Exact constructions with inexact diagrams

John Mumma
Division of Logic, Methodology, and Philosophy of Science in the Suppes Center for History and Philosophy of Science, Stanford University

Symposium on Constructive Geometric Reasoning
Stanford University, Oct. 24, 2009

Three Parts of Talk

I Comments from Bernays on constructive and existential approaches in geometry

II Formal proof system *E*, an analysis of Euclid's figure based reasoning

J. Avigad, E. Dean, J. Mumma, *A Formal System for Euclid's Elements*, to appear in the Review of Symbolic Logic.

III Discussion of the picture *E* provides of Euclid's geometric constructions

Bernays on Constructive and Existential Geometry

Comments from

The Philosophy of Mathematics and Hilbert's Proof Theory
1930

Early in the piece, Bernays contrasts Hilbert and Euclid to illustrate the distinctive features of modern axiomatics.

Bernays on Constructive and Existential Geometry

Bernays emphasizes two key features:

- Abstractness

Hilbert	**Euclid**
Uninterpreted primitives	Contentful primitives

- Existential Form

Hilbert	**Euclid**
Pre-existing domain of objects	Constructed objects

Bernays on Constructive and Existential Geometry

The second difference

For Hilbert, ‘points, lines, and planes in their totality’ are ‘fixed in advance.’ Axioms describe relations between these objects.

For Euclid, ‘the geometric figures under consideration’ are always thought of ‘as constructed ones.’ Axioms describe methods of construction, and relations between constructed objects.

Bernays on Constructive and Existential Geometry

The second difference

Hilberts's axiom I,1:

For every two points *A* and *B*, there exists a line *a* that contains each of the points *A*, *B*.

Euclid's first postulate:

To draw a straight line from any point to any point.

The Formal System *E*

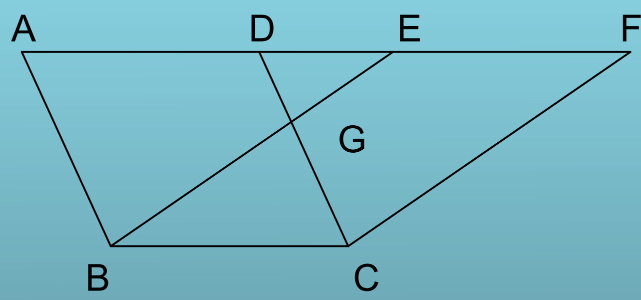
What's involved with existential axiomatic theory of geometry well-understood. (Tarski, *What is Elementary Geometry?*, 1959)

Various options in fleshing out what Euclid's constructive approach to geometry amounts to formally.

Goal of the formal system *E*: provide a formal picture that accounts for the role of geometric diagrams in Euclid's constructions.

Principles behind E

The proof system is based on Ken Manders' observation that Euclid only uses certain properties of geometric diagrams in proofs (roughly its topological properties).



<u>Metric Properties</u>

AB=DC
CBE=BED

<u>Topological Properties</u>

Intersection of BE and DC
Containment of DE in AF

As Manders observed in *The Euclidean Diagram*:

It is only via its non-metric spatial relations that a diagram seems to justify an inference in the *Elements.*

Inferences between spatially separated magnitudes carried out in the text.

Principles behind E

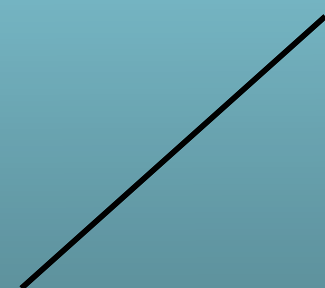
The likely reason behind Euclid's self-imposed restriction:

Geometric constructions produce ideal and exact entities, while diagrams are concrete and inevitably inexact.

<u>Def. 2 of Bk I in *Elements*</u>

A line is a **breadthless** length

<u>Line in diagram</u>



Implicit norm: an inexact diagram ought not be used to reason about exact relations, but only inexact positional ones.

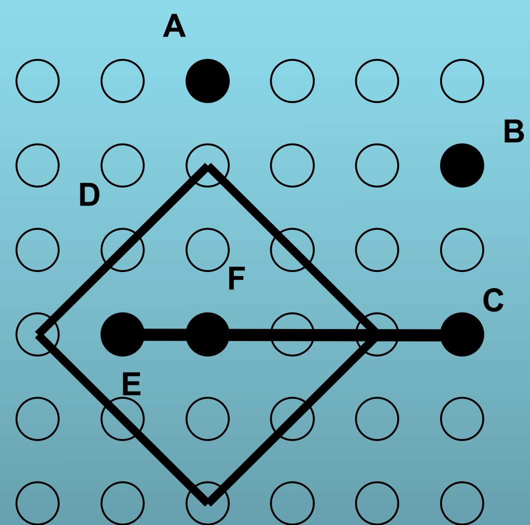
Principles behind *E*

Manders' work first inspired **Eu**, a formal system of proof with:

- Sentential symbols *S*
- Diagrammatic symbols *D*
- Rules *R* for manipulating *S* and *D*

Principles behind *E*

The diagrams *D* of **Eu**



n x n arrays for any n

Rules for well-formed diagram specify how array elements can be distinguished as points, lines, and circles.

