Why Did Geometers Stop Using Diagrams?

Jeremy Heis
University of California, Irvine

PMA Workshop
Notre Dame
26 October 2012
Indeed, if geometry is to be really deductive, the deduction must everywhere be independent of the meaning of geometrical concepts, just as it must be independent of the diagrams; only the relations specified in the propositions and definitions employed may legitimately be taken into account. During the deduction it is useful and legitimate, but in no way necessary, to think of the meanings of the terms; in fact if it is necessary to do so, the inadequacy of the proof is made manifest. (Pasch 1882)

Be careful, since [the use of figures] can be misleading. A theorem is only proved when the proof is completely independent of the diagram. (Hilbert 1894)
Popular but Inadequate Reasons to Reject Diagrammatic Proofs

1. Diagrams illicitly generalize from a particular case to all cases.

2. Diagrammatic proofs mislead even in elementary geometry. E.g. Klein’s “proof” that all triangles are isosceles.

3. Diagrammatic proofs mislead in analysis. E.g. Weierstrass’s continuous but nowhere differentiable curves.
Popular but Inadequate Reasons to Reject Diagrammatic Proofs

1. Diagrams illicitly generalize from a particular case to all cases.

2. Diagrammatic proofs mislead even in elementary geometry. E.g. Klein’s “proof” that all triangles are isosceles.
   • Reply: These “proofs” violate principles of diagrammatic reasoning -- e.g. Manders’s distinction between coexact and exact properties of figures -- that any competent geometer could detect.

3. Diagrammatic proofs mislead in analysis. E.g. Weierstrass’s continuous but nowhere differentiable curves.
   • Reply: These phenomena from analysis don’t explain why geometers would stop using diagrams (Mancosu 2005, Manders 2008).
So Why Did Geometers Stop Using Diagrams?

An answer should cite issues local to geometry: no global problem with diagrams, or problems local to some other area of mathematics.

Mancosu 2005: “These attitudes towards diagrammatic reasoning and visualization have thus a complex history, which still calls for a good historian. Certainly one would have to take into account the importance of the development of projective and non-Euclidean geometries in nineteenth century and the arithmetization of analysis.”
So Why Did Geometers Stop Using Diagrams?

Candidates:

1. Mathematicians want methods that work in every area of mathematics (Pasch 1882).
2. The emergence of Non-Euclidean geometry. (No.)
3. The study of spaces with more points: $n$-dimensional complex projective spaces. (No.)
4. Duality Principles. (No: more on this below.)
5. Independence proofs require reasoning that is independent of both diagrams and the meanings of terms. (Yes, but too late.)
Why Geometers Stopped Using Diagrams: My Hypotheses

• There were in fact local reasons why geometers stopped using diagrams in proofs.

• The limitations of diagrammatic reasoning became clear to geometers in the debate over analytic and synthetic methods in geometry.

• These debates led to the development of kinds of diagrammatic methods that are fundamentally different from the ancient methods studied by philosophers.

• The inferential limitations of even these newer, more powerful kinds of diagrammatic methods became patent in the 1830s with questions about duals of curves of degree 3 or higher.

• Diagrammatic reasoning could not represent general properties of such curves and indeed could even lead into outright error.
Outline of the Talk

I. The limitations of classical diagrammatic methods.
   • Case Study: The theorem that 5 points determine a conic (Apollonius).

II. How algebraic methods overcome these limitations
   • Case Study: The theorem that 5 points determine a conic (Maclaurin).

III. Poncelet’s new diagrammatic method
   • Case Study: The theorem that 5 points determine a conic (Steiner).

IV. The limitations of Poncelet’s new diagrammatic method
   • Case Study: Plücker’s Resolution of the Duality Paradox
Diagrams are “intensional”: the same diagram can exhibit many concepts.

(Just as many words can denote the same referent.)
Euclid I.1: To construct an equilateral triangle.
Euclid I.5: All isosceles triangles have equal angles at their bases.
Euclid I.32: All triangles have interior angles adding to two right angles.
• In Kantian jargon, we say that the same figure “exhibits” the concepts <equilateral triangle>, <isosceles triangle>, and <triangle>.

