Newton’s Quadrilateral Theorem

Waldemar Pompe

Abstract. We discuss a version of Newton’s quadrilateral theorem, where the circle is replaced by a hyperbola. We present two geometric proofs of the theorem in that case. One of them originates from Newton himself.

1. Introduction

Newton’s famous theorem states that if a circle can be inscribed in a convex quadrilateral $ABCD$, then the center of the circle lies on the line joining the midpoints of the diagonals AC and BD of the quadrilateral.

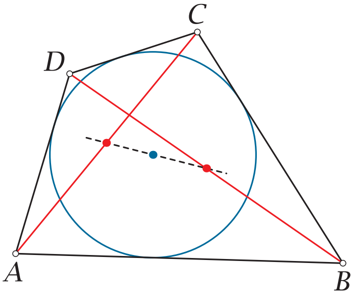

Fig. 1

One of the proofs of this theorem (probably the most famous one) uses the properties of signed areas (we will present this proof later in the text). By looking at it, we easily come to the conclusion that this theorem can be formulated and proved in the same way for any quadrilateral $ABCD$, convex, concave, and even crossed: as long as we assume that there is a circle tangent to the lines AB, BC, CD, DA, its center lies on the line connecting the midpoints of the segments AC and BD.

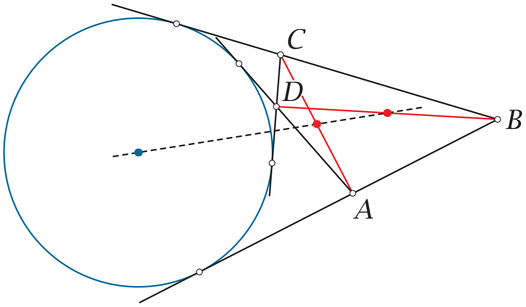

Fig. 2

Moreover, since any affine transformation preserves midpoints, Newton’s theorem can be immediately generalized to the case of an ellipse. The question arises: will Newton’s theorem remain true if we assume that the lines AB, BC, CD, DA are tangent to a certain hyperbola? Specifically, does the center of the hyperbola then lie on the line connecting the midpoints of AC and BD?

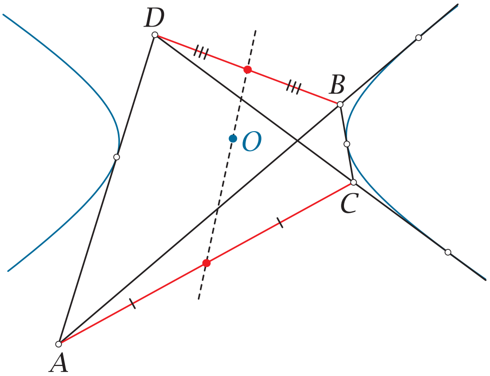

Fig. 3

Surprisingly, it turns out this is true, despite the fact that a hyperbola cannot be transformed affinely into a circle. Also, the use of a projective transformation, which maps a hyperbola onto a circle, will not bring this configuration to Newton's theorem with a circle, since the midpoints of AC and BD will not be preserved under such a transformation.

Newton's theorem with a circle is well-known and can be found in many sources (see for example [1] or Google), but its above version with a hyperbola is (almost) nowhere to be found. Internet search engines only point to one recent paper [3]. It contains a calculational proof of Newton's theorem for an ellipse or a hyperbola (see Conjecture 1.1 and Theorem 4.3 in [3]) and it refers to only one source [2] — a video on a YouTube channel — where the theorem for an ellipse or a hyperbola was formulated, but not proved. To the best knowledge of R. Kaldybayev — the author of [3], his paper [3] is the first to prove Newton's theorem for an ellipse and a hyperbola.

