Charles L. Silver
Berkeley, CA
October, 2010

STEINER-LEHMUS AGAIN

ABSTRACT

The intention is to provide a direct proof of the Steiner-Lehmus theorem geometrically, using only compass and straightedge, based entirely on notions found in Euclid's Book I *of* The Elements*. A quick version of the proof is presented first, just to give the idea behind it. This is followed by a more detailed proof for which every single step is justified by principles in* Book I. *However a definition that he uses is missing as well as a Common notion that he also uses. These principles are discussed and employed in the present proof.*

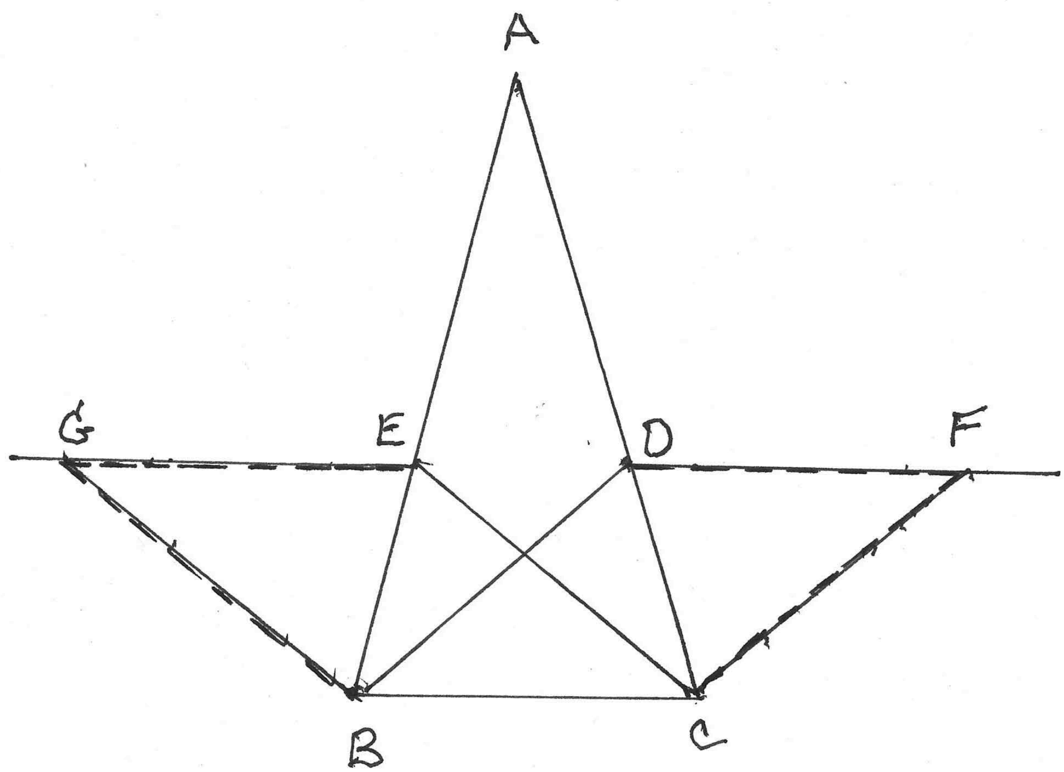
The Steiner-Lehmus theorem has been of interest since 1840 when a schoolteacher named Daniel Lehmus noticed an obvious, elementary fact in geometry that he'd found surprisingly difficult to prove. He wrote a letter to Jacques Sturm, posing the problem to him and asking for a purely geometric proof of it. Sturm passed his request to other geometers, and subsequently the Swiss geometer Jakob Steiner is said to have produced the first proof of it, followed by a proof by Lehmus himself However, it the first proof of the theorem to appear in publication was by the French Mathematician Rougevain in 1842. Besides the indirect proofs using only geometric principles, there are many using trigonometry and/or algebra. But could a direct geometric proof be found—one using only a straight edge and a compass? This itself became controversial. Could such an innocent-looking problem even have a direct, geometric proof?

