in the classroom

Generalization and Specialization

The Strange Case of the Pythagorean Theorem

Shailesh Shirali

The word 'generalize' is extremely dear to mathematicians; they are always looking for ways to generalize something or the other! This should not come as a surprise, because generalization is utterly basic to mathematics. (It is just as basic to the notion of language, but we won't go into that here.) In this article we examine what 'generalization' means, along with its complementary action, 'specialization'. Using a few simple examples, we show that even in very elementary contexts there lurk strange paradoxes.

What does it mean to 'generalize'?

Say we have a proposition Q which contains some free variables or parameters. When those parameters are given particular values, or some constraints are placed on them, let the proposition take on a new form P. *Then,* P *is said to be a* ***specialization*** *of* Q*, and* Q *is said to be a* ***generalization*** *of* P*.*

(1) Consider the following pair of propositions:

(P) $10^2 - 1 = 9 \times 11$.

(Q) $n^2 - 1 = (n-1)(n+1)$ for all numbers n.

P is clearly a particular case of Q in which n has been given the value 10; so P is a specialization of Q, while Q is a generalization of P.

(2) Consider the following pair of propositions:

(P) The area of a circle with radius r is πr^2.

(Q) The area of an ellipse with semi-axes a and b is πab.

P is a particular case of Q in which the semi-axes have equal length r (this is so because the circle is a special case of an ellipse; see Figure 1); hence P is a specialization of Q, while Q is a generalization of P.

(3) Consider the following statements P and Q which refer to an arbitrary positive integer n and the remainder when n is divided by 9. The symbol $s(n)$ denotes the sum of the digits of n when it is written in base ten.

(P) n is divisible by 9 if and only if $s(n)$ is divisible by 9.

(Q) The remainder when n is divided by 9 is the same as the remainder when $s(n)$ is divided by 9.

Statement P is the familiar test for divisibility by 9. You may not be familiar with statement Q, but it is true and is proved in exactly the same way that P is proved. Here is an illustration of Q in action: if $n = 175$, then $s(n) = 13$. Please check that 175 and 13 leave the same remainder (namely, 4) under division by 9.

Here, P is a special case of Q, for it corresponds to the case when the remainder is 0. So P is a specialization of Q, while Q is a generalization of P. (Loosely speaking, "Q contains 9 times as much information as P".)

(4) Consider the following pair of statements which refer to a circle ω with centre O and distinct points A, B, C on ω.

(P) If BC is a diameter of ω, then $\angle BAC$ is a right angle; see Figure 2 (a).

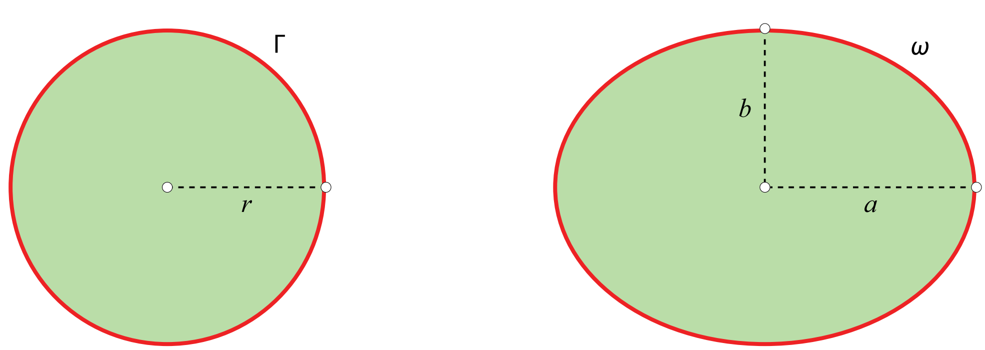
(Q) If A and O lie on the same side of BC, then $\angle BOC = 2\angle BAC$; see Figure 2 (b).

Here, P is a special case of Q in which $\angle BOC = 180°$.

(5) Consider the following statements which refer to a convex quadrilateral $ABCD$ with sides $AB = a, BC = b, CD = c, DA = d$, semi-perimeter $s = (a + b + c + d)/2$ and area Δ.