• Why can the same figure stand in for all triangles, all equilateral triangles, and all isosceles triangles?
Clearly, the same figure falls under all of these concepts.

But this is not sufficient. What’s needed is that any figure that falls under <triangle> in the diagram for I.32 should count as the “same” diagram as ABC. (Netz 1999)

A diagram exhibits a concept iff every diagram that falls under that concept is “equivalent” to it.
The limitations of classical diagrammatic reasoning

• Mumma 2010 makes “equivalence” precise: Two labeled diagrams are equivalent if they agree on whether:
  – A point is or is not on a line or circle
  – Two points are on the same or different sides of a line
  – A point is inside or outside a circle
  – Two lines intersect or don't intersect
  – A line intersects a circle in two points, 0 points, or is tangent
  – Two circles intersect in two points, 0 points, or are tangent

• Can a figure exhibit every concept it falls under? No.
• Does every concept with non-empty extension have a figure that exhibits it? No.
My example of a concept that cannot be exhibited by a figure in classical diagrammatic reasoning is <conic section>.
Conic Sections as sections of cones:
The section is an **ellipse** when $\angle BED > \angle BAC$.
Conic Sections as sections of cones:
The section is a parabola when $\angle BEG = \angle BAC$. 
Conic Sections as sections of cones: The section is a \textit{hyperbola} when $\angle \text{BEG} < \angle \text{BAC}$. 
Apollonius’ *Conics*:
A **parabola** is a curve such that $LM^2 = EM \cdot EH$

$$y^2 = px$$

$$EH = \frac{BC^2 \cdot EA}{BA \cdot AC}$$
Apollonius’ *Conics*:

A hyperbola is a curve such that \( LM^2 = EM \cdot EH + \text{rectangle } HX \)

\[
y^2 = px + \frac{p}{ED} x^2
\]

\[
EH = \frac{BK \cdot KC \cdot DE}{AK^2}
\]
Apollonius’ *Conics*:
An ellipse is a curve such that $LM^2 = EM \cdot EH - \text{rectangle } HX$

$$y^2 = px - \frac{p}{ED} x^2$$

$$EH = \frac{BK \cdot KC \cdot DE}{AK^2}$$
Notice: Ellipses are defined in terms of the diameter ED and the point K.

\[ y^2 = px - \frac{p}{ED} x^2 \]

\[ EH = \frac{BK \cdot KC \cdot DE}{AK^2} \]
But ED does not exist in a parabola.

\[ y^2 = px \]

\[ EH = \frac{BC^2 \cdot EA}{BA \cdot AC} \]
And ED is outside the cone ABC in the case of a hyperbola.

\[ y^2 = px + \frac{p}{ED}x^2 \]

\[ EH = \frac{BK \cdot KC \cdot DE}{AK^2} \]
Similarly, the point K and line AK are inside the cone ABC for a hyperbola, but outside for an ellipse.

\[ y^2 = px + \frac{p}{ED} x^2 \]

\[ EH = \frac{BK \cdot KC \cdot DE}{AK^2} \]
Inequivalent Diagrams

• The diagrams of ellipses, hyperbolas, and parabolas are inequivalent.
• So no diagram could exhibit <conic section>.
• Netz 1999: "the overwhelming rule in Greek mathematics is that propositions are individuated by their diagrams."
• So there cannot in general be a proposition of the form “For any conic section, ....”
• This is illustrated by Apollonius’s failure to prove that any five points determine a conic.
Heath on Why Apollonius Never Proved that 5 Points Determine a Conic

Since Apollonius was in possession of a complete solution of the problem of constructing the four line locus referred to the sides of a quadrilateral of any form, it is clear that he had in fact solved the problem of constructing a conic through five points… The problem of the construction of a conic through five points is, however, not found in the work of Apollonius … The explanation of the omission may be that it was not found possible to present the general problem in a form sufficiently concise to be included in a treatise embracing the whole subject of conics. This may easily be understood when it is remembered that, in the first place, a Greek geometer would regard the problem as being in reality three problems and involving separate construction for each of the three conics, the parabola, the ellipse, and the hyperbola.
Heath on Why Apollonius Never Proved that 5 Points Determine a Conic

