# Generalizing Diagrams: Alternative Media and Mathematical Justification

Thomas Sachen

Recent work by De Toffoli and others [1, 2] has begun to explore the justificatory role that certain diagrams play in mathematical justification. That is, there is reason to believe that diagrams subject to certain conditions carry real, non-redundant mathematical content and play a legitimate epistemic role in justifying mathematical beliefs. In this sense, these mathematical diagrams constitute legitimate mathematical notation, on the same standing as traditional arguments rendered in prose and conventional mathematical notation.

Henceforth, unless specified otherwise the term "prose" will describe a combination of prose in the naive sense with conventional mathematical notation, so for example a sentence like "*Therefore* $\sqrt{\int |\psi|^2}$ *is finite, so* $\psi \in L^2(\mathbb{R}^d)$" is termed an instance of mathematical "prose." Similarly, the term *diagram* describes a still diagram in the usual sense, found in settings such as Euclid's *Elements* or contemporary mathematics journals. Finally, the term "animation" can refer to either an animation in the naive sense (a series of still images played in succession to give the illusion of movement) or an interactive model, an example of which we'll see later on. The precise meaning of the term in any particular instance will be clear from context.

And while informal mathematical proofs (for these purposes, arguments that are considered rigorous by relevant audiences and that might appear in a mathematical journal) are for reasons peculiar to human technological development almost always presented in forms amenable to two-dimensional printing, taking the form of something that can be bound in a physical volume or downloaded as a PDF, it is not clear that this practice has any *philosophical* grounds for excluding forms of media apart from prose and diagrams. The practice of including diagrams in mathematical arguments owes in large part to the history of mathematics and its rich Euclidean tradition, though it's easily conceivable that alien mathematicians might not engage in such a practice. For example, the current framework imposes a *de facto* restriction of diagrams to dimension 2, even though 3D models (*i.e.* either 2- or 3-dimensional objects embedded in 3D space or projections of 4-dimensional objects in 3D space) could conceivably be permissible on the same grounds as 2D diagrams in informal mathematical proof, even in the current absence of requisite technological capabilities.

So, motivated by De Toffoli's conception of diagrams as forming mathematical notation that plays a "non-redundant role in proofs" and the notion that "there is a plausible conception of proof are not just abbreviations . . . but are in fact essential to those proofs," [2] I aim to expand the mathematician's toolbox beyond only that which can be printed on paper or drawn on a chalkboard. Such a departure from the methods of justification that mathematicians hold dear is quite a grand enterprise, so this paper will focus on only a few of the most accessible forms of alternative media in mathematics. I'll focus on the key differences animations and interactive models, concluding that many interactive models *do* constitute well-formed mathematical notation, while non-interactive animations do not.

I'll also explore the related notion of an interactive model, which serves to generalize and improve upon diagrammatic reasoning in mathematics. The key difference between an interactive model and a conventional diagram is that an interactive model allows the reader to protect against erroneous overgeneralization that might result from a misleading particular instance of a diagram. This leads me to conclude that interactive models are the "correct" generalization of diagrams in mathematics, as opposed to animations. This will be expanded upon later, but first we address a critical objection to the use of animations to justify mathematical belief.

## Slice Movies and Redundant Animations

This section addresses a natural objection to the idea that animations play a redundant role in mathematical justification, or in other words that an animation can be faithfully transformed into a combination of prose and diagrams with no loss of information. It also serves as an important illustration of a key deficiency of animations, which we'll flesh out more and try to remedy later, in our discussion of interactive models.

Right away, one might point out that any animation in the usual sense (say a .GIF or .MP4 file) is indeed nothing more than a rather fast-moving sequence of diagrams. In this respect, it's possible that there is no real difference between including an animation in a mathematical proof and simply printing each frame of the animation in sequence, regarding each one as a conventional mathematical diagram. Of course this scheme is flawed in practice — printing each frame of a 30 second animation at a frame rate of 40fps would require hundreds of pages. Even so, it is relatively common practice in *e.g.* knot theory to illustrate certain transformations using a well-chosen sequence of suggestive diagrams that demonstrate distinct stages of some topological process.

One example of such a sequence of diagrams is somewhat ironically called a "slice movie," common in the subject of knot concordance (see Figure 1). The slice movie is a particularly convenenient example, as it is a series of diagrams

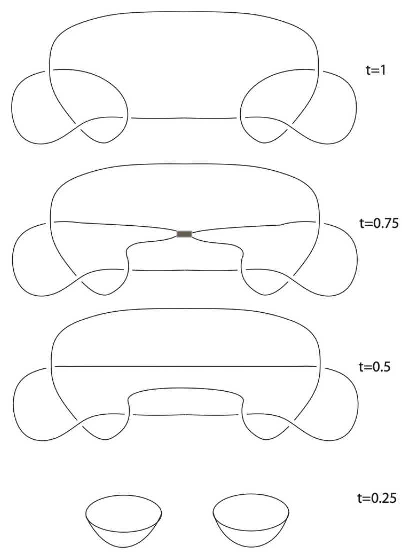


Figure 1: A slice movie for the square knot.*
∗ Figure adapted from [3]

endowed with an inherently time-dependent structure: the vertical direction in Figure 1 represents the time dimension, and the four diagrams are "snapshots" (the timings of which are denoted by the numbers on the right side) of a mathematical object undergoing a transformation over the period from time $t = 0.25$ to $t = 1$.

