equivalence relation called setiod equality and quotienting a setiod amounts to changing the setoid equality to a broader one. A pragmatic approach to quotients in Coq/SSReflect is recently described by Cohen [2]. Besides quotients, when porting our formalization to other provers, library support for complex numbers, trigonometric functions and abstract algebra (automated reasoning in fields, groups, vector and metric spaces, etc.) would be very welcome.

In our further work we plan to use these results for formalizing non-Euclidean geometries and their models (especially, spherical model of the elliptic geometry and the Poincaré disc and upper half-plane models of hyperbolic geometry). We also plan to generalize our linear algebraic results to arbitrary dimensions, as such library could be useful in other contexts.

Acknowledgements The authors are grateful to Pascal Schreck, Pierre Boutry, Julien Narboux, and anonymous reviewers of AMAI journal for many valuable suggestions and advice.

References

1. Gilles Barthe, Venanzio Capretta, and Olivier Pons. Setoids in type theory. *Journal of Functional Programming*, 13(2):261–293, 2003.
2. Cyril Cohen. Pragmatic quotient types in Coq. In Sandrine Blazy, Christine Paulin-Mohring, and David Pichardie, editors, *Interactive Theorem Proving*, volume 7998 of *Lecture Notes in Computer Science*, pages 213–228. Springer Berlin Heidelberg, 2013.
3. Christophe Dehlinger, Jean-François Dufourd, and Pascal Schreck. Higher-Order Intuitionistic Formalization and Proofs in Hilberts Elementary Geometry. In *Automated Deduction in Geometry*, volume 2061 of *Lecture Notes in Computer Science*. Springer, 2001.
4. Jean Duprat. Constructors: a ruler and a pair of compasses. In *TYPES 2002*. 2002.
5. Jean-David Génevaux, Julien Narboux, and Pascal Schreck. Formalization of Wu's simple method in Coq. In *CPP*, volume 7086 of *Lecture Notes in Computer Science*. Springer, 2011.
6. Herman Geuvers, Freek Wiedijk, and Jan Zwanenburg. A Constructive Proof of the Fundamental Theorem of Algebra without Using the Rationals. In *Types for Proofs and Programs*, volume 2277 of *Lecture Notes in Computer Science*. Springer, 2002.
7. Benjamin Grégoire, Loïc Pottier, and Laurent Théry. Proof certificates for algebra and their application to automatic geometry theorem proving. In *Automated Deduction in Geometry*, volume 6301 of *Lecture Notes in Computer Science*. Springer, 2008.
8. Frédérique Guilhot. Formalisation en Coq et visualisation d'un cours de géométrie pour le lycée. *Technique et Science Informatiques*, 24(9), 2005.
9. Florian Haftmann and Makarius Wenzel. Constructive type classes in isabelle. In *Types for Proofs and Programs*, volume 4502 of *Lecture Notes in Computer Science*, pages 160–174. Springer Berlin Heidelberg, 2007.
10. John Harrison. A HOL Theory of Euclidean Space. In *TPHOLs*, volume 3603 of *Lecture Notes in Computer Science*. Springer, 2005.
11. John Harrison. Without loss of generality. In *Proceedings of the 22nd International Conference on Theorem Proving in Higher Order Logics, TPHOLs 2009*, volume 5674 of *LNCS*, Munich, Germany, 2009. Springer-Verlag.
12. John Harrison. The HOL Light Theory of Euclidean Space. *J. Autom. Reasoning*, 50(2), 2013.
13. David Hilbert and E.J. Townsend. *The Foundations Of Geometry*. Kessinger Publishing, 2006.
14. Einar Hille. *Analytic Function Theory*. AMS Chelsea Publishing. American Mathematical Society, 1973.
15. Brian Huffman and Ondřej Kunčar. Lifting and Transfer: A Modular Design for Quotients in Isabelle/HOL. In *Certified Programs and Proofs*, volume 8307 of *LNCS*. Springer International Publishing, 2013.
16. Predrag Janičić, Julien Narboux, and Pedro Quaresma. The Area Method. *Journal of Automated Reasoning*, 48(4), 2012.
17. Gilles Kahn. Constructive geometry according to Jan von Plato. *Coq contribution, Coq V5.10*, 1995.
18. Cezary Kaliszyk and Christian Urban. Quotients revisited for Isabelle/HOL. In *Proceedings of the 2011 ACM Symposium on Applied Computing*, SAC '11, pages 1639–1644, New York, NY, USA, 2011. ACM.
19. Nicolas Magaud, Julien Narboux, and Pascal Schreck. Formalizing Projective Plane geometry in Coq. In *Automated Deduction in Geometry*, volume 6301 of *Lecture Notes in Computer Science*. Springer, 2011.
20. Timothy James McKenzie Makarios. A mechanical verification of the independence of Tarski's Euclidean axiom. Master's thesis, Victoria University of Wellington, 2012.
21. Filip Marić and Danijela Petrović. Formalizing analytic geometries. In *Automated Deduction in Geometry*. 2012.
22. Filip Marić, Ivan Petrović, Danijela Petrović, and Predrag Janičić. Formalization and implementation of algebraic methods in geometry. In *THedu*, volume 79 of *EPTCS*, 2011.

23. Laura Meikle and Jacques Fleuriot. Formalizing Hilberts Grundlagen in Isabelle/Isar. In *Theorem Proving in Higher Order Logics*, volume 2758 of *Lecture Notes in Computer Science*. Springer, 2003.
24. Robert Milewski. Fundamental theorem of algebra. *Formalized Mathematics*, 9(3), 2001.
25. Julien Narboux. Mechanical Theorem Proving in Tarski's Geometry. In *Automated Deduction in Geometry*, volume 4869 of *Lecture Notes in Computer Science*. Springer, 2007.
26. Tristan Needham. *Visual Complex Analysis*. Oxford University Press, 1998.
27. Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*, volume 2283 of *LNCS*. Springer, 2002.
28. Roger Penrose and Wolfgang Rindler. *Spinors and Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1987.
29. Wolfram Schwabhäuser, Wanda Szmielew, Alfred Tarski, and Michael Beeson. *Metamathematische Methoden in der Geometrie*. Springer, Verlag, 1983.
30. Hans Schwerdtfeger. *Geometry of Complex Numbers*. Dover Books on Mathematics. Dover Publications, 1979.
31. Phil Scott. Mechanising Hilberts Foundations of Geometry in Isabelle. Master's thesis, University of Edinburgh, 2008.
32. Christian Sternagel and René Thiemann. Executable matrix operations on matrices of arbitrary dimensions. *Archive of Formal Proofs*, June 2010. `http://afp.sf.net/entries/Matrix.shtml`, Formal proof development.
33. Jan von Plato. The axioms of constructive geometry. *Annals of Pure and Applied Logic*, 76(2), 1995.
34. Makarius Wenzel. Isabelle/Isar — a generic framework for human-readable proof documents. In *From Insight to Proof — Festschrift in Honour of Andrzej Trybulec, Studies in Logic, Grammar, and Rhetoric*, volume 10(23). University of Bialystok, 2007.