Constructive Geometry

Proof theory and straightedge-and-compass constructions

Michael Beeson

San José State University
beeson@cs.sjsu.edu
www.MichaelBeeson.com/Research

September, 2008

Euclid in Proclus's words (450 CE)

Euclid . . . put together the *Elements*, arranging in order many of Eudoxus's theorems, perfecting many of Theaetetus's, and also bringing to irrefutable demonstration the things which had been only loosely proved by his predecessors.

They say that Ptolemy once asked him if there were a shorter way to study geometry than the *Elements*, to which he replied that there was no royal road to geometry.

Pythagoras and Euclid

- The first "foundational crisis" was the discovery of the irrationality of $\sqrt{2}$.
- Euclid's *Elements* are to Pythagoras as *Principia Mathematica* is to Russell's paradox.
- This according to Max Dehn, *Die Grundlegung der Geometrie in Historischer Entwicklung*, in Moritz Pasch's *Vorlesungen ber Neuere Geometrie*.

Postulates *vs* Axioms (according to Geminus and Dehn)

- Postulates set forth our abilities to make certain constructions.
- Axioms merely state (static) properties
- Aristotle and Proclus offer different explanations of the difference, but I like this explanation.
- The idea is not Dehn's but is already attributed to Geminus by Proclus.
- Example: (Postulate 3) To describe a circle with any center and distance.

The Parallel Postulate

- As an axiom: Given a line L and a point P not on L, there exists exactly one line through P that does not meet L.
- As Postulate 5 [Heath translation]: If a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles.

Euclid's 48 Constructions

- The last book culminates in the construction of the Pythagorean solids
- 38 of these are in Books I-IV
- We will study the foundations today, not advanced geometry

Primitive Constructions

- IntersectLines(A,B,C,D) (AB meets CD)
- IntersectLineCircle1(A,B,C,D) (C is center)
- IntersectLineCircle2(A,B,C,D)
- IntersectCircles1(c_1,c_2)
- IntersectCircles2(c_1,c_2)

Constructors and Accessors

- Constructors: Line(A,B)
- Circle(A,B) (center A, passing through B)
- Circle3(A,B,C): circle with center A and radius BC
- Accessors: center(C), pointOnCircle(C), point1On(L), point2On(L).

Four Views of Euclid's Constructions

- **Algebra**: definability of some constructions in terms of others

Four Views of Euclid's Constructions

- **Algebra**: definability of some constructions in terms of others
- **Computer Science**: a programming language for Euclidean constructions

Four Views of Euclid's Constructions

- **Algebra**: definability of some constructions in terms of others
- **Computer Science**: a programming language for Euclidean constructions
- **Logic**: A formal theory close to Euclid, close to textbooks, useful for computerization.

Four Views of Euclid's Constructions

- **Algebra**: definability of some constructions in terms of others
- **Computer Science**: a programming language for Euclidean constructions
- **Logic**: A formal theory close to Euclid, close to textbooks, useful for computerization.
- **Constructive mathematics**: Axiomatization of constructive geometry.

Four Views of Euclid's Constructions

- **Algebra**: definability of some constructions in terms of others
- **Computer Science**: a programming language for Euclidean constructions
- **Logic**: A formal theory close to Euclid, close to textbooks, useful for computerization.
- **Constructive mathematics**: Axiomatization of constructive geometry.
- Today's talk will focus on the last two topics.

Circles of Zero Radius

- Let Circle3(A,B,C) be the circle with center A and radius BC.
- Should we require $B \neq C$ or not?
- That is, should we allow circles of zero radius?
- Argument in favor: it seems reasonable to allow the points of the compass to coincide.
- Possible argument against: maybe we can define it without assuming it.

A theorem in the algebraic setting

Suppose that we do require $B \neq C$ for the definedness of Circle3(A, B, C). Then:

Theorem (with Freek Wiedijk) *Circle3 is not definable from the other operations.*

Proof. Let t be a term definable without Circle3, with one free point variable x. Then t becomes undefined when one of the variables is set equal to a constant. For example, Circle(x, β) is undefined when $x = \beta$. But not so for Circle3(x, α, β).

