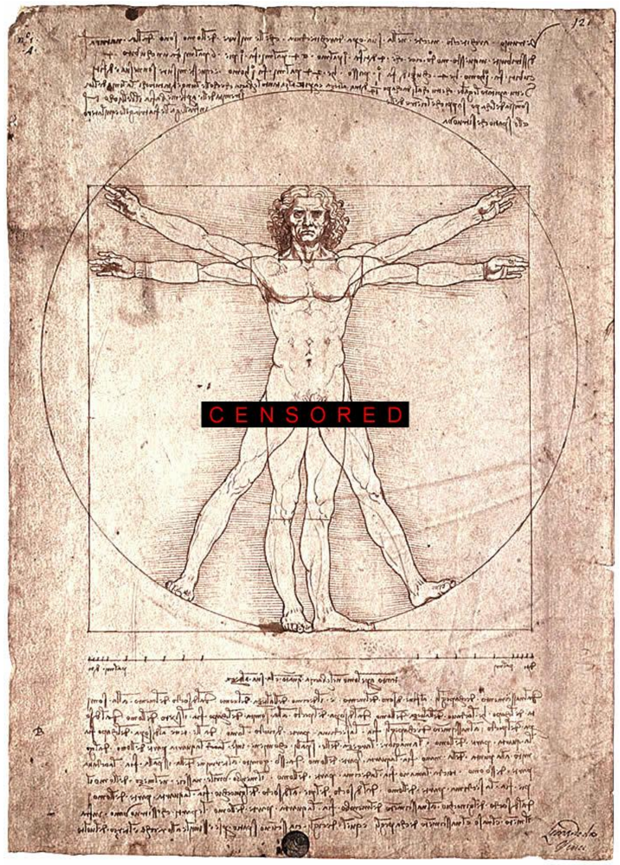
What is Symmetry?

You probably have an intuitive idea of what symmetry means and can recognise it in various guises. For example, you can hopefully see that the letters **N**, **O** and **M** are symmetric, while the letter **R** is not. But probably, you can't define in a mathematically precise way what it means to be symmetric — this is something we're going to address.



Symmetry occurs very often in nature, a particular example being in the human body, as demonstrated by Leonardo da Vinci's drawing of the Vitruvian Man on the right.[1] The symmetry in the picture arises since it looks essentially the same when we flip it over. A particularly important observation about the drawing is that the distance from the Vitruvian Man's left index finger to his left elbow is the same as the distance from his right index finger to his right elbow. Similarly, the distance from the Vitruvian Man's left knee to his left eye is the same as the distance from his right knee to his right eye. We could go on and on writing down statements like this, but the point I'm trying to get at is that our intuitive notion of symmetry — at least in geometry — is somehow tied up with the notion of distance.

What is an Isometry?

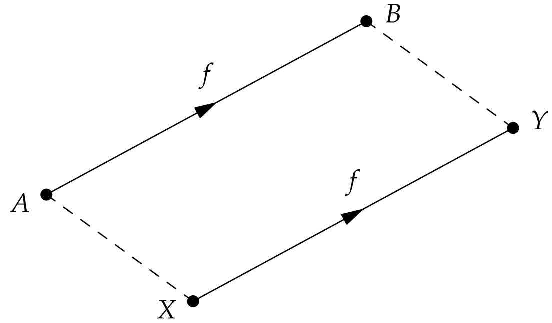
The flip in the previous discussion was a particular function which took points in the picture to other points in the picture. In particular, it did this in such a way that two points which were a certain distance apart would get mapped to two points which were the same distance apart. This motivates us to consider functions f which map points in the plane to points in the plane such that the distance from $f(P)$ to $f(Q)$ is the same as the distance from P to Q for any choice of points P and Q. Any function which satisfies this property is called an *isometry*. This comes from the ancient Greek words "isos", meaning equal, and "metron", meaning measure.

Example. The best way to get a feeling for what an isometry looks like is to consider some examples.