He would then discover that the construction was not always possible for a parabola, since four points are sufficient to determine a parabola; and the construction of a parabola through four points would be a completely different problem… Further, if the curve were an ellipse or hyperbola it would be necessary to find a *diorismos* expressing the conditions that must be satisfied by the particular points in order that the conic might be one or the other. If it were an ellipse, it would have been considered necessary to provide against its degeneration into a circle.
Section II: How algebraic methods overcome these limitations

- The limitations of diagrammatic reasoning were recognized in the modern period. These limitations were avoided in algebra.
- Poncelet 1822: “Algebra uses abstract signs; it represents absolute magnitudes by characters which by themselves have no value, and [thus] leave the magnitudes as undetermined as possible. Consequently, algebra necessarily operates upon, and reasons about, signs of non-existence just as it does with quantities which are always absolute and real. For example, if a and b represent two arbitrary quantities, it is not possible to take their relative magnitude into account during the computation; one is forced, willy-nilly, to reason about expressions such as a – b and \( \sqrt{(a - b)} \) as if they were quantities which were always absolute and real. Thus, the result must have that same generality, and extend to all possible cases, to all possible values of the letters which occur in it.”
MacLaurin’s proof (*Treatise on Algebra*, 1748) that any 5 points, no two collinear, determine a conic.

First, MacLaurin proved that any conic is represented by a second degree equation, and that any (non-degenerate) secondary equation represents a conic.
Take 5 points D, P, S, E, and C.

Connect CS, CP, and PS.

Draw an arbitrary line $l$ through E, meeting CS in A, CP in N.

Draw DN, meeting PS in Q.
Draw lines through D, P, and Q parallel to EA, meeting CS.
Draw EQ, meeting CS in B, and PH parallel to EQ and meeting CS in H.

Let CM = \(x\) and MP = \(y\).

Label CS = \(a\), CA = \(b\), SB = \(c\), DF = \(k\), AF = \(l\), AE = \(d\), BE = \(e\), and AB = \(f\).
Maclaurin then shows that

\[
\begin{align*}
[(bf(c + l + f))y^2 + & [c(kf - ld) + bd(f + d) + kf^2]xy \\
- & [bad(l + f)]y + [kfad]x - [kfd]x^2 = 0
\end{align*}
\]

Since this is a second degree equation, the variable point P (which changes position when \( l \) is modified) traces a conic.
Maclaurin then shows that

\[
[(bf(c + l + f))y^2 + [c(kf - ld) + bd(f + d) + kf^2]xy
\]

\[
- [bad(l + f)]y + [kfad]x - [kfd]x^2 = 0
\]

Since this is a second degree equation, the variable point P (which changes position when \( l \) is modified) traces a conic.
The algebra doesn’t care whether, say, $a > b$ (CS > CA): that is, it does not care whether A is between C and S. The fact that different values for the constants correspond to inequivalent diagrams is irrelevant.
Section III: Poncelet’s new diagrammatic method

• The project of Poncelet 1822 to find new diagrammatic methods so that “ordinary geometry would turn out to rival analytic geometry in many respects.”

• Two innovations in diagrammatic method:
  1. extended the equivalence relation so that two figures are equivalent if they are projections of one another or can be derived from one another by “continuous” motion
  2. allowed the proof to exploit special features of the drawn figure.

• Innovation 1 has been widely discussed. #2 has not been.
Poncelet proved general results by exploiting the particular features of special cases

Poncelet 1822, §99: When we wish to establish a certain property regarding a given figure, it is sufficient to prove that the property exists in the case of any one of its projections. Among all the possible projections of a figure, some one may exist that may be reduced to the simplest conditions and from which the proof or the investigations which we have proposed may be made with the greatest ease. It may require but a brief glance or at most, the knowledge of certain elementary properties of geometry to perceive it or to know it. For example, to take a particular case, suppose the figure contains a conic section, this may be considered the projection of another [figure] in which the conic section is replaced by the circumference of a circle and this single statement is sufficient to shift the most general questions regarding conic sections to others that are purely elementary.
Poncelet proved general results by exploiting the particular features of special cases

Coolidge (History of Geometrical Methods): The first writer to have a definite idea of using a transformation to find the properties of a general figure from the simpler properties of a special one was Poncelet. …[T]he procedure is as follows:

1. Simplify the figure by projection.
2. Prove a theorem for the simplified figure.
3. State the result in a projectively invariant form.
Steiner’s Proof (1832) that any 5 points determine a conic

Euclid III.21: In a circle the angles in the same segment equal one another.