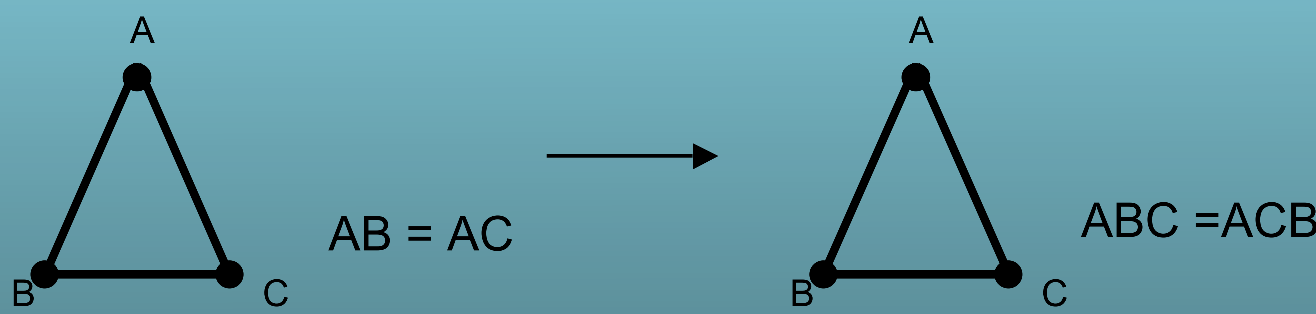
Principles behind *E*

Derivable claims of **Eu** are of the form

$$\Delta_1, A_1 \rightarrow \Delta_2, A_2$$

where Δ_1, Δ_2 are in *D* and A_1 , A_2 are in *S*.

Example: I, 5 of *Elements*



Symbols *D* track positional information of proof.
Symbols *S* track metric information.

Principles behind *E*

Structure of **Eu** proof has structure of a proof in *Elements*. A proof has a *construction* stage and a *demonstration* stage.

Example: Proposition 5, Book I

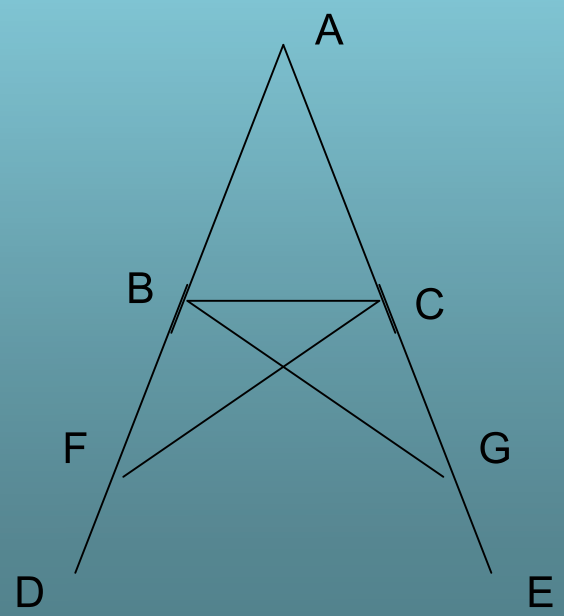
In an isosceles triangle the angles at the base are equal to one another.

Principles behind *E*

Proposition 5, Book I

In an isosceles triangle the angles at the base are equal to one another.

PROOF



Let ABC be an isosceles triangle.
Extend AB to D and and AC to E.
Pick an arbitrary point F on BD, and cut off a line AG equal to AF on AE.
Join FC and GB.
By SAS, FC = BG, ∠ABG =∠ACF and ∠ABG =∠ACF.
Equals subtracted from equals are equal, so BF=CG.
By SAS again, ∠CBF =∠BCG.
And again since equals subtracted from equals are equal, ∠CBF =∠BCG.

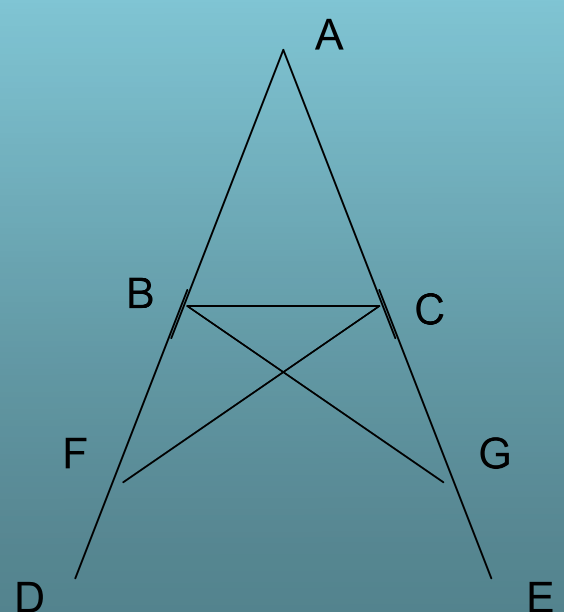
QED

Principles behind *E*

Proposition 5, Book I

In an isosceles triangle the angles at the base are equal to one another.