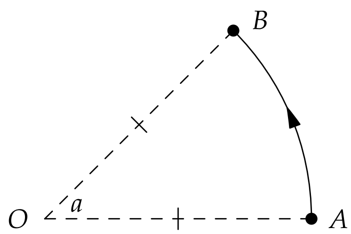
- *Identity* : The identity is the function which simply takes a point P in the plane to the same point P. In other words, it does nothing, so hopefully you can see that it's an example of an isometry — in fact, the simplest example of an isometry. We'll sometimes denote the identity map by I.
- *Translation* : A translation is a function which takes every point in the plane and slides it in a certain direction by a certain distance. In other words, if f is a translation such that $f(A) = B$ and $f(X) = Y$, then the quadrilateral $ABYX$ will always be a parallelogram. This means that if you want to specify a

[1]Leonardo da Vinci was very interested in symmetry and later on we're going to see a theorem which bears his name.

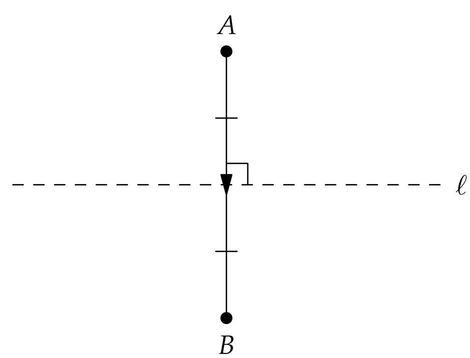
translation, then you only have to specify a point A and the point $f(A) = B$. We'll sometimes denote the translation which takes the point A to the point B by T_{AB}.



- *Rotation* : A rotation is a function which takes the whole plane and turns it about a certain point O through a certain angle a. In other words, if f is a rotation such that $f(A) = B$, then the centre of rotation O and the angle of rotation a satisfy $OA = OB$ and $\angle AOB = a$. This means that if you want to specify a rotation, then you only have to specify the centre of rotation and the angle of rotation. We'll sometimes denote the rotation about the point O through an angle a by $R_{O,a}$.



- *Reflection* : A reflection is a function which takes the whole plane and flips it over a certain line ℓ which we also refer to as a mirror. It's important to observe that if a reflection maps the point A to the point B, then ℓ will be the perpendicular bisector of the line segment AB. This means that if you want to specify a rotation, then you only have to specify the mirror. We'll sometimes denote the reflection through the line ℓ by M_ℓ.

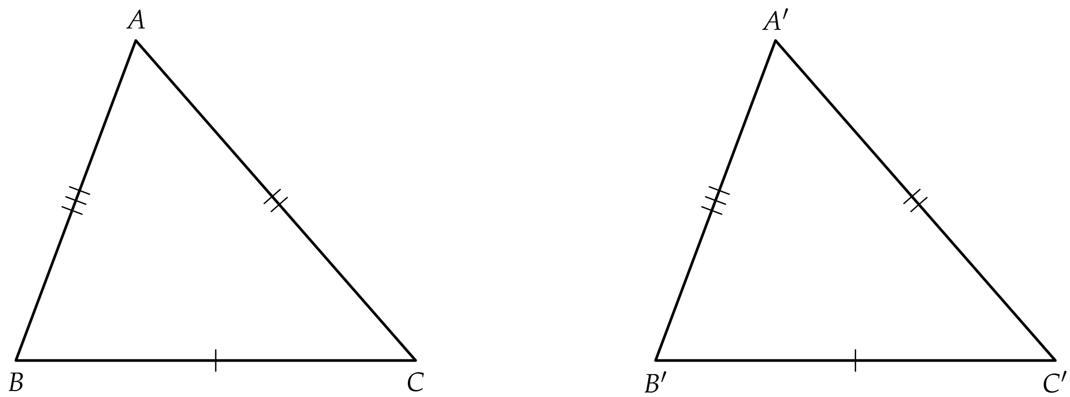


Suppose that P and Q are two points in the plane which are distance d apart from each other. If we apply an isometry to both points, then the result will be two points which are still distance d apart from each other. If we apply another isometry to the resulting points, then the new result will be two points which are still distance d apart from each other. The application of two or more isometries in a row is called *composition of isometries*. From what we've just said, you can see that the composition of two or more isometries will always be an isometry. This means that we can compose the examples we've listed above in an attempt to discover new isometries.

Interesting Facts about Isometries

One of our guiding questions will be to determine what sorts of isometries there are. For example, is it true that the examples we discussed earlier — identity, translation, rotation and reflection — account for every single possible isometry? Before we answer this question, we need to prove some basic facts.