If corresponding lines of two pencils make equal angles, then the two pencils are projective. (Steiner theorem §13.II)

Conclusion: the pencils of lines through two points $P$ and $P'$ on the circumference of a circle are projective (theorem §37)
The pencils of lines through two points $P$ and $P'$ on the circumference of a circle are projective (theorem §37).

Since any two sections of the same cone are projectively related (theorem §36),
The pencils of lines through two points $P$ and $P'$ on the circumference of a circle are projective (theorem §37)

Since any two sections of the same cone are projectively related (theorem §36), a conic can be generated by two projective pencils of lines through two points $P$ and $P'$. (theorem §38.IV)
A projectivity is given when any three pairs of projective elements are given. (theorem §10.β)

So, a conic is given by five points. (§41.I)
Poncelet proved general results by exploiting the particular features of special cases

• This feature of Poncelet’s and Steiner’s method violated fundamentally a key feature of classical diagrammatic methods.

• Kant, *Critique of Pure Reason*, A713/B741: “The individual drawn figure is empirical, and nevertheless serves to express the concept without damage to its universality, for in the case of this empirical intuition we have taken account only of the action of constructing the concept, to which many determinations, e.g. those of the magnitude of the sides and the angles, are entirely indifferent, and thus we have abstracted from these differences, which do not alter the concept of the triangle.”
Poncelet proved general results by exploiting the particular features of special cases

- Poncelet 1822 (also Carnot 1801, 1803; Chasles 1837) described the levels of generality in a diagrammatic proof:

  1. **The physical, drawn object**: has non-geometric properties, like color, weight, temporal location.

  2. **Figure**: #1, abstracting away from all non-geometric properties.

  3. **States of a system**: #2, abstracting away from geometric properties that differ among figures that are “equivalent” in the classical sense from Mumma 2010.

  4. **The system**: #3, abstracting away from properties that change under projection or continuous deformation.

- Classical diagrammatic reasoning infers from **figure** to **state**.
- Poncelet added an extra layer, from **state** to **system**.
Poncelet’s extended notion of “equivalent” figures

• Two figures are equivalent if they are projections of one another
  • A diagram of a conic is equivalent to a diagram of a circle
  • A line at finite distance to a line at infinity.

• Two figures are equivalent if one can be derived from the other under a continuous motion, maintaining the “relevant” properties
  • A conic C and intersecting line $l$ are equivalent to a circle and $l_{\infty}$.
  • Any two conics are projectively equivalent to two circles.
  • Any two doubly tangent conics are equivalent to two concentric circles.
Poncelet, *Traité*: Any conic and line can be projected to a circle and the line at infinity.

Obvious when the conic and line do not intersect. What if they do intersect?
Poncelet claimed that theorems proved for a system with a conic and an intersecting line hold also for a system with a conic and a non-intersecting line.

He reasoned by the principle of continuity: “if in keeping the same given properties, we begin to vary the original figure by insensible degrees, or if we subject certain parts of this figure to a continuous motion of any sort, is it not evident that the properties and relations found for the first system remain applicable to successive stages of this system?”
A chord PP’ of a central conic like an ellipse has the property

\[ QP^2 = \gamma (QO)(QO') \]

where OO’ is the diameter that bisects PP’ at Q.

As Q moves continuously past O’, there will still be segments RR’ that have this property. Poncelet calls these segments “ideal chords” of the conic.
The points $R, R'$ trace out a new conic, the “supplementary hyperbola,” whose real chords are ideal chords of the original conic, and whose diameter and center are the same as the original conic’s.

Thus any line intersects any conic, forming either a real or ideal chord.
So any conic and line can be projected to a circle and the line at infinity.