In search of a source and information about the version of Newton's theorem with a hyperbola, I initiated a discussion [5] on this topic in the Facebook group *Romantics of Geometry*, which brings together over 18,000 geometry enthusiasts from around the world. In response, based on Wikipedia, Fedor Bakharev pointed out that Newton proved his theorem for *any* conic, but no reference to Newton's proof was given on Wikipedia. The breakthrough information was provided by Vladimir Dubrovsky that Newton published his proof in his famous work *Principia* [4]. It turns out that Newton's book [4] is available on the Internet and I have verified that it indeed contains a proof for an ellipse and a hyperbola ([4] Cor. 3, page 146)!

In this paper, I will present two geometric proofs of Newton's theorem for an ellipse or a hyperbola. One is my own, while the second is based on Newton's original idea. Thanks to the use of theorems that are now widely known, I was able to simplify and write Newton's proof in a very concise way.

1. Signed Areas

The theory presented in this chapter is classical and well-known. I include it here for the convenience of the reader.

To any triangle in the plane we may assign a positive or a negative orientation, depending on the order we list its vertices in. Namely, we say that triangle ABC is *positively oriented*, if the order $A \to B \to C \to A$ determines a counterclockwise orientation in the plane (Fig. 4); otherwise we say that triangle ABC is *negatively oriented* (Fig. 5).

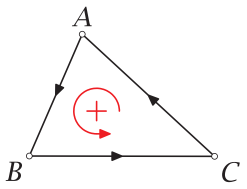

Fig. 4

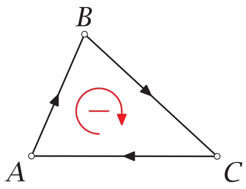

Fig. 5

If triangle ABC is positively oriented, so are triangles BCA and CAB, while triangles BAC, ACB, and CBA are then negatively oriented.

To each oriented triangle ABC we assign a *signed area*, which is simply the area of ABC, if ABC is positively oriented, or minus the area of ABC, if triangle ABC is negatively oriented. We will denote the signed area of oriented triangle ABC by $S(ABC)$. If points A, B, C are collinear, then we set $S(ABC) = 0$. Therefore, for any triangle ABC we have $S(ABC) = S(BCA) = S(CAB)$, while $S(ABC) = -S(BAC)$.

Sometimes the signed area is more convenient to use than the traditional area. For example, *if ABC is any triangle and X is any point in the plane, then*

$$S(ABC) = S(ABX) + S(BCX) + S(CAX).$$

Note that an analogous formula for the traditional area is true only if X lies inside triangle ABC. An immediate induction yields now the following

Theorem 1.1.

Let $A_1, A_2, \dots, A_n$ be arbitrary points. Then the sum

$$\sum_{k=1}^{n} S(A_k A_{k+1} X)$$

does not depend on the choice of point X. (We assume $A_{n+1} = A_1$.)

This theorem allows to define a signed area for any oriented closed broken line with consecutive vertices $A_1, A_2, \dots, A_n$. Namely, we set

$$S(A_1A_2\dots A_n) = \sum_{k=1}^{n} S(A_kA_{k+1}X),$$

where X is an arbitrary point. Note that the broken line may be self-intersecting.

In the proof of the next theorem we shall use another useful property of the signed area, which fails for the traditional area. Namely, let $A \neq B$ be points and let a be a real number. Then *the set of all points X in the plane, such that $S(ABX) = a$ is a line parallel to line AB.*

Theorem 1.2.

Let A, B, C, D be points such that $ABCD$ is not a parallelogram and let a be a real number. Then the set of all points X in the plane, such that

$$S(ABX) + S(CDX) = a$$

is a line (Fig. 6).

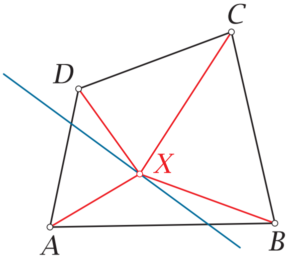

Fig. 6

Proof

Observe that point X satisfies $S(ABX) + S(CDX) = a$ if and only if it satisfies $S(BCX) + S(DAX) = a'$, where $a' = S(ABCD) - a$. Therefore, since $ABCD$ is not a parallelogram, we may without loss of generality assume that lines AB and CD are not parallel and intersect at P.