Proposition. *If f is an isometry such that $f(A) = A'$, $f(B) = B'$, $f(C) = C'$, then triangles ABC and $A'B'C'$ are congruent.*



Proof. By the very definition of an isometry, we have the equal lengths $AB = A'B'$, $BC = B'C'$ and $CA = CA'$. So it follows by SSS that the triangle ABC and the triangle $A'B'C'$ are congruent. □

What if I give you two triangles ABC and $A'B'C'$ and ask you whether you can find an isometry such that $f(A) = A'$, $f(B) = B'$ and $f(C) = C'$? The previous proposition says that you couldn't possibly be able to do this unless the triangles ABC and $A'B'C'$ are congruent. So if we suppose that they actually are congruent, then can you definitely find an isometry satisfying the conditions? If so, then how many can you find exactly? The next proposition states that you can always find one and only one.

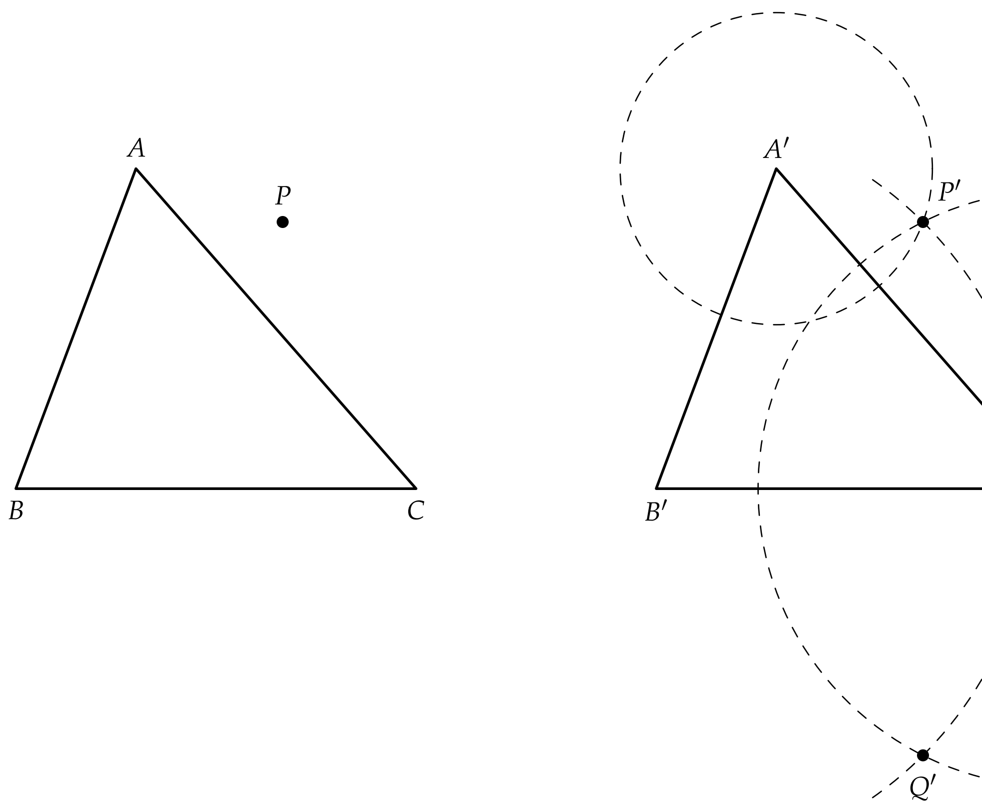
Theorem. *If triangles ABC and $A'B'C'$ are congruent, then there is a unique isometry such that $f(A) = A'$, $f(B) = B'$ and $f(C) = C'$.*

Proof. To show that there is at least one such isometry, we can construct it explicitly. First, we translate triangle ABC until A lines up with A'. Then we rotate triangle ABC around the point A until B lines up with B'. Now if C and C' don't already line up, then we reflect through the line AB so that they do. Since the composition of these two or three isometries is also an isometry, we have constructed an isometry satisfying the conditions of the problem. Note that we need triangles ABC and $A'B'C'$ to be congruent so that everything lines up exactly as we have described.

Now we need to show that there is at most one such isometry. Take any point P in the plane — we will show that there is only one possible point P' for $f(P)$. This is because the definition of an isometry forces the point P' to satisfy the conditions $A'P' = AP$, $B'P' = BP$, and $C'P' = CP$.