Therefore, we also consider Circle4(A,B,C), which is like Circle3 except that it is defined for all A, B, and C, producing a circle of zero radius if B and C are the same point.

www.dynamicgeometry.org

- The 48 Euclidean constructions in Euclid's words, animated (Ralph Abraham)
- An applet Diagrammer allows you to make your own constructions. (Chris Mathenia and Brian Chan)
- An applet Constructor provides a visual interpreter for Geoscript. You can step through or into the 48 Euclidean scripts. (with some help from Thang Dao.)

Book I, Prop. 2

- Euclid's compass is "collapsible"
- You cannot use it directly to construct Circle3(A,B,C), the circle with center A and radius BC.
- Book I, Prop. 2 is intended to show that Circle3 need not be assumed, because
- Given A, B, and C, we can construct a point D with $AD = BC$.

Book I, Prop. 2

- Euclid's compass is "collapsible"
- You cannot use it directly to construct Circle3(A,B,C), the circle with center A and radius BC.
- Book I, Prop. 2 is intended to show that Circle3 need not be assumed, because
- Given A, B, and C, we can construct a point D with $AD = BC$.
- But Euclid's construction assumes not only $B \neq C$, but also $B \neq A$, and the point constructed does not depend continuously on B as B tends to A.
- This is a bad omen for constructive geometry, because computable points must depend continuously on their parameters.

Uniform I.2

- The uniform version of this proposition says that, given A, B, and C, with $B \neq C$, we can construct $D = e(A, B, C)$ with $AD = BC$, without assuming $A \neq B$.
- Given such a term e, we could define Circle3(A,B,C) = Circle(A,$e(A, B, C)$).
- Given Circle3, we could define $e(A, B, C) =$ pointOnCircle(Circle3(A,B,C).
- Having Circle3 is equivalent to "realizing" uniform I.2.

Projection

The properties of projection are

- project(P,L) is on L
- P lies on the perpendicular to L at project(P,L).
- project(P,L) is defined whether or not P is on L.

Using projection, we can assign coordinates on two perpendicular lines to any point P.

- Projection is continuous, like the other basic constructions.
- Projection is computable-we can compute project(P,L) to any desired accuracy.
- Projection is definable using Circle4
- For that it is crucial that we allow circles of zero radius.

Models of the Elementary Constructions

- The "standard plane" $\mathbb{R}^2$
- The "recursive plane". Points are given by recursive functions giving rational approximations to within $1/n$.
- The minimal model, the points constructible by ruler and compass
- The algebraic plane, points with algebraic coordinates
- The Poincaré model. These constructions work in non-Euclidean geometry too.

Coordinatization

- Every model is a plane over some ordered field.
- Because of quantifier elimination (Tarski) every real-closed field gives a model of Tarski geometry.
- Euclidean fields (every positive element has a square root) correspond to the geometry of constructions.

A problem of Tarksi

- Is the geometry of constructions decidable?
- That is, the theory of Euclidean fields (ordered fields in which positive elements have square roots)?
- Ziegler (1980) says not. Indeed any finitely axiomatizable field theory that has $\mathbb{R}$ or the p-adics as a model is undecidable.
- His proof is only 11 (difficult) pages.
- I have translated this paper if anyone wants an English version.

Another problem of Tarski

- Is the smallest Euclidean field $Q(\sqrt{})$ decidable?
- Goes beyond J. Robinson's famous results for Q and the algebraic number fields, because $Q(\sqrt{})$ is not of finite degree over Q.
- Still an open problem.

Hilbert, Tarski, Borsuk, and Szmielew

- Hilbert introduced the primitives of betweenness (A is between B and C) and congruence (of segments), and considered points, lines, and planes with an incidence relation.
- Tarski's theory has variables for points only. Congruence of segments AB and CD becomes the equidistance relation $\delta(A, B, C, D)$.
- Details are in Borsuk and Szmielew

Constructive Geometry

- Straightedge-and-compass constructions
- Constructive (intuitionistic) logic
- Is there a connection?