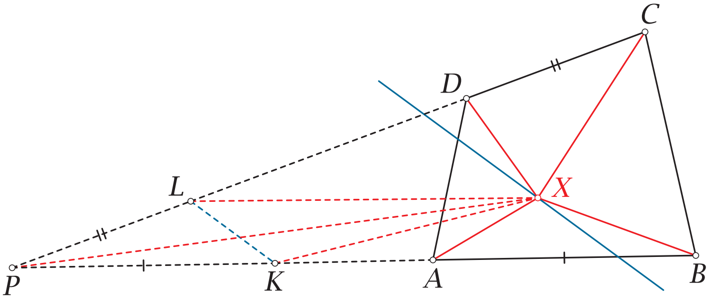

Fig. 7

Let K and L be the points determined by $\overrightarrow{PK} = \overrightarrow{AB}$ and $\overrightarrow{LP} = \overrightarrow{CD}$ (Fig. 7).

Then we obtain

$$S(ABX)+S(CDX)=S(PKX)+S(LPX)=S(KLP)-S(KLX).$$

Therefore, $S(ABX)+S(CDX)=a$ if and and only if $S(KLX)=b$, where $b=S(KLP)-a$. This means point X lies in a line parallel to line KL.

2. Proof of Newton's Theorem for a circle

We are ready to prove Newton's Theorem for a circle. The proof we are going to present is classical and well-known.

For simplicity we assume that $ABCD$ is convex and has an excircle (Fig. 8). The other circular configurations, including the most classical one shown at Figure 1, can be proved analogously.

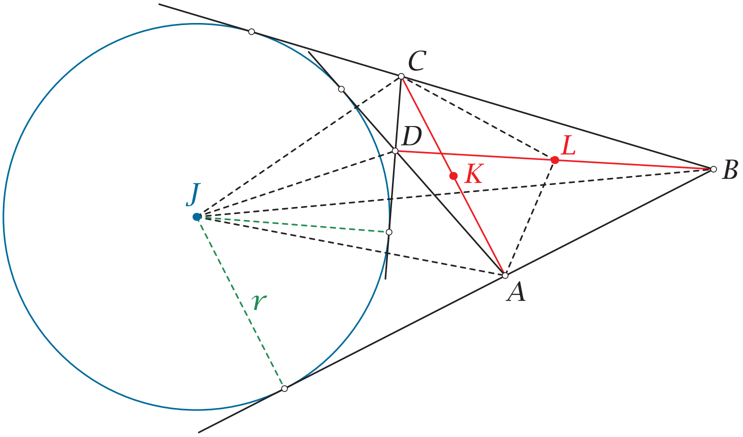

Fig. 8

Let J be the excenter of $ABCD$ and let K, L be the midpoints of diagonals AC, BD, respectively. By Theorem 1.2 it suffices to show that the sum $S(ABX)+S(CDX)$ attains the same value for $X=J,K$, and L.

Choose $X=L$ first. Then we have

$$S(ABL)+S(CDL)=\tfrac{1}{2}S(ABD)+\tfrac{1}{2}S(CDB)=\tfrac{1}{2}S(ABCD).$$

Similarly, we show that $S(ABK)+S(CDK)=\frac{1}{2}S(ABCD)$.

Let $X=J$. Since $ABCD$ has an excircle, we have $DA+AB=BC+CD$, or $AB-CD=BC-DA$. Therefore, we obtain

$$\begin{aligned}S(ABJ)+S(CDJ)&=\tfrac{1}{2}AB\cdot r-\tfrac{1}{2}CD\cdot r\\&=\tfrac{1}{2}BC\cdot r-\tfrac{1}{2}DA\cdot r=S(BCJ)+S(DAJ).\end{aligned}$$

But $S(ABJ)+S(CDJ)+S(BCJ)+S(DAJ)=S(ABCD)$, which implies that

$$S(ABJ)+S(CDJ)=\tfrac{1}{2}S(ABCD).$$

This completes the proof.