Without engaging in the precise mathematical goings-on of this diagram,[1] we may interpret Figure 1 simply as a depiction of two unlinked unknots undergoing the transformations of ambient isotopy from time $t = 0.25$ to $t = 0.75$ in conjunction with a connected sum at time $t = 0.75$ and another ambient isotopy from time $t = 0.75$ to $t = 1$. The reason for the choice of a snapshot at $t = 0.75$ is that it represents a critical point of a so-called Morse function for an associated "slice disk". The technical implications of that distinction aren't

[1] Livingston [3] gives a good introduction to the relevant mathematics.

relevant at the moment, apart from the fact that in some sense this sequence of diagrams adequately captures the only topologically "interesting" moment in the transformation of the pair of unlinked unknots at time $t = 0.25$ to the single square knot at time $t = 1$. As links are in many respects considered equivalent up to the ambient isotopies taking place at all other moments of this slice movie, the transformation at $t = 0.75$ is in some sense the only topologically relevant occurrence: this slice movie is really showing us that at time $t = 0.75 - \epsilon$ we still have two unlinked (though tangled up) knots, while at time $t = 0.75 + \epsilon$ we have a topologically distinct object – a single knot obtained via a connected sum.

In this case, nothing of the continuous nature of this transformation is lost by including only a discrete few diagrams. De Toffoli writes on this topic to justify the role of diagrams as mathematical notation in the first place, claiming that "[t]he use of diagrams triggers a form of manipulative imagination that gets enhanced by the practice. Thanks to this imagination, knot diagrams become an effective notation to make operations and calculations: According to specific aims, we can form sequences of diagrams connected by specific moves"[1]. In other words, these diagrams are so suggestive that they force the mind to "fill" in the blanks between these snapshots, and importantly they encode all relevant topological processes taking place: a connected sum at time $t = 0.75$ and ambient isotopies (possibly the identity isotopy) taking place at any other instant from $t = 0.25$ to $t = 1$.

Consider in contrast an honest-to-goodness slice "movie", *i.e.* a 2D animation that depicts the isotopies between these diagrams by increasing the resolution to a frame for every, say, increment of 0.1 in the $t$ dimension. It seems obvious in this case that an animation *does not* provide any additional epistemic benefit over the sequence of diagrams in Figure 1 – in fact it may even be less useful as a heuristic or pedagogical tool. The reasons for this are several-fold. There is really only one critical moment that the creator of this slice movie seeks to emphasize: the connected sum occurring at time $t = 0.75$. Animating this transformation rather lowers the emphasis placed on time $t = 0.75$, seemingly placing it at level of importance equal to every other frame of the animation.

And if somehow the creator of the animation indicates the importance of the snapshot at time $t = 0.75$ (maybe including a flash of color or slowing the animation down at this moment), it's unclear whether this will even be useful! A point mentioned by both Burgess and De Toffoli is that it's difficult in general to gain mathematical understanding from such an animation, as choices like speed or color of an animation are made by the creator independent of the mode of understanding best suited for a viewer. If a viewer finds an animation of a slice movie too fast-paced, he will be forced to pause the animation at any critical moments to grasp the key ideas at his own pace, essentially reducing the animation to a diagram. This apparent deficiency of animations is important to keep in mind, as we'll explore the notion of interactivity in the next section.

Hence an animation of this slice movie plays merely a heuristic role, as opposed to one carrying actual justificatory weight. The reason for this is that the main point of emphasis in this slice movie is the critical point at the *instant* $t = 0.75$, rather than properties of the time dimension itself such as the nature of a specific type of motion or another time-dependent process.

In general, I claim that an animation or model is merely heuristic (*i.e.* it plays no non-redundant role in justifying mathematical belief) if its main point of emphasis can be equally well communicated by a well-chosen and arranged series of frames interpreted as diagrams. Of course, the selection of diagrams should be made in good faith and intended to be viewed *as a discrete series of diagrams.* For example, a "well-chosen" subset of the frames of an animation cannot just be *all* the frames of an animation, intending for the reader to interpret this sequence as an object like the original animation. Instead this well-chosen subset should be a relatively small choice of particularly suggestive frames that unambiguously encode all relevant mathematical processes at play.