Constructive Geometry

- Straightedge-and-compass constructions
- Constructive (intuitionistic) logic
- Is there a connection?
- What is constructively proved to exist (in a suitable theory IEGC) should be constructible with straightedge and compass

Constructive Geometry

- Straightedge-and-compass constructions
- Constructive (intuitionistic) logic
- Is there a connection?
- What is constructively proved to exist (in a suitable theory IEGC) should be constructible with straightedge and compass
- IEGC is to straightedge-and-compass as HA is to recursive functions.

A (classical) theory EGC of the Elementary Geometry of Constructions

- Quantifier-free, disjunction-free axiomatization
- Terms for the primitive geometric constructions.
- Models are planes over Euclidean fields
- Conservative over Tarski's geometry of constructions.
- Axioms merely state the properties of the elementary constructions

An issue in the axiomatization

- LineCircle1(A,B,C,D) is one intersection point of Line(A,B) and Circle(C,D).
- Not trivial to distinguish LineCircle1(A,B,C,D) from LineCircle2(A,B,C,D)
- ABC should be a "left turn" or a "right turn".
- How to define that?

Right and left handedness

- There are three distinguished noncollinear points (given by constants) α, β, and γ.
- Arbitrarily we say $\alpha\beta\gamma$ is "left" and $\alpha\gamma\beta$ is "right".
- Still have to define "ABC has the same handedness as DEF"
- That can be done, but it takes too much time to explain here.
- This doesn't seem to be in geometry books.
- I have verified in detail the mutual intepretatibility of EGC and the theory in Greenberg's textbook *Euclidean and Non-Euclidean Geometries*

Multi-sorted theories

- Six possible sorts: Point, Line, Circle, Segment, Ray, Arc
- Angles treated as triples of points
- Function symbols for the elementary constructions IntersectLines, IntersectLineCircle1, IntersectLineCircle2, IntersectCircles1, IntersectCircles2, and Circle(A,B,C)
- Also for Circle3(A,B,C), which constructs the circle with center A and radius BC.
- Also for the accessor functions center(C), pointOnCircle(C), point1On(L), point2On(L).
- Logic of partial terms (LPT) because these functions are partial.

Separability

- Arcs, rays, and segments can be defined in terms of points, lines, and circles
- Even circles and lines can be eliminated, e.g. Hilbert-Bernays works only with points.
- We work with points, lines, and circles.

Intuitionistic Geometry

- Use intuitionistic logic
- In intuitionistic logic, we do not have $a < b \vee a = b \vee b < a$ for points on a line.
- In view of that, several of Hilbert's axioms are not correct with intuitionistic logic.
- Reductions to field theory need reconsideration.
- Intuitionistic RCF is undecidable (Gabbay)
- We are interested in the intuitionistic geometry of constructions
- Axioms the same as for GC
- Handedness can be defined using intuitionistic logic.

Decidable Equality

- Decidable equality means $A = B$ or $A \neq B$.
- If points are given by real numbers there is no algorithm to decide equality.
- If points are given by rational or Euclidean numbers then there is an algorithm, but not a geometric construction, i.e. no Geoscript program, to decide equality.
- Euclid, as made right by Proclus, uses proof by cases (and often only one case is illustrated in Euclid).

Apartness

- Apartness (introduced by Heyting) is a positive version of inequality.
- $A\#B$ means (intuitively) that we can find a lower bound on the distance from A to B.
- Axiomatically one could add $\#$ as a primitive relation with natural axioms. In particular
- $\neg A\#B \supset A = B$
- $A\#B \supset A \neq B$
- $B\#C \supset A\#B \vee A\#C$
- $f(u,v)\#f(a,b) \supset u\#a \vee v\#b$ for primitive constructions f.