3. Proof of Newton's Theorem for a hyperbola (and an ellipse)

The signed area is also used in my proof of Newton's Theorem for a hyperbola and an ellipse. We start with another application of Theorem 1.2.

Theorem 3.1.

Let A, B, C, D be points such that lines DA and BC are not parallel and intersect at P (Fig. 9). Let Q be any point. Then the midpoints of segments AC, BD, and PQ are collinear if and only if $S(ABQ) = S(DCQ)$.

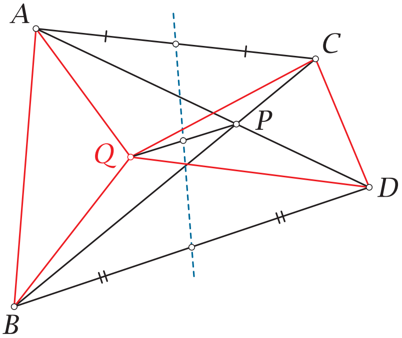

Fig. 9

Proof

Denote by K, L, and O the midpoints of segments AC, BD, and PQ, respectively (Fig. 10). Moreover, let X, Y be symmetric points to P with respect to points K, L, respectively. Then points K, L, O are collinear if and only if point Q lies on line XY.

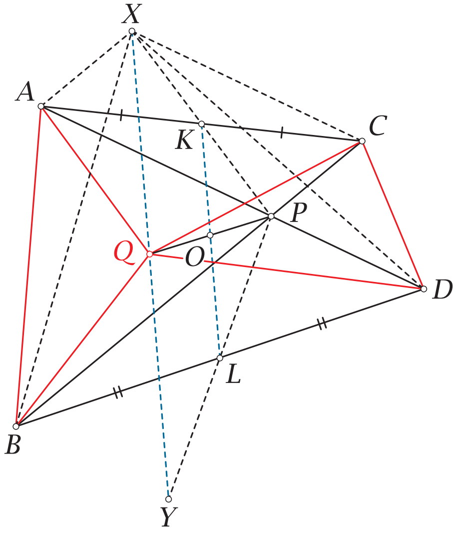

Fig. 10

Since $AXCP$ is a parallelogram, we get $S(ABX) = S(ACX) = S(DCX)$, or

$S(ABX)+S(CDX)=0$. Similarly, we obtain $S(ABY)+S(CDY)=0$. Therefore, applying Theorem 1.2 for $a=0$, we infer that point Q lies on line XY if and only if $S(ABQ)+S(CDQ)=0$, i.e. $S(ABQ)=S(DCQ)$. This completes the proof.

Theorem 3.2.

Two tangent lines a, b to a hyperbola (or to an ellipse) with center O meet at point P. Let Q be a symmetric point to P with respect to O (Fig. 11). A variable tangent line to the hyperbola (or to the ellipse) intersects lines a and b at X and Y, respectively. Then the signed area of triangle QXY is constant.

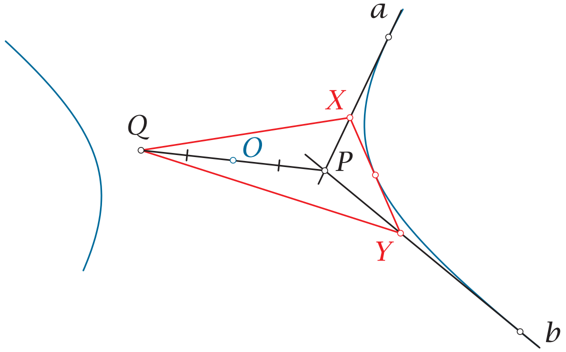

Fig. 11

Proof

Let a' and b' be symmetric lines to a and b, respectively, with respect to O (Fig. 12). Then a' and b' are tangent to the hyperbola (or to the ellipse). Let C be the touch point lying on a'. Set $K=a'\cap b$ and $L=a\cap b'$.