## Viviani's Theorem and Interactive Models

In order to address many of the concerns raised about the animation described in the last section, this section will explore the similar but distinct notion of *interactive models* in informal proofs. An example of such an object due to Doyle et al. [4] can be found at this link – in a moment I'll describe what this particular model is meant to show.

But first some definitions. An *interactive model* is a mathematical diagram with the added structure that the viewer is able to manipulate some of its *enabling* features while leaving its *constitutive* features fixed. These terms are meant in a sense analogous to De Toffoli definition of *enabling* and *constitutive* features of diagrams, just upgraded for the case of interactive models. A *constitutive* feature is one that has a specific and univocal mathematical meaning, while an *enabling* feature is one that is incidental to any one choice of diagram and whose precise value holds no mathematical content.[1]

For an interactive model designed to prove a certain theorem, the constitutive features will often correspond to the hypothesis of the theorem, whereas the enabling features correspond to the entire class of objects in the theorem's domain. Importantly, this means that exactly which features are considered constitutive and enabling is entirely context-dependent. Those features of an interactive model which are considered important and to which we ascribe precise mathematical meaning might differ based on the specifics of a theorem or proof we have in mind. An example is helpful here. Consider Viviani's Theorem, a well-known result in classical geometry which is as follows:[2]

> **(Viviani's Theorem)** *The sum of the perpendicular distances from a point in or on an equilateral triangle is equal to the altitude of the triangle.*

It's worth playing around with the interactive model linked above to convince yourself this is true. By the end of this section I'll conclude that the interactive model actually constitutes non-redundant *proof* of this theorem

In the interactive model linked above the primary constitutive feature is that the triangle is equilateral – indeed the truth of the theorem relies on this fact. If there were a slider that modified one of the triangle's side lengths then this wouldn't be a legitimate interactive model under this definition, as the viewer is able to freely manipulate a constitutive feature. However, the position of the point $P$ within the triangle is an important enabling feature. Note that this isn't an enabling feature of the interactive model itself, but rather an enabling feature of the diagram to which that particular "snapshot" of the interactive model corresponds. The theorem holds for any particular choice of the position of $P$ and this position isn't intended to be quantified or given precise mathematical meaning. Yet the role $P$ plays is an important one and is perhaps the key distinction between an interactive model and a diagram: it convinces the viewer that the theorem holds *in all generality* and provides assurance that there is no particularly misleading choice of $P$ that, if translated to a sequence of diagrams, would erroneously show the theorem true. The power of the interactive model is thus the ability to consider all possible manifestations of certain enabling features of a corresponding non-interactive diagram. For consider the "proof without words" of Viviani's Theorem in Figure 2, due to Kawasaki [6].

[2]this formulation is due to Alsina and Nelsen, from the book *Charming Proofs: A Journey Into Elegant Mathematics* [5]

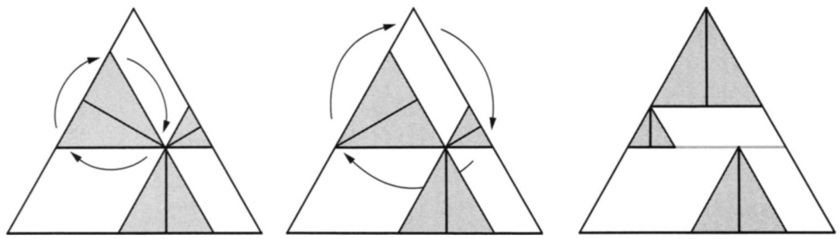

Figure 2: A series of diagrams illustrating Viviani's Theorem.*
∗ Figure due to Kawasaki [6]

Kawasaki was careful to choose a point of the triangle that "feels" arbitrary, in that it isn't too near the center or too near the boundary and no two of the three inner triangles are the same size. However such a "proof" as the one given in Figure 2 cannot be used as justification for Viviani's Theorem, as there is no reference to why the chosen point in particular accounts for the full generality of the problem. It certainly serves as a good heuristic to inform a belief of *why* the theorem might be true, but the reasoning it suggests doesn't adequately account for, for instance, cases where the point $P$ lies on the boundary of the triangle (Figure 3). Likewise *any* particular choice of $P$ would run into this problem of underspecification. It is worth emphasizing again that this is the key reason why interactive models cannot in all cases be faithfully transformed into diagrams or series of diagrams. Interactive models such as this account explicitly for the generality that diagrams try to evoke and hence serve as far stronger justificatory grounds.

