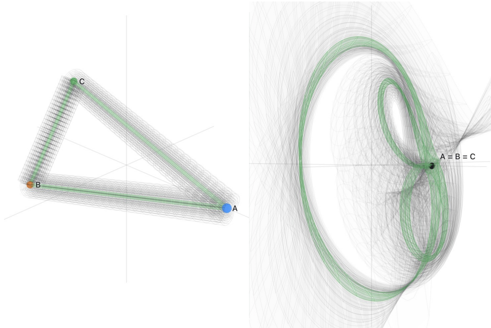

A Simplified View of the Jacobian Conjecture

The full conjecture is stated over abstract fields, but the counterexample is a concrete 3D function that we can explain and visualize using familiar geometric ideas and a little algebra.

Publication note
Originally published in Towards Data Science. Author's copy on Substack.

It’s likely that you’ve heard about a new mathematical result announced by Levent Alpöge that came out of work with Fable a few days ago concerning something called the Jacobian conjecture. Unfortunately, many discussions of the result simply repeat parts of the conjecture’s formal definition, with little in the way of a widely accessible explanation.

It’s an understandable problem because mathematicians tend to use very formal language. That language has the advantage of being precise, but the disadvantage of using terms that don’t mean anything to most people. Unfortunately, explaining all of the surrounding terminology from scratch can take about as long as earning a PhD in math.

However, in this case the mathematical result has a relatively simple explanation that can be understood without a PhD. Some precision is inevitably lost by expressing it simply, but the core idea becomes much easier to understand. The explanation still requires some math, but at the level of a typical STEM undergraduate rather than a math PhD.

The Jacobian Conjecture asks whether a particular type of function must always be invertible. The disproof of the conjecture is simply a counterexample: a function that meets all the conjecture’s requirements but provably cannot be inverted.

As a reminder, a function can be invertible only if its output uniquely determines its input. If a function produces the same output for two different inputs, then it cannot be reversed. There are also other requirements for a function to be invertible, but a function without this necessary “one-to-one” property cannot be inverted.

So, if we can find a function that satisfies the Conjecture’s requirements, but the function is nevertheless not one-to-one, then we have a contradiction: The function would match what the conjecture describes, but it can’t be invertible, so the conjecture would be false.

The formal conjecture applies to polynomial functions over any “field of characteristic zero”. However, the counterexample only relies on ordinary real numbers in a three-dimensional space. That means that we can explain this counterexample using concepts from the ordinary three-dimensional space we all live in and are familiar with.

Imagine a big block of infinitely stretchy rubber. Technically, this would be an infinitely large block of rubber that fills the whole universe. Obviously, that's not possible, but we can still *imagine* a giant block of rubber that never ends.

Now imagine reshaping that rubber by stretching it, bending it, twisting it, and otherwise deforming it. Every point in the rubber starts at some (x,y,z) location in space, and we deform the block by moving each point to some new (x,y,z) location. This function we've been talking about essentially tells us where to move each point. A function that tells us how to reshape an object is a fairly common concept and is often called a deformer or deformation function.

If we allowed just any function, then it could do nasty things to the rubber. It might tear apart neighboring points, produce sharp creases, or collapse an entire volume down to a degenerate surface, line, or single point. The conjecture does not allow those sorts of things.

The conjecture requires that the function have its "Jacobian determinant" to be constant and nonzero. In this context, the Jacobian can be thought of as a set of vectors that describe how our rubber block is locally stretched/compressed in different directions. The determinant of the Jacobian then tells us how much the volume of the rubber is expanded/compressed at a given point. If the determinant is zero somewhere, then that is a location where, essentially, a volume of the rubber would be compressed down to zero.

This requirement that the determinant can't be zero anywhere, combined with the previous requirement that the function be polynomial, means that the rubber can stretch or compress, but it cannot pinch, crease, tear, or collapse. It must stay smooth everywhere. This rubber block fills all of space because the conjecture concerns a function defined on all of space. This means the function must be smooth and well defined everywhere. Otherwise, we could just cut any problem locations out of the block and pretend they weren't there.

Using this perspective, we can now express the simplified version of the Jacobian conjecture in plain language: Can we smoothly deform the infinite block of rubber so that at least two distant parts of it end up occupying the same place, even though nothing is ever locally pinched, creased, torn, or collapsed?

Keep in mind that even though this is a simplified version of the Jacobian Conjecture, it is still covered by the full formal version. If our simple version is not true, then the full formal version can't be true either.

What Alpöge found with Fable is actually rather clean and simple. It is a function that meets the Jacobian Conjecture's requirements, but it's still not invertible. Specifically:

- It is a polynomial function that maps 3D space to 3D space, so we can use it to deform our rubber block, *and*
- It's a polynomial with Jacobian determinant nonzero and constant, so the deformation doesn't have any pinches, creases, or other problems, but
- It also produces the same output for multiple inputs, so our deformed block of rubber would end up overlapping itself.

The first two points establish that the function satisfies the conjecture's requirements, but the third point contradicts the conjecture's conclusion. Therefore, the conjecture must be false.

At this point, we can also get a sense of why the Jacobian conjecture was so interesting to mathematicians and scientists. The condition I've described somewhat informally as "no pinch, crease, tear, or collapse" assures what mathematicians call local invertibility. Imagine zooming in closely enough to a little region anywhere in the rubber block, as though you were using a melon baller to scoop out a suitably small bite. This condition means that wherever you scooped, the bite would be free of self-intersections.

If the conjecture had been true, then knowing this about every little bite would have told us that the entire infinite block was globally intersection-free. The counterexample shows that this simply isn't true: every little bite can be perfectly well-behaved, but distant parts of the block can still end up occupying the same place. We lose the guarantee some mathematicians had hoped for, but now we know, somewhat surprisingly, that this strange kind of deformation is actually possible.

One concern with other mathematical results being produced by AI is that they may be very difficult for humans to understand and verify. However, that is not an issue here. We can verify the disproof using straightforward algebra and minimal calculus: check that the Jacobian determinant is constant and nonzero, and then check that at least two different inputs produce the same output.

The hard part was finding the right function to use as a counterexample. Mathematicians worked on the problem for nearly a century without finding a suitable function. A human working with AI solved it, in Alpöge's words, "during the World Cup final."

The counterexample works in three dimensions, so it would also work in four or more dimensions. Interestingly, the two-dimensional version of the conjecture, which one might have expected to be easier target, is still open. Perhaps you can coach an LLM into solving the 2D problem?

I hope this explanation is helpful to non-mathematicians who would like to better understand this interesting result. As stated up front, I've focused on the concrete three-dimensional counterexample that disproves the conjecture rather than the full formal framework surrounding it. I've used what I hope is an intuitive physical analogy, but, like any analogy, the match is imperfect. In the interest of keeping this broadly accessible, I've also avoided or glossed over some technical details. If you'd like a more rigorous discussion with all the crunchy math, this explanation by Terence Tao might be good reading for you.

About Me: James F. O'Brien is a Professor of Computer Science at the University of California, Berkeley. His research interests include computer graphics and animation, simulation of physical systems, human perception, rendering, artificial intelligence, computer vision, virtual reality, digital privacy and security, and image and video forensics.

If you would like to read my other articles, consider subscribing. You can also find me on LinkedIn, Instagram, and at UC Berkeley.

Disclaimer: Any opinions expressed in this article are only those of the author as a private individual. Nothing in this article should be interpreted as a statement made in relation to the author's professional position with any institution.