PROOF



Let ABC be an isosceles triangle.
Extend AB to D and and AC to E.
Pick an arbitrary point F on BD, and cut off a line AG equal to AF on AE.
Join FC and GB.
By SAS, FC = BG, ∠ABG =∠ACF and ∠ABG =∠ACF.
Equals subtracted from equals are equal, so BF=CG.
By SAS again, ∠CBF =∠BCG.
And again since equals subtracted from equals are equal, ∠CBF =∠BCG.

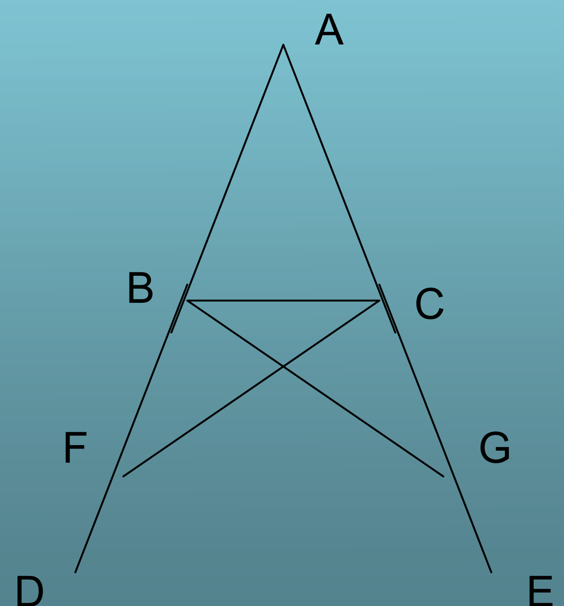
QED

Principles behind *E*

Proposition 5, Book I

In an isosceles triangle the angles at the base are equal to one another.

PROOF



Let ABC be an isosceles triangle.
Extend AB to D and and AC to E.
Pick an arbitrary point F on BD, and cut off a line AG equal to AF on AE.
Join FC and GB.

By SAS, FC = BG, $\angle ABG = \angle ACF$ and $\angle ABG = \angle ACF$.

Equals subtracted from equals are equal, so BF=CG.

By SAS again, $\angle CBF = \angle BCG$.

And again since equals subtracted from equals are equal, $\angle CBF = \angle BCG$.

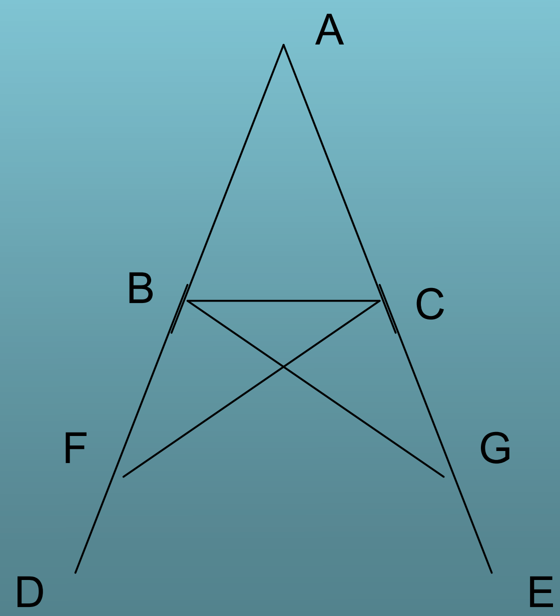
QED

Principles behind *E*

Proposition 5, Book I

In an isosceles triangle the angles at the base are equal to one another.

PROOF

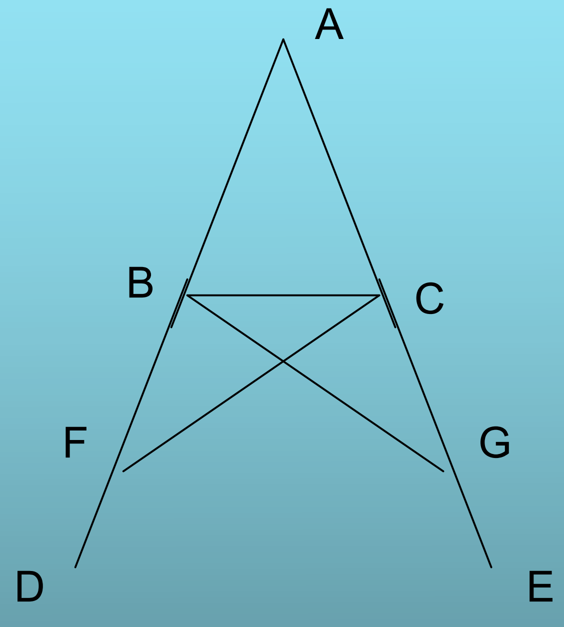


Let ABC be an isosceles triangle.
Extend AB to D and and AC to E.
Pick an arbitrary point F on BD, and cut off a line AG equal to AF on AE.
Join FC and GB.
By SAS, FC = BG, ∠ABG =∠ACF and ∠ABG =∠ACF.
Equals subtracted from equals are equal, so BF=CG.
By SAS again, ∠CBF =∠BCG.
And again since equals subtracted from equals are equal, ∠CBF =∠BCG.