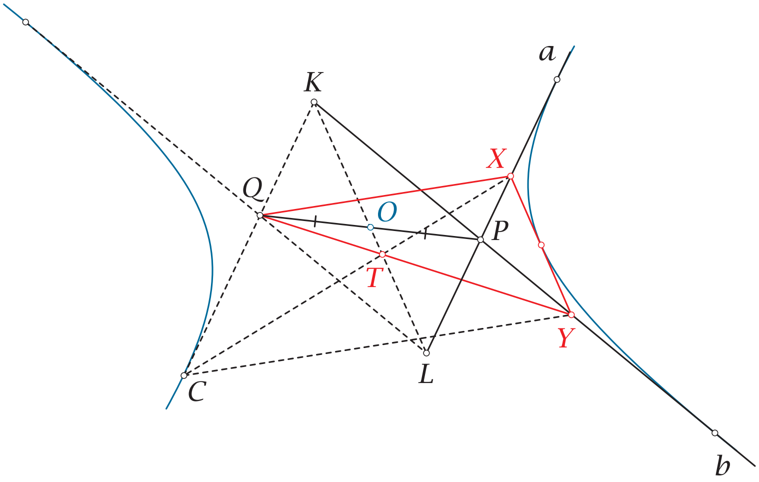

Fig. 12

Brianchon's Theorem applied for hexagon $KCQLXY$ implies that lines

KL, CX, and QY are concurrent at some point T. Moreover, since $XL \parallel CK$ and $LQ \parallel KY$, triangles XLQ and CKY are homothetic (or translated, if T is an infinity point). This yields $QX \parallel YC$. Therefore, $S(QXY) = S(QXC) = S(QPC)$, which is constant.

Theorem 3.3. (Newton's Theorem for a hyperbola and an ellipse)

Let A, B, C, D be points such that lines AB, BC, CD, DA are tangent to a hyperbola (or to an ellipse) with center O (Fig. 13). Then the line passing through the midpoints of AC and BD passes through O.

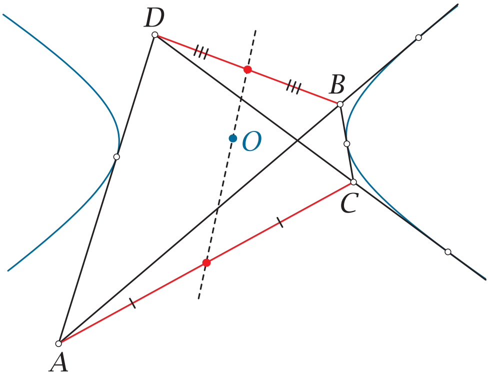

Fig. 13

Proof

Set $P = AB \cap CD$. Moreover, let Q be the point symmetric to P with respect to O (Fig. 14). By Theorem 3.2 we infer that $S(QBC) = S(QAD)$, which by Theorem 3.1 implies that O and the midpoints of AC and BD are collinear. This completes the proof.

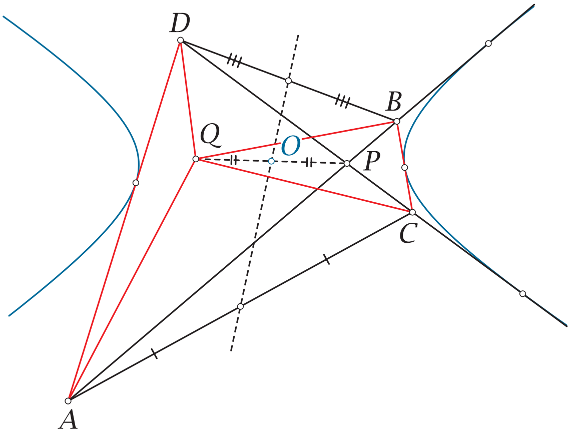

Fig. 14

4. Newton's Proof of Newton's Theorem

Newton's original proof of Theorem 3.3 (see [4], p. 146) is based on projective properties of a conic. Using today's knowledge and terminology it can be presented in a concise way as follows.