Another way to say this is that P' has to

- lie on the circle with centre A' and radius AP;
- lie on the circle with centre B' and radius BP; and
- lie on the circle with centre C' and radius CP.



But hopefully you can see that the latter two circles can only meet at two points which we've labelled P' and Q', while the first circle can only pass through one of these two. And this single point where the three circles meet is the only possible location for the point $f(P)$. So we have shown that for every point P in the plane, there is only one possible location for the point $f(P)$. This means that there is at most one isometry f such that $f(A) = A'$, $f(B) = B'$ and $f(C) = C'$.

Since we've proven that there is at least one isometry satisfying the given conditions as well as the fact that there is at most one isometry satisfying the given conditions, we can put these two statements together to deduce that there's a unique isometry satisfying the given conditions. □

One very important consequence of this theorem is the fact that two isometries f and g must be the same if they satisfy $f(A) = g(A)$, $f(B) = g(B)$ and $f(C) = g(C)$ for some triangle ABC.

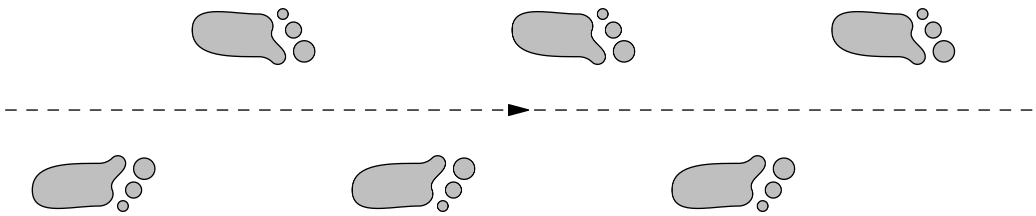
Composition of Isometries

We've already mentioned that you can take isometry f followed by an isometry g and that the result is another isometry called the composition of f and g. It's going to be very useful to have some notation for this idea. So let's denote the result of "doing f followed by doing g" as $g \circ f$. Yes, that's right — even though we do f first and g second, we write the result with g to the left of f.[2]

$$P \xrightarrow{f} f(P) \xrightarrow{g} g(f(P)) = g \circ f(P)$$

[2]The reason for this is because in function notation, we would normally write the result of doing f followed by doing g to a point P as $g(f(P))$ — note that even though we do f first and g second, we write the result with g to the left of f.

As our first application of composition, I'm going to introduce a new isometry which didn't appear in our original list of examples. It's the isometry which you get when you do a reflection in a line ℓ followed by a translation parallel to ℓ. Using our earlier notation, we could write such a function as $T_{AB} \circ M_{\ell}$, where AB is a line segment parallel to ℓ. An isometry which takes this form is called a *glide reflection* and we'll sometimes denote it by G_{AB}. The process of repeatedly applying a glide reflection is something you've no doubt all done before when walking along the beach. Each successive footprint that you leave in the sand is a glide reflection applied to the previous footprint.



Since you can compose different isometries, a natural question to ask is what the result is. For example, the fact below follows immediately from the definition of a translation and doesn't really need a proof.

Proposition. *The composition of a translation with another translation is always a translation.*[3] *In fact, we have the rather obvious formula* $T_{BC} \circ T_{AB} = T_{AC}$.

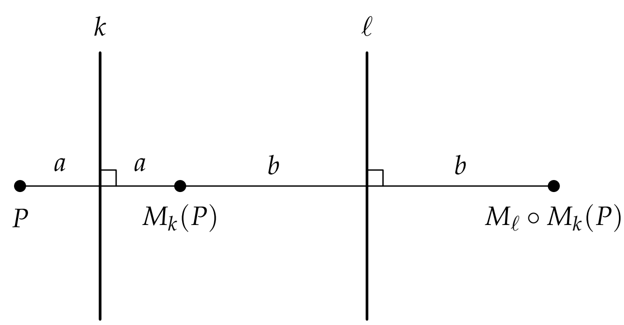
To write down a similar fact concerning the composition of two reflections is a more difficult matter — but we can still do it.