QED

Principles behind *E*

PROOF



Let ABC be an isosceles triangle.
Extend AB to D and and BC to E.
Pick an arbitrary point F on BD, and cut off a line AG equal to AF on AE.
Join FC and GB.
By SAS, FC = BG, ∠ABG =∠ACF and ∠ABG =∠ACF.
Equals subtracted from equals are equal, so BF=CG.
By SAS again, ∠CBF =∠BCG.
And again since equals subtracted from equals are equal, ∠CBF =∠BCG.

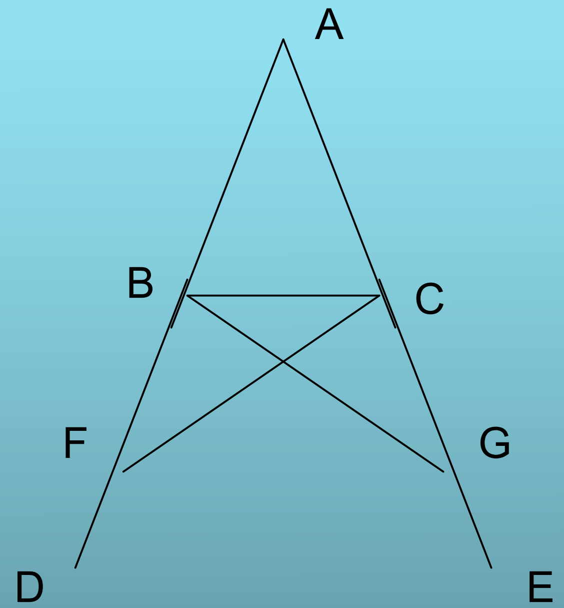
QED

<u>**Eu** analysis</u>

•Rules of construction stage specify ways the initial pair Δ_1, A_1 can be augmented with new objects.

Principles behind *E*

PROOF



Let ABC be an isosceles triangle.
Extend AB to D and and BC to E.
Pick an arbitrary point F on BD, and cut off a line AG equal to AF on AE.
Join FC and GB.
By SAS, FC = BG, ∠ABG =∠ACF and ∠ABG =∠ACF.
Equals subtracted from equals are equal, so BF=CG.
By SAS again, ∠CBF =∠BCG.
And again since equals subtracted from equals are equal, ∠CBF =∠BCG.

QED

__**Eu** analysis__

- Rules of construction stage specify ways the initial pair Δ_1, A_1 can be augmented with new objects.
- Rules of demonstration stage specify the ways inferences can be drawn from the Δ, A pair produced by construction.

Principles behind *E*

Eu leaves open the question:

How does Euclid’s diagrammatic method relate to modern, logical axiomatizations of geometry?

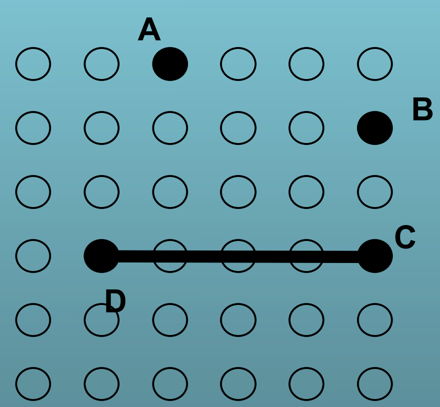
Specifically, can we prove with the method everything that ought to be provable according to a modern axiomatization?

The system *E* was developed to address this question.

Formal system *E*

Just as in **Eu**, the data for a geometric configuration in *E* has a diagrammatic and a metric component.

The key difference is the way diagrams are modeled.



Eu diagram

A, B, C, D points
L line
A≠B
C≠D
D on L
C on L
Sameside(A,B,L)

E diagram

With such data lists, *E* formalizes the constructive and inferential moves Euclid makes with diagrams.

The Formal System *E*

Sorts

Points: p,q, r … Lines: L, M, N,… Circles: α, β, γ, …

Segments: **seg**(pq), **seg**(rs).. Angles: **ang(**pqr).. Areas: **area**(pqr)…

Functions

seg, **ang, area**, and an addition function **+** on segment, angle and area sorts.

Relations

Equality: =

Diagrammatic Relations	Metric Relations
on(x,y) x a point, y a line or circle	**seg**(xy) > **seg**(zw)
sameside(x,y,z) x, y points, z a line	**seg**(xy) = **seg**(zw)
inside(x,y) x a point, y a circle	**ang**(xy) > **ang**(zw)
between(x,y,z) x,y,z points	**ang**(xy) = **ang**(zw)
center(x,y) x a point, y a circle	**area**(xy) > **area**(zw)
intersect(x,y) x,y line or circle	**area**(xy) = **area**(zw)

Literals

Atomic relation or negation of atomic relation.