(P) If quadrilateral $ABCD$ is cyclic, its area Δ is given by

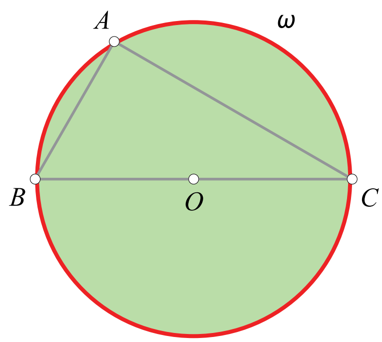
$$\Delta^2 = (s-a)(s-b)(s-c)(s-d).$$



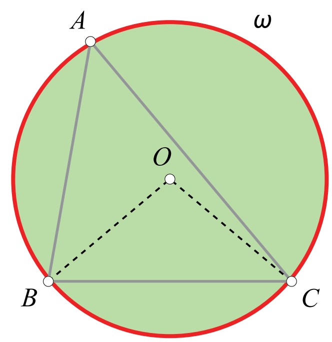
(a): Circle with radius r

(b): Ellipse with semi-axes a, b

Figure 1. A circle may be regarded as a particular case of an ellipse



(a): The case when $\measuredangle A = 90°$



(b): The general case

Figure 2. Two circle theorems: (a) is a particular case of (b)

(Q) If $ABCD$ is a general quadrilateral, its area Δ is given by

$$\Delta^2 = (s-a)(s-b)(s-c)(s-d) - abcd\cos^2\tfrac{1}{2}(A+C).$$

Here, P is a special case of Q; for, if the quadrilateral is cyclic, then $A + C = 180°$, so the cosine term in Q vanishes (because $\cos 90° = 0$) and we get formula P.

Some of you may know that P was first found by Brahmagupta.

Note that Q implies much more than P. Here is a lovely corollary of Q which derives from the fact that the cosine term comes in squared form and so is never negative: If the sides a, b, c, d of a quadrilateral are fixed, then its area is largest when the quadrilateral is cyclic.

Three Apparent Generalizations Of The Pythagorean Theorem

The title of this section is "Three Apparent Generalizations Of The PT" (we use the short-form 'PT' for 'Pythagorean Theorem'). We should add the following as a subtitle: "Or Are They 'Mere' Corollaries?" We present three propositions which can be thought of as generalizations of the PT, but which have an odd, paradoxical feature associated with them. The first one is the **cosine rule**.

Theorem 1 (Cosine rule). In $\triangle ABC$, we have the following relationship:

$$a^2 = b^2 + c^2 - 2bc\cos A.$$

The cosine rule clearly contains the PT as a special case: if $\angle A = 90°$ then $\cos A = 0$, therefore $a^2 = b^2 + c^2$.

It also implies the *converse* of the PT; for, the cosine of an angle of a triangle is 0 precisely when the angle is a right angle. Hence if it happens that $a^2 = b^2 + c^2$, then it must be that $\cos A = 0$ (since bc is not zero), and therefore that $\angle A = 90°$.

Indeed, the cosine rule yields still more: it also implies the *inequality form of the PT*, drawing on the fact that the cosine of an angle is positive if the angle is acute, and negative if the angle is obtuse. Therefore, if $\angle A < 90°$ then $a^2 < b^2 + c^2$, and if $\angle A > 90°$ then $a^2 > b^2 + c^2$. So there is a lot of information contained in that simple rule.

So surely the cosine rule can be regarded as a genuine generalization of the PT?

Next we study the **theorem of Apollonius**. This is named after the third century AD Greek geometer Apollonius (often described as "the greatest geometer of antiquity", and known for his work on the conic sections), and it states the following:

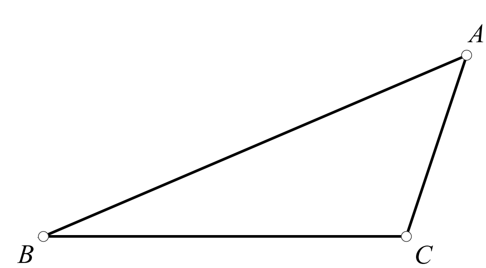
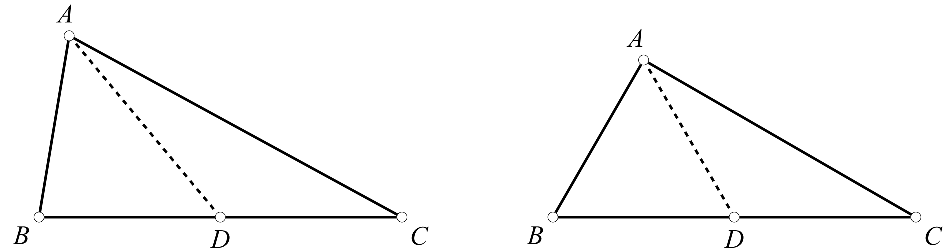


Figure 3. The cosine rule



(a): General case

(b): The case when $\measuredangle A = 90^\circ$

Figure 4. The theorem of Apollonius

Theorem 2 (Apollonius). In $\triangle ABC$, let D be the midpoint of BC. Then:

$$AB^2 + AC^2 = 2\left(AD^2 + BD^2\right).$$

See Figure 4(a). In the special case when $\angle A = 90^\circ$ (see Figure 4 (b)), D is the centre of the circle through A, B, C, so $AD = BD$. The statement now reduces to: $AB^2 + AC^2 = 4BD^2$. Since $2BD = BC$, this may be written as $AB^2 + AC^2 = BC^2$. So this theorem too yields the PT as a special case.

With some ingenuity one can derive the inequality form of the PT from the theorem of Apollonius. (We urge you to try doing so.)

A symmetrical and more pleasing form of this theorem is the following: *The sum of the squares of the four sides of a parallelogram is equal to the sum of the squares of the diagonals.* This is equivalent to the theorem of Apollonius because of the easily-proved result that the diagonals of a parallelogram bisect one another.

Apollonius's theorem may itself be expressed in a stronger form, and we have the following theorem first found by a Scottish mathematician of the eighteenth century, Mathew Stewart (see Figure 5):

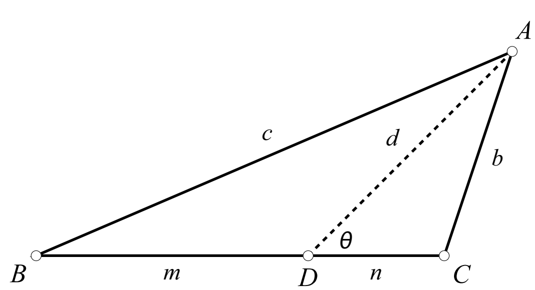


Figure 5. Stewart's theorem

Theorem 3 (Stewart). In $\triangle ABC$, let D be a point on side BC. Let a, b, c be the lengths of the sides of the triangle, and let d, m, n be the lengths of AD, BD, CD. Then:

$$b^2 m + c^2 n = a\left(d^2 + mn\right).$$

If D is the midpoint of BC then we have $m = n = a/2$, hence the statement reduces to $\left(b^2 + c^2\right) a/2 = a\left(d^2 + a^2/4\right)$, giving $b^2 + c^2 = 2d^2 + a^2/2$. This is equivalent to the theorem of Apollonius. (Do you see why?) Hence the theorem of Apollonius may be regarded as a specialization of Stewart's theorem. This implies that the theorem of Pythagoras may be considered a specialization of Stewart's theorem.

Here is a way of proving Stewart's theorem. Let $\angle ADC = \theta$; then $\angle ADB = 180^\circ - \theta$. Using the cosine rule together with the identity $\cos(180^\circ - \theta) = -\cos\theta$, we have:

$$b^2 = d^2 + n^2 - 2dn\cos\theta,$$

$$c^2 = d^2 + m^2 + 2dm\cos\theta.$$

If we multiply the first relation by m and the second one by n and then add them, the terms containing $\cos\theta$ are eliminated and we then get: $mb^2 + nc^2 = md^2 + mn^2 + nd^2 + nm^2$. Since $m + n = a$, this may be written in a more convenient way:

$$mb^2 + nc^2 = ad^2 + amn = a\left(d^2 + mn\right).$$

Those of you who are familiar with the math Olympiads will know that Stewart's theorem is part of the staple diet for all mathletes.