If the conic and line intersect, then their common chord is real, but not ideal.
Suppose that the line forming an ideal chord with our conic is $l_\infty$. 
Suppose that the line forming an ideal chord with our conic is $l_\infty$.

Then CR and CR' are the asymptotes of the supplementary hyperbola.
Suppose that the line forming an ideal chord with our conic is $l_\infty$.

Then CR and CR’ are the asymptotes of the supplementary hyperbola.

R and R’, the imaginary infinite points that lie on the conic, will move around $l_\infty$ as the direction of the asymptotes change.
R and R’ have two degrees of freedom:

1. They change position as the ratio of the major and minor axes of the conic changes. (Moving from one ellipse to a different ellipse dissimilar to it.)
R and R’ have two degrees of freedom:

1. They change position as the ratio of the major and minor axes of the conic changes.

2. They change position as we consider different conjugate axes of the same ellipse.
But if the conic is a circle, these degrees of freedom are eliminated: every circle has the same ratio of diameters (1:1), and is symmetric under rotation.

So R and R’ (“the circular points”) will be the same imaginary infinite points for any circle, and there is an ideal chord RR’ common to every circle.
Poncelet used the circular points to prove some very surprising projective equivalences. Such as:

*Any two conics are projectively equivalent to two circles.*

**Proof.** Consider the special case of two conics with one real common cord and one ideal chord I’.

Project I’I onto \( l_\infty \) and one of the conics onto a circle. Since I and I’ are the intersection points of a circle with \( l_\infty \), they are the circular points.

Since I’ and I are common to both conics, the projection of the second conic also passes through the circular points and so is also a circle.
Any two tangent conics are projectively equivalent to two concentric circles.

Proof. By our previous result, the two conics are projectively equivalent to two circles. Since they are tangent, they have only 2 intersection points and only one common chord, and since their projections are circles, RR’ must be projected onto the circular points.
Since the two circles are tangent to one another at $R$ and $R'$ on $l_\infty$, their supplementary hyperbolae have common asymptotes.

Since asymptotes of supplementary hyperbolae pass through the centers of conics, the two circles are concentric.
Pasch, 1882 (again): “If, however, a theorem is rigorously derived from a set of propositions—the basic set—the deduction has value which goes beyond its original purpose. For if, on replacing the geometric terms in the basic set of propositions by certain other terms, true propositions are obtained, then corresponding replacements may be made in the theorem; in this way we obtain new theorems as consequences of the altered basic propositions without having to repeat the proof.”

Ernst Nagel (1939) famously argued that projectively duality undermined diagrammatic forms of reasoning and motivated the modern idea that the validity of deductive arguments are independent of the meanings of terms.
In projective geometry, theorems come in pairs:

Pascal’s theorem:

*If the vertices of a hexagon lie on a conic, then the points intersecting the lines connecting the vertices are collinear.*
In projective geometry, theorems come in pairs:

Brianchon’s theorem:

*If the sides of a hexagon lie on a conic, then the lines connecting the points of intersection of the sides are copunctual.*
Against Nagel’s explanation

• His argument: the reasoning in the proofs of dual theorems is essentially the same despite the fact that the diagrams are different. So the reasoning must be independent of the diagram.

• Nagel’s explanation for why geometers stopped using diagrams in proofs has the advantage of being local to geometry.

• But this argument would cut no ice with Poncelet’s diagrammatic method. If a diagram of two concentric circles can stand in for every two tangent conics, then there is no obstacle at all to a geometer transforming a diagram of a point conic into its envelope of tangent lines.
Poncelet’s Trouble with Duality

• There is no in principle difficulty posed to diagrammatic reasoning by duality.

• Instead, the problem is that a certain kind of duality (duals to curves of degree 3 or higher) eludes diagrammatic reasoning:
  – the properties of dual curves to general higher plane curves cannot be represented diagrammatically
  – diagrammatic reasoning can lead into fallacies when reasoning about curves dual to higher plane curves
The dual of a curve $C$ is formed by taking the points $P$, polars with respect to an arbitrary chosen conic $C'$, of tangent lines $p$. 
The dual of a curve $C$ is formed by taking the points $P$, polars with respect to an arbitrary chosen conic $C'$, of tangent lines $p$. 
Incidence is maintained on polars:

If A is on p, then P is on a.