The Formal System *E*

What's derived in *E*:

$$\Gamma \Rightarrow \exists\, \mathbf{q},\mathbf{M},\beta \;\; \Delta$$

where Γ, Δ are both lists of literals, and **q**, **M**, β are tuples of point, line, and circle variables.

It represents the following geometric claim:

Given a figure satisfying the conditions in Γ*,*
one can construct points q, lines M and circles β
satisfying the conditions of Δ*.*

The Formal System *E*

What's derived in *E*:

$$\Gamma \Rightarrow \exists\, \mathbf{q},\mathbf{M},\beta \;\; \Delta$$

where Γ, Δ are both lists of literals, and **q**, **M**, β are tuples of point, line, and circle variables.

Example:

a on L, b on L $\Rightarrow$ $\exists$ c c on L, between(acb), seg(ac) = seg(cb)

Representation in *E* of I,10 in the *Elements*.

The Formal System *E*

As Euclid does not use much logic in his proofs, the logic whereby expressions

$$\Gamma \Rightarrow \exists_{\mathbf{q},\mathbf{M},\beta}\ \Delta$$

is restricted.

Euclid's general logic pattern:

Conditional proof where objects satisfying certain conditions are introduced (construction stage), and inferences about the resulting complex of objects are drawn (demonstration stage).

The Formal System *E*

As Euclid does not use much logic in his proofs, the logic whereby expressions

$$\Gamma \Rightarrow \exists_{\mathbf{q},\mathbf{M},\beta}\ \Delta$$

is restricted.

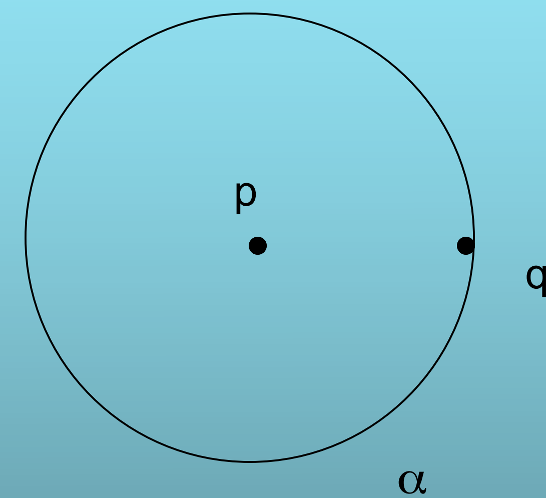
Construction stage steps codified by introduction rules for literals with new objects.

Demonstration stage steps codified by inference rules on list of literals.

The Formal System *E*

<u>Construction Rules of *E*</u>

$$\frac{\Gamma \Rightarrow \exists\, \mathbf{x}\ \Delta}{\Gamma \Rightarrow \exists\, \mathbf{x}, \mathbf{x}'\ \Delta, \Delta'}$$



Example:

Circle construction

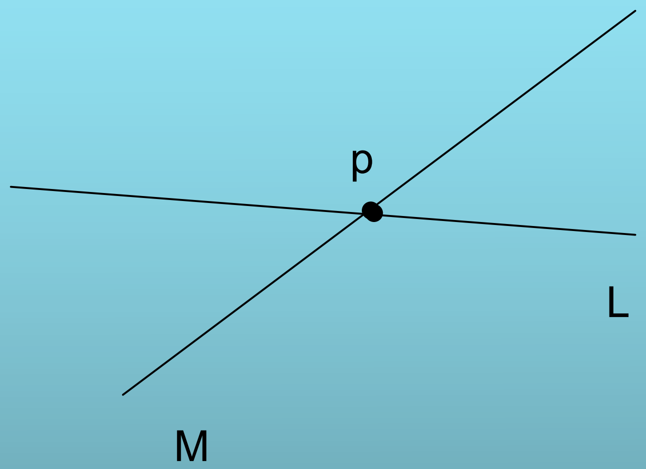
Prerequisite: $p \neq q$ (in Γ, Δ)

Conclusion: p is inside α, q is on α (in Δ')

The Formal System *E*

Construction Rules of *E*

$$\frac{\Gamma \Rightarrow \exists\, \mathbf{x}\ \Delta}{\Gamma \Rightarrow \exists\, \mathbf{x}, \mathbf{x}'\ \Delta, \Delta'}$$



Example:

Point of line intersection

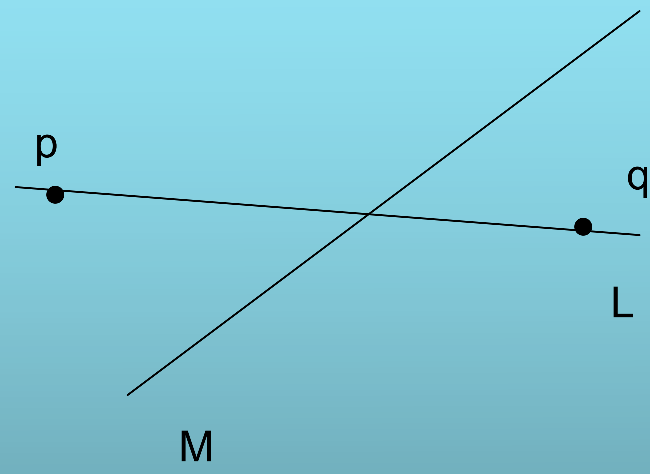
Prerequisite: M, L intersect (in Γ, Δ)