Proposition. *Let* M_k *denote a reflection in the line* k *and* M_ℓ *denote a reflection in the line* ℓ*. Then the composition* $M_\ell \circ M_k$ *of a reflection in* k *followed by a reflection in* ℓ *is*

- *the identity if* k *and* ℓ *are the same line;*
- *a translation if* k *and* ℓ *are parallel lines — the direction of translation is perpendicular to* k *and* ℓ *while the distance of translation is twice the distance between* k *and* ℓ*; or*
- *a rotation if the two lines meet — the centre of rotation is the intersection of* k *and* ℓ *while the angle of rotation is twice the angle from* k *to* ℓ*.*

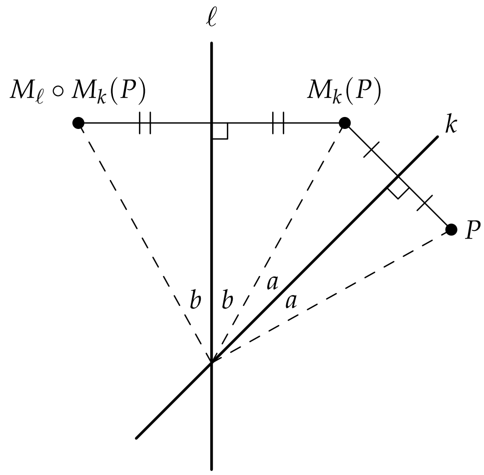
Proof.

- This one's obvious because if you reflect twice through the same line, you end up where you started.
- In the following diagram, you can see that the result of applying $M_\ell \circ M_k$ to P is to move it by a distance of $2a + 2b$, where $a + b$ is the distance between k and ℓ. Furthermore, the direction of translation is perpendicular to both k and ℓ.



[3] You might think that we haven't accounted for the fact that the composition of two translations could possibly be the identity. However, you can and should consider the identity as a translation which moves every point by zero distance.

- In the following diagram, you can see that the result of applying $M_\ell \circ M_k$ to P is to rotate by an angle $2a + 2b$, where $a + b$ is the angle from k to ℓ. Furthermore, the centre of rotation is the intersection of k and ℓ.



Although we've demonstrated these facts for only one particular point P, the argument would have worked for several different choices of P. By the theorem we proved earlier, we only need to demonstrate these facts for three points which form a triangle to know that they hold for all points. □

We'll also state a couple more results concerning composition of isometries. We won't provide the proofs, since they aren't particularly interesting and they use ideas very similar to those in the previous proof. However, you should definitely draw a few diagrams to convince yourself that they are true.

Proposition. *Let R_1 be a rotation by angle a_1 and R_2 be a rotation by angle a_2. Then the composition $R_2 \circ R_1$ is*

- *a rotation by angle $a_1 + a_2$ if $a_1 + a_2 \neq 360°$; or*
- *a translation if $a_1 + a_2 = 360°$.*

Proposition. *A reflection followed by a translation is a reflection or a glide reflection. A reflection followed by a rotation is a reflection or a glide reflection.*

Classification of Isometries

Since the composition of two isometries is again an isometry, you can try to build every possible isometry out of some simple building blocks.[4] The next result tells us that we can take the reflections as our building blocks.

Proposition. *Every isometry is the composition of at most three reflections.*

Proof. Consider any isometry f, any triangle ABC, and let $f(A) = A'$, $f(B) = B'$ and $f(C) = C'$. All we need to prove is that it takes at most three reflections to send A to A', B to B' and C to C'. We simply reflect in the perpendicular bisector of AA' so that A ends up coinciding with A'. Now we simply reflect in the perpendicular bisector of BB' so that B ends up coinciding with B'. The congruence of triangles ABC and $A'B'C'$ ensures that A still coincides with A'. Now either the two triangles coincide and the job took two reflections, or we need one more reflection through $A'B'$ to finish off the job, so that C coincides with C'. □

[4] This reeks of the reductionist approach that we used to begin our journey into Euclidean geometry.

Theorem. *Every isometry is a translation, a reflection, a rotation or a glide reflection.*

Proof. We just proved that every isometry is the product of at most three reflections, so the proof can be divided into the following cases.

- The only isometry which is the product of zero reflections is, of course, the identity isometry. As usual, we can think of the identity isometry as a translation which moves every point by zero distance.
- The only isometries which are the product of one reflection are, of course, reflections themselves.
- The proposition which we proved above tells us that the product of two reflections is the identity, a translation or a rotation.
- Now let's consider the product of three reflections as a reflection followed by the product of two reflections. Thus, the product of three reflections can be the product of a reflection and the identity, the product of a reflection and a translation, or the product of a reflection and a rotation. In the first case we obtain a reflection, in the second case we obtain a reflection or a glide reflection, and in the third case we obtain a reflection or a glide reflection.