So if a line l meets a curve in n points $P_1,...,P_n$, then the polar point L is the intersection of n lines $p_1,...,p_n$ tangent to the dual curve.
The Paradox of Duality

• Plücker (1835) proved that for any curve $C_n$ of degree $n$, the number of tangents to $C$ from a point $P$ (the “class” of $C_n$) is:

$$k = n(n - 1)$$

• Since incidence is maintained when taking polars, the class of a dual curve = the degree of the original curve.
• And the dual of the dual of a curve $C_n$ should also have degree $n$.
• But the dual of a dual has degree:

$$n(n - 1)[(n(n - 1) - 1]$$

• And, unless $n = 2$, this won’t equal $n$. 
Plücker’s Solution

• Cayley called the discovery of Plücker's equations "the most important one beyond all comparison in the entire subject of modern geometry."

• Felix Klein: "While the French geometers had generally restricted themselves to the linear-quadratic domain, and Poncelet had run into difficulties in the first attempts at going beyond it, Plücker was now able to make the first successful advance in the general theory of algebraic plane curves. I would say that his principal achievement was the "Plücker formulae," which connect the order $n$ of a curve (the degree of its equation in point coordinates) with its class $k$ (the degree of its equation in line coordinates) and its simple singularities.”
Plücker’s Solution

• The attempt to solve this paradox was a test of the power of the new diagrammatic methods to reason consistently about curves of degree 3 or higher. They failed.
• Algebraic methods for doing the projective geometry of Steiner and Poncelet succeeded.
• Coolidge: “[Plücker] had an unshakeable belief that for most purposes, algebraic methods were infinitely preferable to the purely geometric ones recently brought into fashion by Poncelet and Steiner. He went a good way towards proving the correctness of his belief.”
A double point P is a point where every line through P meets C twice.

Geometrically, at the point two branches of the curve cross.
A double point P is a point where every line through P meets C twice.

Geometrically, at the point two branches of the curve cross.
Algebraically:

\[ A + (B\cos\theta + C\sin\theta)\rho + \left(D\cos^2\theta + E\cos\theta\sin\theta + F\sin^2\theta\right)\rho^2 + \ldots = 0 \]

Suppose \(A=B=C=0\). Then the first degree term will vanish for any value of \(\theta\), which means that the equation will have two roots at the origin for every value of \(\theta\), and every line through the origin intersects the curve twice at the origin.
Plücker’s Equations

\[ k = n(n-1) - 2d - 3r \]
\[ n = k(k-1) - 2t - 3w \]
\[ w = 3n(n-2) - 6d - 8r \]
\[ r = 3k(k-2) - 6t - 8w \]

n = degree
k = class
d = double points
r = cusps
t = double tangents
w = points of inflection
Why are dual curves of degree 3 or higher resistant to diagrammatic reasoning?

• Some singular points are indetectible to diagrammatic methods. Salmon 1852 mentions triple points where one tangent is real and two are imaginary: “The last kind of triple point is worth notice, since to the eye it does not appear to differ from any other point on the curve.” [Salmon [1852], p. 30]

• Plücker’s formulae hold for any algebraic curve of any degree $n$. He could prove these formulae algebraically because there are ways of representing polynomials of arbitrary degree:

\[ A + (B\cos\theta + C\sin\theta)\rho + \left(D\cos^2\theta + E\cos\theta\sin\theta + F\sin^2\theta\right)\rho^2 + \ldots = 0 \]

• There is no diagrammatic way to represent a curve of arbitrary degree.
Why Geometers Stopped Using Diagrams: My Hypotheses

• There were in fact local reasons why geometers stopped using diagrams in proofs.

• The limitations of diagrammatic reasoning became clear to geometers in the debate over analytic and synthetic methods in geometry.

• These debates led to the development of kinds of diagrammatic methods that are fundamentally different from the ancient methods studied by philosophers.

• The inferential limitations of even these newer, more powerful kinds of diagrammatic methods became patent in the 1830s with questions about duals of curves of degree 3 or higher.

• Diagrammatic reasoning could not represent general properties of such curves and indeed could even lead into outright error.