Apartness holds computably but not continuously

We can invent a "construction" to correspond to apartness: apart(A,B,C) should satisfy:

- If $B\#C$ then $P =$apart(A,B,C) is defined
- $P = B \vee P = C$
- $P\#A$
- Just compute A, B, and C to an accuracy less than $1/3$ of $|B - C|$.
- But apart, although computable, is not extensional and not continuous.
- For that reason I am interested in theories without apartness.
- Also apartness does not occur in Euclid, so if we want a theory that is close to Euclid, we should not include apartness.
- If one does include apartness, one can prove uniform I.2

Stability Axioms

- We take "Markov's principle" $\neg\neg P \supset P$ for atomic P (betweenness, equidistance, equality, and definedness).
- The axiomatization is quantifier-free and disjunction-free.
- Seems to correspond better to Euclid's Elements than using apartness
- "Markov's principle" is $b \neq c \supset b \# c$.
- In the presence of Markov's principle, we can use $\neq$ for $\#$.
- In that case we do not need an extra symbol for apartness, only the axioms.

Projection and coordinatization

- It's not so obvious that projection is enough to define addition, multiplication, and sqrt without needing test-for-equality, but it is!
- Example lemma: para(p,L) constructs a line through p that is parallel to L if p is not on L, and equal to L if p is on L. Such a construction para can be defined using project.

Axioms of IEGC

- Axioms for the elementary constructions including Circle4
- Stability axioms
- Intuitionistic logic

Research plan

IECG is to straightedge-and-compass constructions as HA is to recursive functions.
We define suitable versions of the tools used in the proof theory of HA:

- cut elimination
- realizability
- the Dialectica interpretation

With these tools we obtain nice metatheorems about IEGC.

Extraction of Algorithms from Proofs

We know how to extract terms for computable functions from proofs in number theory or analysis. Now we want to extract geometrical constructions from proofs in EGC and related theories.

Theorem. Suppose IEGC proves

$$\forall x\,(P(x) \supset \exists y\, A(x,y))$$

with P negative. Then there exist a term $t(x)$ such that IEGC proves

$$\forall x\,(P(x) \supset A(x,[y := t(x)]))$$

Here x can stand for several variables.

Terms of IEGC correspond directly to (uniform) Euclidean constructions; a term can be directly compiled to a Geoscript program.

Uniform validity for IEGC

Corollary. Suppose IEGC proves

$$\forall x\,(P(x) \supset \exists y\, A(x,y))$$

with P negative. Then this theorem is uniformly valid, i.e. y depends locally continuously on x.

Proof. All the terms of IEGC are continuous on their domains of definition.

Since the same proof works with Circle3 instead of Circle4, and since uniform I.2 is equivalent to Circle4, it follows that uniform 1.2 is not provable in the restricted version of IEGC that has only Circle3.

Proof by cut elimination

- Standard proof method, appealing to permutability of inferences (Kleene 1951)
- Consider a cut-free proof of $\Gamma \Rightarrow \exists\, y\, A(x, y)$, where Γ is a list of universal closures of axioms and the negative hypotheses P.
- The last step can be assumed to introduce the quantifier, so the previous step gives the desired conclusion.
- Doesn't work if apartness is used because the apartness axioms involve disjunction and such inferences do not permute.

Constructions and classical logic

Theorem. Suppose EGC with classical logic proves

$$\forall x\,(P(x) \supset \exists y\, A(x,y))$$

with P and A quantifier-free. Then there exist terms $t_1(x), \ldots, t_n(x)$ of EGC such that EGC proves

$$\forall x\,(P(x) \supset A(x,[y := t_1(x)]) \vee \ldots \vee A(x,[y := t_n(x)])$$

Proof. By Herbrand's theorem.

Realizability

A tool used in the metatheory of intuitionistic systems. We define $e$realizes$A$, written $e\mathbf{r}A$, for each formula A. Here e can be a term or a program (e.g. index of a recursive function). The key clauses are

- $e\mathbf{r}(A \supset B) \iff \forall q\,(q\mathbf{r}A \supset Ap(e,q)\mathbf{r}B)$
- $e\mathbf{r}\exists x\,A \iff p_1(e)\mathbf{r}A[x := p_0(e)]$.