Conclusion: p is on M, p is on L (in Δ')

The Formal System *E*

Demonstration Inference Rules of *E*

$$\frac{\Gamma \Rightarrow \exists \mathbf{x}\, \Delta}{\Gamma \Rightarrow \exists \mathbf{x}\, \Delta, \Delta'}$$



Example:

Line intersection

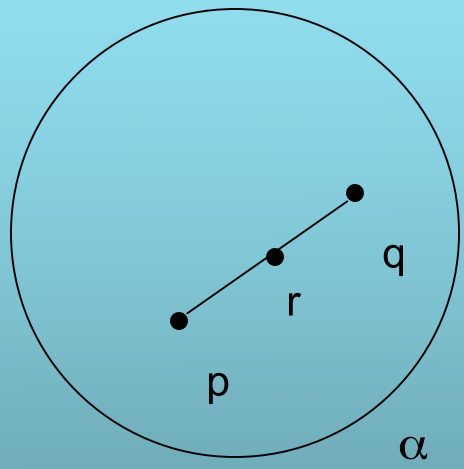
Premises: p, q on different sides of M
p, q on L (in Γ, Δ)

Conclusion: L,M intersect (Δ')

The Formal System *E*

<u>Demonstration Inference Rules of *E*</u>

$$\frac{\Gamma \Rightarrow \exists\, \mathbf{x}\, \Delta}{\Gamma \Rightarrow \exists\, \mathbf{x}\, \Delta, \Delta'}$$



Example:

A circle rule

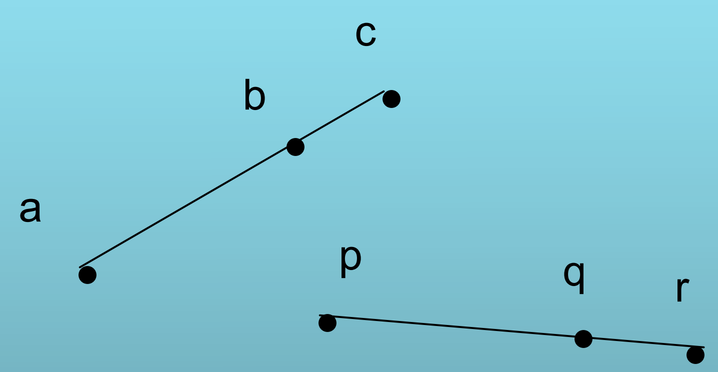
Premises: p, q inside α
r between p and q (in Γ, Δ)

Conclusion: r inside α (Δ')

The Formal System *E*

<u>Demonstration Inference Rules of *E*</u>

$$\frac{\Gamma \Rightarrow \exists\, \mathbf{x}\, \Delta}{\Gamma \Rightarrow \exists\, \mathbf{x}\, \Delta, \Delta'}$$



Example:

Equals added to equals are equal

Premises: seg(ab)=seg(pq), seg(bc) =seg(qr)
between(abc), between(pqr) (in Γ, Δ)

Conclusion: seg(ac) = seg(pr) (Δ’)

The Formal System *E*

Not all of Euclid's reasoning can be formalized in a linear way as the introduction of literals with new objects and the inference of literals.

In some proofs Euclid reasons by cases, and in others by contradiction.

Further, one needs a way to apply a previously proven result.

Logical rules

Case splitting
Ex falso
Theorem application

The Formal System *E*

Case splitting:

$$\frac{\Gamma \Rightarrow \exists \mathbf{x}\, \Delta \qquad \Gamma, \varphi \Rightarrow \exists \mathbf{y}\, \Delta' \qquad \Gamma, \neg\varphi \Rightarrow \exists \mathbf{y}\, \Delta'}{\Gamma \Rightarrow \exists \mathbf{x}, \mathbf{y}\, \Delta, \Delta'}$$

Ex falso:

$$\frac{\varphi \ldots \neg\varphi}{\bot} \qquad\qquad \frac{\bot}{\Gamma \Rightarrow \exists \mathbf{x}\, \Delta}$$

where $\Gamma \Rightarrow \exists \mathbf{x}\, \Delta$ is any expression of *E.*

Proof by contradiction formalized as a case splitting argument where all but one of the cases leads to contradiction.

The Formal System *E*

Logical Rules of *E*

Theorem application:

$$\frac{\Gamma \Rightarrow \exists\, \mathbf{x}\, \Delta \qquad \Pi \Rightarrow \exists\, \mathbf{y}\, \Theta}{\Pi \Rightarrow \exists\, \mathbf{y}, \mathbf{z}'\, \Theta, \Theta'}$$

where $\Gamma' \Rightarrow \exists\, \mathbf{x}'\, \Delta'$ is a renaming of the variables in $\Gamma \Rightarrow \exists\, \mathbf{x}\, \Delta$ where Γ' is contained in Π, and $\mathbf{z}'$ and Θ' are contained in Δ'.