In 1882 Sylvester offered a proof that a direct proof of the Steiner-Lehmus theorem is impossible. Though Sylvester's impossibility proof was itself found to be incorrect, his conclusion was still accepted that no direct proof of the theorem could be found (for example, a proof using only intuitionistic logic). The famous Princetonian mathematician, John Horton Conway has said that there is no "equality-chasing type" proof of this result. However, it is mysterious what exactly Conway means. It seems that he is denying that there exists an algebraic proof using complex numbers. He concluded his argument against any equality-chasing proof with the comment that if there were one, "it would prove the triangle equilateral, which it isn't. So there's no such proof!" In 1970 a purportedly direct proof was published in a mathematical journal, but the same journal four years later published a refutation of it.

Apparently by now there are over sixty proofs of this "obvious theorem. Below, I provide a compressed proof, which may seem completely correct on the surface, but it's not unless every single step can verified. And the proof attempted here the proof is nor just direct, but every verifying principle can be found only in Book I of Euclid's elements. As the reader will soon see, it is quite difficult to separate off and verify each and every step in the proof of this theorem. Another comment should be mentioned: some persons have proved the "contrapositive" of this theorem, which is simply an indirect proof in disguise.

Stated by Lehmus, the theorem is that if the two internal angle bisectors of a triangle are themselves equal, the triangle itself is isosceles. Without further ado, a compressed proof will be provided first to orient the reader to the ideas in the proof, especially the constructions.

Here is a compressed version of the Steimus-Lehmus proof. Following this compressed version, a detailed proof is provided, where every step of reasoning is justified by a principle found in Book I of Euclid's Elements.



GIVEN: Triangle ABC with angle ABD bisecting angle ABC and angle ACE bisecting angle ACB. And the bisecting lines BD and CE are equal.

TO PROVE: Triangle ABC is isosceles.

PROOF:

From D construct to the right a line parallel to BC which also equals BC. This right line ends at point F.

Similarly, from E construct to the left a line parallel to BC again, also equaling BC. This left line ends at point G. There are now two congruent parallelograms (BDFC and CEGB, (by "application," see below)). Since the opposite angles of two equal parallelograms are equal, the angle at F equals angle CBD. And the angle at G equals angle BCE. Therefore, twice these "half-angles" are equal, making the base angle ABC = angle ACB. Thus, triangle ABC is isosceles. ❑

This marks the end of the intuitive proof. However, the quick proof given above compresses steps, provides incomplete justification of them, and does not show that *every single step* follows from principles of Book I of The Elements.[1] The detailed one below is intended to do this.

First, an assumption and a Common Notion are needed. Euclid nowhere defines a "parallelogram," though he freely uses this notion in propositions 34, 35, and 36. Since to prove proposition 34, Euclid assumes that the opposite sides of a parallelogram are parallel, this must have been the following definition he had in mind:

Def. 22*: "Of quadrilateral figures, a *parallelogram* is that which has its opposite sides parallel". (The asterisk indicates that this definition should be added to Def. 22.)

In Proposition I.4 Euclid says "For if the triangle ABC be applied to the triangle DEF, and the point A be placed on the point D and the straight line AB on DE, then the point B will also coincide with E, because AB is equal to DE." (He does something similar in Prop. I.8.) Euclid has no Common Notions corresponding to "applied to" and "placed on" among those listed. Therefore, since he uses them, a common notion needs to be added to account for them. This common notion is usually called "superposition,"[2] though it's called "application" here.

---

[1] I take Euclid's notions from The Thirteen Books of Euclid's Elements, translated by Sir Thomas L. Heath in The Great Books of the Western World. A slightly different version of Euclid, also translated by Heath, can be found in a Dover edition. The Dover edition is preferable in general, as it is a more modern translation by Heath and contains much interesting supplementary material. However, often its supplementary material is interspersed among Euclid's principles, making them difficult to consult in a single viewing. To follow the proof most conveniently, this internet site may be best: http://aleph0.clarku.edu/~djoyce/java/elements/bookI/bookI.html (though this version differs slightly also).