And now ... a paradox!

We did not present the proofs of the cosine rule and the theorem of Apollonius, as most 11th standard mathematics textbooks give these proofs. But if we study these proofs, and the one given above for Stewart's theorem, we find a paradoxical situation. *Namely, these proofs are based on the Pythagorean theorem!* A good exercise for you would be to study these proofs and find the exact point(s) where the PT has been used. It may well come in a disguised form! For example, you may opt for a vector proof of the cosine rule as follows: Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ denote the vectors $\overrightarrow{AB}$, $\overrightarrow{AC}$, $\overrightarrow{BC}$ respectively (see Figure 3). Then $\mathbf{w} = \mathbf{u} - \mathbf{v}$, and by squaring we get:

$$\mathbf{w} \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} - 2\mathbf{u} \cdot \mathbf{v}.$$

But $\mathbf{w} \cdot \mathbf{w} = a^2$, $\mathbf{u} \cdot \mathbf{u} = b^2$, $\mathbf{v} \cdot \mathbf{v} = c^2$ and $\mathbf{u} \cdot \mathbf{v} = bc \cos A$. Hence

$$a^2 = b^2 + c^2 - 2bc \cos A.$$

Similarly, for the theorem of Apollonius (Figure 4), let $\mathbf{u}$ and $\mathbf{v}$ denote the vectors $\overrightarrow{DB}$ and $\overrightarrow{DA}$ respectively; then $\overrightarrow{DC} = -\mathbf{u}$, so $\overrightarrow{AB} = \mathbf{u} - \mathbf{v}$, $\overrightarrow{AC} = -\mathbf{u} - \mathbf{v}$. Now we obtain expressions for AB^2, AC^2, AD^2, BD^2 and confirm that the stated claim is true.

But where has the PT been used in these two derivations? We leave this puzzle for you to crack.

So we have here a situation where a theorem appears to lead to its own generalization; how apparently paradoxical! What are we to make of it? In such a case, should we not regard the generalization as "only a corollary" and not a generalization at all?

Or should we say that the cosine rule can be regarded as a generalization of the PT only if we can find a proof for the rule that is not based on the PT? We leave you to ponder the matter.

Exercises

(1) Show how the theorem of Apollonius implies the inequality form of the PT.

(2) If a pair of propositions P and Q have the property that P $\implies$ Q and also Q $\implies$ P, what word is appropriate to describe the relationship between P and Q?

(3) Find some nice specializations of the cosine rule. (See what you get if you put $\theta = 60°$ or $120°$.)

(4) Consider the following pair of propositions:

(P) $(1 + x)^2 = 1 + 2x + x^2$.

(Q) $(a + b)^2 = a^2 + 2ab + b^2$.

We certainly get the impression that P is a specialization of Q; for, by putting $a = 1$, $b = x$ in Q, we get P. So it would seem that Q is a generalization of P. But is this so? Consider what happens in P if we put $x = b/a$ in P. We get:

$$\left(1 + \frac{b}{a}\right)^2 = 1 + \frac{2b}{a} + \frac{b^2}{a^2},$$
$$\therefore (a + b)^2 = a^2 + 2ab + b^2.$$

So we have P $\implies$ Q as well as Q $\implies$ P. In the light of this, would you still say that Q is a generalization of P?

See [1] for more such examples. Chapter 2 in [2] has an illuminating discussion on the process of generalization.

References

[1] Bogomolny, A. Generalizations in Mathematics, from *Interactive Mathematics Miscellany and Puzzles*, http://www.cut-the-knot.org/Generalization/epairs.shtml, Accessed 10 March 2014

[2] Pólya, G. *Mathematics and Plausible Reasoning*, Vol 1, Induction and Analogy in Mathematics. Princeton University Press, 1954.

SHAILESH SHIRALI is Director and Principal of Sahyadri School (KFI), Pune, and Head of the Community Mathematics Centre in Rishi Valley School (AP). He has been closely involved with the Math Olympiad movement in India. He is the author of many mathematics books for high school students, and serves as an editor for *At Right Angles*. He may be contacted at shailesh.shirali@gmail.com.