So what you can hopefully see is that products of up to three reflections always result in a translation, a reflection, a rotation or a glide reflection. Furthermore, since every isometry is the composition of at most three reflections, this accounts for all possible isometries. □

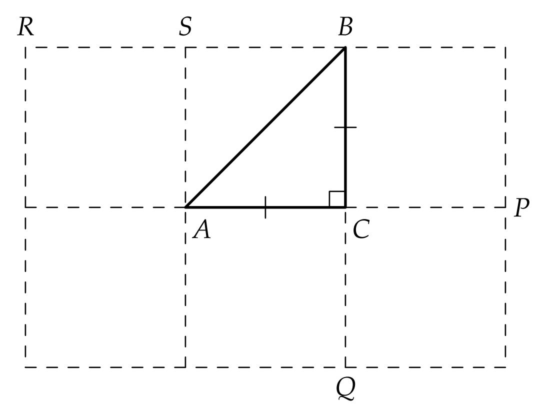
Problems

Problem. *Let ABC be a triangle with the vertices labelled clockwise such that $AC = BC$ and $\angle ACB = 90°$. Let M_{AB} be the reflection in the line AB, M_{AC} be the reflection in the line AC, and R be the rotation by $90°$ counterclockwise around B. Identify the composition $R \circ M_{AB} \circ M_{AC}$.*

Proof. The idea is to use the theorem we proved earlier which states that

> *if triangles ABC and $A'B'C'$ are congruent, then there is a unique isometry such that $f(A) = A'$, $f(B) = B'$ and $f(C) = C'$.*

What this means is that we can solve problems like this one using the following simple strategy. Find three points which form a triangle and see where the composition of isometries takes them. Next, it's time to guess what the isometry is. If your guess is correct for the three vertices of the triangle, then it must be correct. And this is because the theorem above guarantees that if you know what an isometry does to three corners of a triangle, then you know what the isometry does to every point in the plane.



The diagram above shows triangle ABC drawn on a grid of squares. Since we want to choose three points which form a triangle, we may as well choose the points A, B and C.

- It's easy to check that $M_{AC}(A) = A$, $M_{AB}(A) = A$ and $R(A) = P$.
  In other words, $R \circ M_{AB} \circ M_{AC}(A) = P$.

- It's easy to check that $M_{AC}(B) = Q$, $M_{AB}(Q) = R$ and $R(R) = Q$.
  In other words, $R \circ M_{AB} \circ M_{AC}(B) = Q$.
- It's easy to check that $M_{AC}(C) = C$, $M_{AB}(C) = S$ and $R(S) = C$.
  In other words, $R \circ M_{AB} \circ M_{AC}(C) = C$.

So can you think of an isometry which takes A to P, B to Q and C to C? If you think hard enough, you should realise that it's just a rotation by 180° around C. So we've managed to deduce that the composition $R \circ M_{AB} \circ M_{AC}$ is a rotation by 180° around C. □

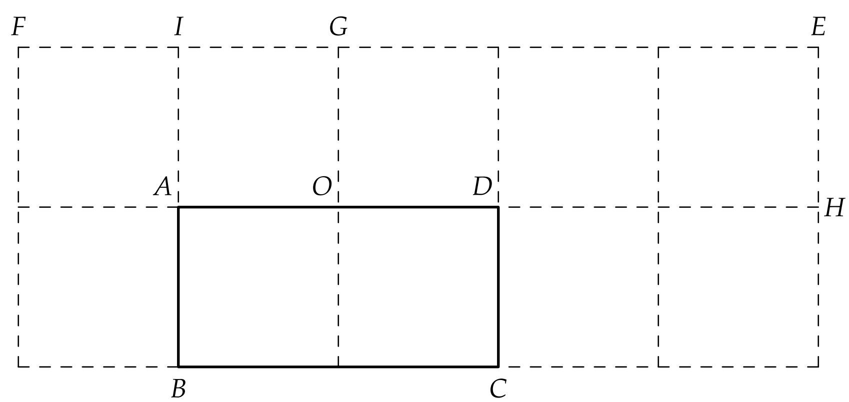
Problem. *Let $ABCD$ be a rectangle with the vertices labelled counterclockwise such that $BC = 2AB$. Let*

- *M_{AB} be the reflection in the line AB;*
- *R_B be the counterclockwise rotation by 90° about B;*
- *T_{DB} be the translation which takes D to B; and*
- *G_{CD} be the glide reflection in the line CD which takes C to D.*

Identify the composition $M_{AB} \circ R_B \circ T_{DB} \circ G_{CD}$.