We assume a conic tangent to lines AB, BC, CD, DA is an ellipse or a hyperbola, with center O.

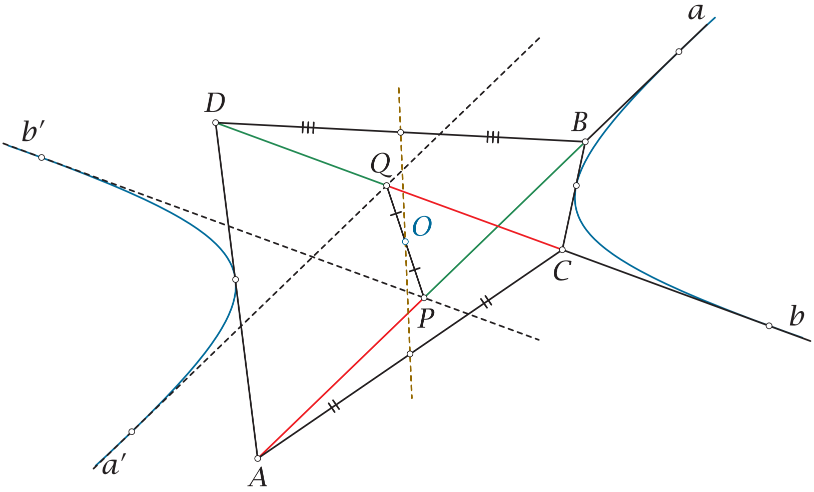

Fig. 15

Label the tangents AB and CD by a and b, respectively. Let a', b' be lines symmetric to a, b, respectively, with respect to O. Then a' and b' are also tangent to the conic. Set $P = a \cap b'$, $Q = a' \cap b$. Lines a, b, a', b' bound a parallelogram, so O is the midpoint of PQ.

A projection from line a to b through the conic preserves the double ratio, so

$$(ABP\infty) = (DC\infty Q) = (CDQ\infty),$$

implying

$$\frac{\overrightarrow{AP}}{\overrightarrow{PB}} = \frac{\overrightarrow{CQ}}{\overrightarrow{QD}}.$$

Thus degenerated triangles ABP and CDQ are similar, so the midpoints of AC, BD, and PQ are vertices of a triangle similar to ABP and CDQ. In particular, they are collinear, which completes the proof.

Remarks

Of course Newton didn't use the double ratio and the similar triangles property in his arguments. He derived the special cases of these properties directly and stated them as lemmas.

References

[1] T. Andreescu, L. Gonzalez, C. Pohoata, *Newton and Midpoints of Diagonals of Circumscriptible Quadrilaterals*, Mathematical Reflections **1** (2014).

[2] CodeParade, *Extraordinary Conics: The Most Difficult Math Problem I Ever Solved*, `https://youtu.be/X83vac2uTUs`, March 2020.

[3] R. Kaldybayev, *A generalization of Newton's quadrilateral theorem and an elementary proof of Minthorn's quadrilateral theorem*, `https://arxiv.org/abs/2211.09764`, November 2022.

[4] I. Newton, *The Mathematical Principles of Natural Philosophy,* `https://redlightrobber.com/red/links_pdf/Isaac-Newton-Principia-English-1846.pdf`, English translation, New York 1846.

[5] W. Pompe, *Problem 12266*, Romantics of Geometry Facebook Group, `https://www.facebook.com/groups/parmenides52/posts/6099435526836769`, April 2023.