Here p_0 and p_1 are projection functions:

$$x = \langle p_0(x), p_1(x)\rangle \qquad \text{if } x \text{ is a pair}$$

Ap is application, as in lambda calculus or combinatory logic. Pairing and Ap are not available in IECG.

Lambda logic

- I have already studied in general what happens when we add lambda terms to a first order theory.
- Lambda Logic was introduced (for other purposes) in IJCAR-2004.
- See papers on my website www.MichaelBeeson.com/Research
- Lambda logic = Type-free lambda calculus plus first-order logic.
- IEGC in lambda logic contains lambda, Ap, beta-reduction as well as IEGC. (We need unary predicates Point, Line, etc. because lambda logic is not multisorted.)
- Let GT be IEGC plus lambda logic.

Lambda logic does not prove new geometrical theorems

- Lambda logic is conservative over FOL plus the schema "there exist at least N things" (for each N).
- But IEGC already proves there exists at least N things.
- Hence IEGC + lambda logic is conservative over IEGC.
- Hence GT is conservative over IEGC.

Soundness of Realizability

Theorem. If IEGC proves A then GT proves $t\mathbf{r}A$ for some normal term t whose free variables are among those of A. Similarly for q-realizability.

Corollary Extraction of algorithms for IEGC + Apartness.
Suppose IEGC + Apartness proves

$$\forall x\,(P(x) \supset \exists y\, A(x,y))$$

with P a conjunction of atomic formulae. Then there is a term t of IEGC + Apartness such that IEGC proves

$$\forall x\,(P(x) \supset A(x, y := t(x)))$$

We cannot do this by cut-elimination since the apartness axioms are not disjunction-free.

Double negation interpretation

A^- defined as usual for predicate calculus.

- $P^-(x) := \neg\neg P(x)$ for atomic P
- $(A \vee B)^- := \neg(\neg A^- \wedge \neg B^-)$
- $\exists x\, A)^- : -\neg\forall x\, \neg A^-$
- $^-$ commutes with $\wedge$, $\neg$, $\supset$, and $\forall$

If classical ECG proves A then IECG proves A^-. This works because of the stability axioms (Markov's principle, if you like).

The Dialectica interpretation

The Gödel Dialectica interpretation A^0 can be defined for any theory, into that theory plus lambda logic.

$$A^0 : -\exists z \, \forall x \, A_0(x, y)$$

Here x and y are functionals of finite type over the base type of (in our case) points.
If (as in our case) we do not have decidable equality for the base type, the definition is more complicated. We use the "Diller-Nahm variant" of the Dialectica interpretation.
Since the axioms of IEGC are quantifier-free and disjunction-free, they are their own interpretations.

What comes out of the dialectica interpration

Theorem. Suppose (classical) EGC proves $\forall x \exists y A(x, y)$ for quantifier-free, disjunction-free A. Then IEGC + Apartness proves this theorem also, and there are terms $t_1(x), \ldots t_n(x)$ such that it proves

$$\forall x A(y : t_1(x)) \vee \ldots \vee A(y : t_n(x))$$

Proof. As usual we compose the double-negation interpretation and the Dialectica interpretation. This gets us to GT; then we use the fact that normal terms of type (0,0) in GT are geometrical terms, and apply the conservativity of GT over IEGC. There are many details to check, but this is the outline. Because we had to use the Diller-Nahm variant, we get n terms instead of just one.

Conclusion

- The algorithmic and axiomatic viewpoints have a long history in geometry
- Modern axiomatizations of classical geometry are well understood.
- I studied axiomatizations of geometry with intuitionistic logic and showed that, as in number theory, constructive proof methods imply that things proved to exist can be constructed, in this case with straightedge and compass.
- With appropriate axiomatizations we get not only computable, but continuous, dependence on parameters.
- Light is shed on questions like Euclid I.2 that go back millenia
- Lambda logic is used to extend realizability and the Dialectica interpretation to geometry.