[2] Euclid's use of "superposition" was first pointed out to me by Alasdair Urquhart (who also introduced me to three articles as well as directing my attention to the formal system of Euclid's elements by Avigad et al. cited below). Commentaries on "superposition" range from the fairly positive to the extremely negative. In the Dover translation of Heath, it is claimed that "It looks as though [Euclid] found this method by tradition (we can hardly suppose that if Thales proved that the diameter of a circle divides it into two equal parts, he would do so by any other method than that of superposition)." Also claimed is the view that "the method [of superposition] can hardly be regarded as being, in Euclid, of only subordinate importance; on the contrary, it is fundamental. Nor as a matter of fact, do we find in the ancient geometers any expression of doubt as to the legitimacy of the method. Archimedes uses it to prove that any spheroidal figure cut by a plane through the centre is divided into two equal parts in respect of both its surface and its volume; he also postulates in *Equilibrium of Planes* I. that 'when equal and similar plane figures coincide if applied to one another, their centres of gravity coincide also'." (p. 225 for all of the above). These remarks are fairly positive on superposition. But see the next comments.

Heath also says, "At the same time, it is clear that Euclid disliked the method and avoided it wherever he could, e.g. in I. 26, where he proves the equality of two triangles which have two angles respectively equal to two angles, and one side of the one equal to the corresponding side of the other."

Common Notion 4.5: (Application[3]) If all features of one object can be applied exactly with all features of another object, then the two objects coincide.[4]

GIVEN: Triangle ABC with angle ABD bisecting angle ABC and angle ACE bisecting angle ACB. And the bisecting lines BD and CE are equal.

TO PROVE: Triangle ABC is isosceles.
TO PROVE: Side AB equals side AC. Thus Triangle ABC is isosceles
TO PROVE: Angle ABC equals angle ACB (Thus by Def'n 20: "Of trilateral figures an *isosceles triangle* is that which has two of its sides alone equal".) [my underline]

PROOF:
1) Construct a circle with center D and distance BC. (By Post. 3: "To describe a circle with any centre and distance.")
2) Through D, draw a straight line parallel to BC (that's exterior to triangle ABC). (By Prop. 31: "Through a given point to draw a straight line parallel to a given straight line." (Euclid often uses "exterior" or some other synonym in his proofs. [See footnote 5.])
3) Point F is the intersection of the circle (by 1) and the straight line (by 2) on the other side of DE.[5]
4) Therefore BC is parallel to DF (by 2) (and BC is equal to DF (by 1)).
5) BD is parallel (and equal) to CF (Prop. 33: "The straight lines joining equal and parallel lines (at the extremities) which are in the same directions (respectively are themselves equal and parallel.")
6) BCFD is a parallelogram. (By Def.* 22.)
7) CBGE is a parallelogram (By the same method as steps 1 through 6 for BCFD.)
8) Parallelogram BCFD is congruent to[6] parallelogram CBGE. (By Def. 4.5, they coincide by "application". This is the order of the application.[7] First apply the point F to

Avigad et al. discuss "Euclid's notorious 'superposition inferences,' which vexed the commentators through the ages." They mention three ways of handling superposition. They then offer a "more elegant solution [than the first two they consider]." They say that "what superposition allows one to do is to act as though one has the result of *doing the constructions*…only for the sake of proving things about objects that are already present…." They continue by indicating that in their formal system, "if you can derive a conclusion *assuming the existence of some new objects*, you can [use superposition as an elimination rule permitting you to] infer that the conclusion holds *without* the additional assumption. [my italics] (pp. 25-26). This method is a familiar and an innocuous one that's used extensively in systems of formal logic. Such proofs are "sound" in the sense that they can be *proved* to be correct. In fact, the entire Euclidean system of Avigad et al. can be proved to be sound. Hence, it seems to me Avigad et al make "superposition inferences" less vexatious (though we refer to such inferences as "application"..

[3] Though this principle is virtually always referred to as "superposition," I define it as "application," since in Euclid's language he "applies" one figure to another.