Proof. The diagram below shows rectangle $ABCD$ drawn on a grid of squares. Since we want to choose three points which form a triangle, we may as well choose the points A, B and C.

- It's easy to check that $G_{CD}(A) = E$, $T_{DB}(E) = D$, $R_B(D) = F$ and $M_{AB}(F) = G$.
  In other words, $M_{AB} \circ R_B \circ T_{DB} \circ G_{CD}(A) = G$.
- It's easy to check that $G_{CD}(B) = H$, $T_{DB}(H) = C$, $R_B(C) = I$, and $M_{AB}(I) = I$.
  In other words, $M_{AB} \circ R_B \circ T_{DB} \circ G_{CD}(B) = I$.
- It's easy to check that $G_{CD}(C) = D$, $T_{DB}(D) = B$, $R_B(B) = B$, and $M_{AB}(B) = B$.
  In other words, $M_{AB} \circ R_B \circ T_{DB} \circ G_{CD}(C) = B$.



So can you think of an isometry which takes A to G, B to I and C to B? If you think hard enough, you should realise that it's a rotation, although you might not be sure of where the centre lies. However, we can use the fact that if a rotation takes X to Y, then the centre of rotation must lie on the perpendicular bisector of XY. In particular, the centre of the rotation that we're interested in must lie on the perpendicular bisector of AG as well as the perpendicular bisector of BI. And there's only one point which does that — namely, the point O labelled in the diagram above. It's now easy to deduce that the composition must be a rotation about O by $\angle AOG = 90°$ in the clockwise direction. □

Archimedes

Archimedes of Syracuse was a Greek mathematician who is widely held to be the greatest mathematician of antiquity and one of the greatest of all time. Living from about 287 BC to 212 BC, his thoughts on mathematics were far ahead of his time. He used a technique called the method of exhaustion to calculate areas under parabolas and give a very accurate approximation of the number π. He also applied this technique to prove that a sphere which fits perfectly inside a cylinder has two-thirds of its surface area as well as two-thirds of its volume. He regarded this as the greatest of his mathematical achievements and asked for a diagram of a sphere inside a cylinder to be placed on his tombstone. Archimedes' method of exhaustion was a precursor to the modern day differential and integral calculus which was discovered nearly two thousand years later.

It's quite common for a mathematician to be multitalented and Archimedes was certainly no exception. He is also renowned as a physicist, engineer, inventor, and astronomer. Among his great discoveries and inventions are the foundations of hydrostatics, the explanation of the principle of the lever, machines to be used in siege warfare, and many many more. He is the one who is said to have cried "Eureka!" and ran through the streets of Syracuse naked upon discovering the principle of buoyancy while in the bathtub. Another story about Archimedes is that his servants needed to take him against his will to the baths. And while they bathed him and anointed him with oils, he would be drawing diagrams on his body with the oils, such was his enthusiasm for geometry.

The mathematical writings of Archimedes were not particularly well known throughout antiquity. However, the few copies which survived through to the Middle Ages became an influential source of ideas for scientists. Amazingly, previously unknown works of Archimedes were discovered in 1906. These writings, now known as the Archimedes Palimpsest, provide new insights into how he obtained mathematical results. We are quite lucky to have them because this copy of his writings had been made in the tenth century AD and the pages subsequently erased, folded in half, and reused for a Christian text. Fortunately, the erasure was incomplete and we can now read Archimedes' work after scientific and scholarly work over the past ten years involving digital image processing using ultraviolet, infrared and X-ray technology.

Although Archimedes had built many machines of war to keep the Romans out, they finally managed to capture his home town of Syracuse. Apparently, one Roman soldier happened to find him hard at work on a geometry problem. Archimedes was so transfixed that he never noticed the soldier nor even the fact that the city had been taken. When Archimedes refused to follow the soldier until he had finished solving the problem, the soldier decided to run his sword through him, despite orders to keep him alive. And that was the end of Archimedes.