E's analysis of diagrammatic inference

E provides a precise a characterization of what by Euclid's method is immediate from the diagram.

The characterization based on the diagrammatic inference rules of the system. (Demonstration rules which contain only diagrammatic literals.)

<u>Tarksi betweenness axiom</u>

$B(abc) \& B(bdc) \rightarrow B(adc)$

<u>E betweenness rules</u>

$B(abc) \& B(bdc) \rightarrow B(adc)$

$B(abc) \& B(bcd) \rightarrow B(acd)$

$B(abc) \& B(abd) \rightarrow \neg B(cbd)$

$B(abc) \& \neg B(adc) \rightarrow \neg B(bdc)$

.

.

.

E’s analysis of diagrammatic inference

Given a list of literals Γ the set of the *direct diagrammatic consequences* of Γ is defined as the closure of Γ under the diagrammatic inference rules.

By *E*’s analysis,

What’s immediate from a diagram in a Euclidean proof	=	The direct diagrammatic consequences of the information Γ represented by the diagram in the proof.

E's analysis of diagrammatic inference

By *E*'s analysis,

What's immediate from a diagram in a Euclidean proof = The direct diagrammatic consequences of the information Γ represented by the diagram in the proof.

All the diagrammatic inference rules have a simple logical form: φ_1 &... & $\varphi_n \rightarrow \psi$

Consequently, the question 'is θ a direct diagrammatic consequence of Γ?' can be determined in time polynomial with the number of point, line and circle variables of Γ. (Theorem 3.6 in Avigad et al.)

E's analysis of diagrammatic inference

By *E*'s analysis,

What's immediate from a diagram in a Euclidean proof = The direct diagrammatic consequences of the information Γ represented by the diagram in the proof.

Suitability of analysis depends on whether *E* has enough diagrammatic rules. No conclusive argument at present that it does.

Completeness of *E*

The semantics for *E*

The analytic geometry of F x F, where F is any field closed under square roots (i.e. F is any Euclidean field).

Expressions $\Gamma \Rightarrow \exists\, \mathbf{x}\, \Delta$ interpreted as ∀∃ claims.

Soundness of rules easily verified.

Tarski (1959) contains an axiomatization complete with respect to this semantics.

Idea of proof: show that Tarski's theory is in a sense a conservative extension of *E*.

Completeness of *E*

Proof uses the result of Negri (2003):

Any sequent provable in a sequent calculus with geometric rule schemes has a cut free proof.

<u>Geometric Formula</u>

$\forall \mathbf{x}\ (B(\mathbf{x}) \rightarrow \bigvee\ (\exists \mathbf{y}\ A_i(\mathbf{x},\mathbf{y})))$

<u>Geometric Rule Schemes</u>

$$\frac{A_1(\mathbf{x},\mathbf{y}),\ \Pi \Rightarrow \Theta\ \ldots\ A_n(\mathbf{x},\mathbf{y}),\ \Pi \Rightarrow \Theta}{B(\mathbf{x}),\ \Pi \Rightarrow \Theta}$$

where B and the A_i's are conjunctions of atoms.

Strategy of proof: modify Tarski's theory into a sequent calculus with geometric rule schemes, and define translation functions between the resulting theory (call it *T*) and *E*.

Completeness of *E*

$\Gamma \Rightarrow \exists\, \mathbf{x}\, \Delta$ $\xrightarrow{\quad \pi \quad}$ χ

E *T*

First a translation π which maps valid sequents of *E* to valid *regular* sequents of *T* is defined.

(A regular sequent is a geometric sequent where there is no disjunction in antecedent.)

Completeness of *E*

$\Gamma \Rightarrow \exists \mathbf{x}\, \Delta \xrightarrow{\ \pi\ } \chi$

E   *T*

If χ is valid in every F x F, then χ has a proof in *T*.

Further, since *T* is a geometric theory, by Negri's result χ has a cut free proof in *T*.

Completeness of *E*

$\Gamma' \Rightarrow \exists \mathbf{x}' \Delta'$ ⟵ χ

ρ

E *T*

And so, we can define a translation function ρ which translates such provable χ's in *T* to provable sequents in *E*.

Completeness of *E*

What this plan requires:

•The modification of Tarski's theory into a geometric sequent calculus *T*.

•The definition of translation function π mapping *E* sequents to regular sequents of *T*.

•The definition of translation function ρ mapping provable regular formulas of *T* to provable *E* sequents.

•Proof that if *E* proves $\rho(\pi(\Gamma \Rightarrow \exists \mathbf{x}\, \Delta))$, then it proves $\Gamma \Rightarrow \exists \mathbf{x}\, \Delta$.

Geometric constructions in *E*

What they are:

Sequence of those steps in proof which result from application of construction rule or previously proven construction.

Example:

<u>Construction for I,12</u>
Let *q* be a point on the opposite side of *L* from *p*.
Let α be the circle with center *p* and radius *pq*.
Let *a* and *b* be the points of intersection of α and *L*.
Let *d* bisect the segment *ab*.
Let *M* be the line joining *p* and *d*.