[4] In Byrne's beautiful book, he defines under "Elucidations": "Superposition is the process by which one magnitude may be conceived to be placed upon another, so as to exactly cover it, so that every part of each shall exactly coincide." (XXV)

[5] The notion of "the other side," one point being "the extremity of, being on the other side" is taken for granted by Euclid in several propositions in Bk I. The phrase "on the other side of" is used specifically in the proof of Prop. 12. (However, even more equivalent notions such as "contained,"—therefore "not contained," "opposite," etc. can also be found in proofs in BK I.)

[6] Euclid uses "equals" which we now replace with the more modern "is congruent to" for polygons.

G and apply the straight line FD to GE (By Prop. 31, they are both parallel and equal to BC, thus they coincide. Then the applied point D will coincide with E because both lines are equal and parallel to BC. (This is almost word for word the same as in the proof of Prop. I.4.) Next, place the point C on B and B on C (i.e., just reverse the names of the extremities of the points on the base line.) Then the applied point B will fall on the original C, and C will fall on B. Then the line from (the applied) point F to point C coincides with the original line from G to B (Prop 33 again: "The straight lines joining equal and parallel straight lines (at the extremities which are in the same directions (respectively) are themselves equal and parallel"). Likewise, apply the applied line DB on line EC. Now, all points and lines of parallelogram BCDF coincide with parallelogram BCFD. Thus, the parallelogram BCFD coincides with CBFD. Thus, the parallelogram BCDF is congruent to CBGE (by CN 4. "Things which coincide with one another are congruent ("equal") to one another.) "And the remaining angles will also coincide with the remaining angles and will be equal to them" (quoting a line in the proof of Prop. I. 4) Hence, angle CBD equals angle BCE.
But angle BCD is half of angle ACB (BD bisects the angle ACB). Similarly, angle BCE bisects angle ABC.
10) Thus, angle ABC equals angle ACB (CN 2: "If equals be added to equals, the wholes are equal".
11) Thus, side AB equals side AC (by Prop. 6: "If in a triangle two angles be equal to one another, the sides which subtend the equal angles will also be equal."

Therefore, etc. QED

Note, however, we have not proved the "alone" condition, since according to Euclid's definition of *isosceles triangle* ABC *may not be* equilateral as well. Using this definition, special precautions could be taken to ensure that the legs of the triangle each differ in length from the base, but it seems preferable simply to accept the standard notion that an isosceles triangle does not *exclude* it from being equilateral, rather than asserting that an isosceles triangle must have *only* two sides equal.

Bibliography

Byrne, Oliver (The First Six Books of The Elements of Euclid, (William Pickering, London, 1847);
Coxeter, H. S. M., and Greitzer, S. L. (Geometry Revisited; Mathematical Association of America, 19)
Euclid's Elements, translated by Sir Thomas Heath (The Great Books of the Western World Encyclopedia Britannica, 1952);
Euclid, The Thirteen Books of The Elements, translated with introduction and commentary by Sir Thomas L. Heath (Second Edition, Dover, New York, 1956).
Henderson, Archibald, *A Classic Problem In Euclidean Geometry*, *A Basic Study*

[7] Because of the imprecision in Euclid's use of this method, it is difficult to determine the preferable order of its application. Putting in the base, then its parallel and equal line may be best.

(University of North Carolina, Chapel Hill, NC., Year ?)

Henderson, Archibald, *The Lehmus-Steiner-Terquem Problem in Global Survey* ed. By Nathan Altshiller Court, Notes on Famous Questions, (Scripta Mathematica Vol 21 1955);

URL Internet: http://aleph0.clarku.edu/~djoyce/java/elements/bookI/bookI.html , D.E. Joyce (Clark University 1996)

Acknowledgements

I'd most like to acknowledge Alasdair Urquhart for informing me about details of the Steiner-Lehmus Theorem, for providing many manuscripts concerining it, and for helpful discussion in general.

I appreciate Jeremy Avigad's explanation of first using "superposition" within his formal system and then discharging it later.

I'd like to thank John Horton Conway for many instructive discussions and for his very neat, indirect "scissors proof" of Steiner-Lehmus.

Thanks to Iddo Patt for the beautiful Byrne edition of Euclid.

I'd also like to thank Mrs. Cynthia Law for re-introducing me to Geometry.