Geometric constructions in *E*

What they are:

Sequence of those steps in proof which result from application of construction rule or previously proven construction.

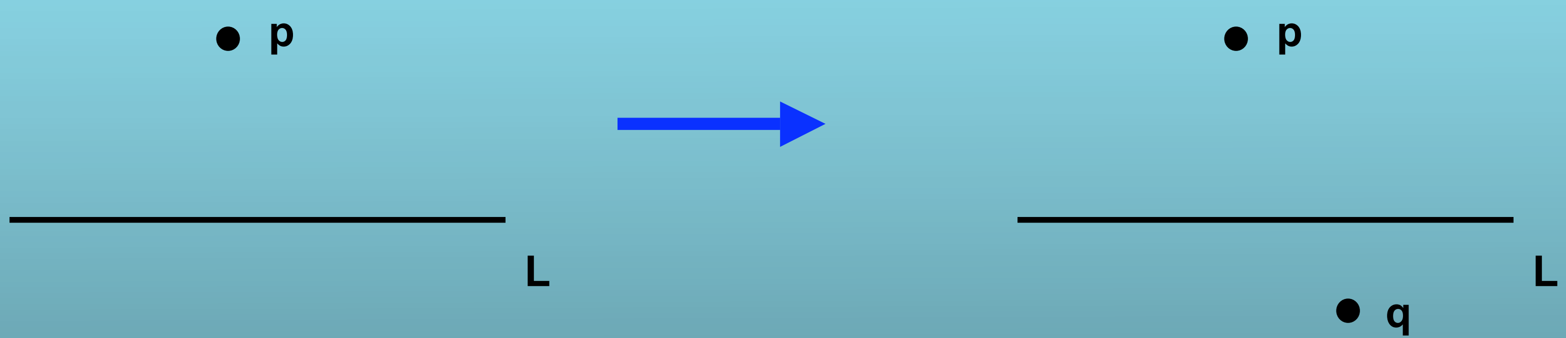
Most such steps produce a unique geometric object. One kind of step, however, does not.

Example:

Construction for I,12
Let *q* be a point on the opposite side of *L* from *p*.
Let α be the circle with center *p* and radius *pq*.
Let *a* and *b* be the points of intersection of α and *L*.
Let *d* bisect the segment *ab*.
Let *M* be the line joining *p* and *d*.

Geometric constructions in *E*

The rules behind these steps license the **free** choice of a point within a region defined by a line or circle. The point is not the result of determinate or exact procedure.



First step in I,12: choice of point q on side of L opposite from p.

The steps are constructive in the sense that the point was not part of the proof's initial configuration (i.e. it is an object added in the proof's construction stage).

Examples in the Elements: I,5 I,9 I,11 I,12 III,1

Geometric constructions in *E*

These proof moves fit naturally into an account where Euclid's diagrams are understood as proof symbols which convey non-metric positional information.

By such an account, the way to express the condition that a point lies in a given region is to place a mark in the representative region in a given diagram.

QUESTION

How does this picture relate to the idea that Euclid's proofs are constructive?

Geometric constructions in *E*

<u>Bernays' Hilbert/Euclid contrast</u>

Hilberts's axiom I,1: $\forall A,B\ \exists a\ C(a, A)\ \&\ C(a,B)$

For every two points *A* and *B*, there exists a line *a* that contains each of the points *A*, *B*.

Euclid's first postulate: $D: P \times P \rightarrow L$

To draw a straight line from any point to any point.

Geometric constructions in *E*

What licenses Euclid's free choice of points seems best represented by a ∀∃ sentence

$\forall\ (x_1 \ldots . x_k)$ defining region R $\exists$ point p Contains(R, p)

to distinguish it from constructions producing exact, unique objects.

General Methodological Mix in Euclid

Constructions	Relations
Postulates	Common notions
Problems	Theorems
Construction stage	Demonstration stage

E's picture of Euclid

Way *E* analyzes this mix

The inexact spatial relations exhibited by diagrams taken as basic.

Constructions understood to occur against a background where such spatial relations obtain.

The point of such constructions is to produce an object with exact relations to given objects within the spatial background.

The free choice of points is attributable to the presumption of this spatial background. The existence/construction of points satisfying inexact relations to given objects is immediate and trivial.

References

J. Avigad, E. Dean, J. Mumma, *A Formal System for Euclid's Elements*, to appear in the Review of Symbolic Logic.

K. Manders, *The Euclidean Diagram,* in Paolo Mancosu, editor, The Philosophy of Mathematical Practice, 2008. Manuscript first circulated in 1995.

J. Mumma, *Proofs, Pictures, Euclid,* to appear in Synthese.

S. Negri, *Contraction-free sequent calculi for geometric theories with an application to Barr's theorem,* Archive for Mathematical Logic, 42, 2003.

A. Tarski, *What is elementary geometry?,* in L. Henkin, P. Suppes and A. Tarski, eds. The Axiomatic Method, with Special Refernece to Geometry and Physics, 1959.