Recalling the deficiencies of the animated slice movie discussed above, consider an animated version of the interactive model for Viviani's Theorem. No such animation could adequately account for all combinations of values of $\alpha$, $\beta$, and the position of $P$ and so it suffers the same fate as the mathematical diagram; a specific instance of a mathematical phenomenon does not justify

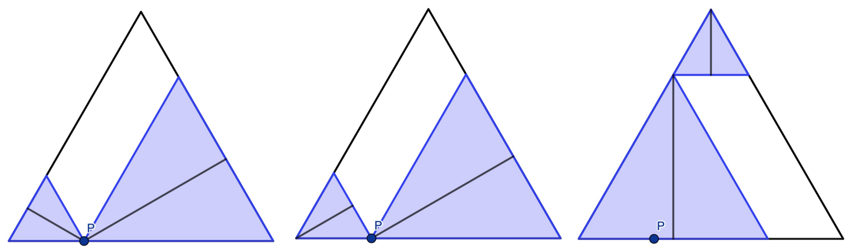


Figure 3: Using the interactive model to analyze a boundary case.

that phenomenon in more generality unless there is some account of *why* that instance is representative of the phenomenon's entire domain.

Similarly, in a way analogous to how changing the position of $P$ in the interactive model makes explicit the universal quantifier in the statement of the theorem, manipulation of the sliders $\alpha$ and $\beta$ account explicitly for inferential steps in the proof. The action of these sliders *can* be faithfully transformed into prose as something like the following:

> *Let $T$ denote the outer equilateral triangle, and let $T_1$, $T_2$, and $T_3$ denote the inner triangles, beginning with the top left triangle and labeling them clockwise. The $\alpha$ slider corresponds to a clockwise rotation of $T_1$ by $120°$ about its centroid. Such a rotation is a symmetry of $T_1$ and corresponds to the permutation of its three medians that sends the median representing the perpendicular distance from $P$ to the edge of $T$ intersecting $T_1$ to the median parallel to the median of $T_2$ representing the perpendicular distance from $P$ to the edge of $T$ intersecting $T_2$. The $\beta$ slider corresponds to a clockwise rotation of of $T_1 \cup T_2$ by $120°$ about the centroid of $T'$, the equilateral triangle by all points of $T$ that are at least as high as $P$. This rotation sends the indicated medians of the $T_i$ to medians of the images (under this rotation) of the $T_i$ that are parallel to the altitude of $T$.*

Note that it wouldn't have been enough to say something as simple as "the $\alpha$ slider rotates $T_1$ clockwise by up to $120°$," as this doesn't adequately capture the inferential steps taking place in an informal proof that the interactive model so eloquently demonstrates. It's also quite difficult to capture in prose *why* the altitude of $T_1$ plus the altitude of $T_2$ is the altitude of $T'$, but it becomes quite obvious after a few seconds using the interactive model.

Finally, I remark that interactive models satisfy De Toffoli's minimal constraints for forming a mathematical notation. In this instance De Toffoli was arguing for the legitimacy of certain types of diagrams as mathematical notation, but her criteria are easily adapted to the setting of interactive models. The point of this is to eventually conclude that interactive models carry epistemic weight and are in the that sense are far more than merely "subjective representations" [1] of mathematical objects.

The minimal constraints for mathematical notation introduced by Toffoli are the following [1]:

1. a notation should be cognitively accessible: its constitutive formal features should be clearly identified, persistent, and stable;
2. a notation should be reproducible: it should be possible for an average practitioner to copy its constitutive formal features with relative ease and reliability, possibly with the aid of different tools such as a straightedge and/or a computer;
3. a notation should support calculations and/or inferences: it should be possible for an average practitioner to perform reasonably simple manipulations corresponding to mathematical operations.

We can use the interactive model for Viviani's Theorem again, and check each of these criteria in turn.

1. Since the constitutive features of an interactive model correspond to the constraints given in the hypothesis of a given theorem, they are easily identifiable. Furthermore, the constitutive features of a particular interactive models are fixed by definition, and are therefore persistent and stable.

2. Simple geometric models like the one for Viviani's Theorem can be reproduced with minimal to no background in computer programming. This particular model was created using Geogebra, an easy to use software for creating geometric models. Certainly it's reasonable to assume that mathematicians could reproduce such a model quite easily.

3. The interactive nature of an interactive model is precisely what enables it to mathematical inferences. The manipulations of the interactive model induced by the $\alpha$ and $\beta$ sliders correspond to well-defined rotations. These rotations and manipulations like it are epistemic actions that exist are independent of the viewer's mind.

Thus in this sense we can reasonably conclude that an interactive model is a legitimate (and useful) form of mathematical notation.

## Conclusion

The point of this discussion isn't necessarily to advocate for the immediate use of interactive models in serious mathematics, but instead to continue the exploration of media to be considered rigorous in mathematical practice. I've taken a negative view toward animations in general, as in no cases I could think up did they provide actual justificatory benefit beyond well-chosen diagrams, though in many cases they work very well as heuristic tools to promote mathematical understanding.

My goal was to suggest that interactive models are the “correct” generalization of mathematical diagrams, as opposed to animations. And while I’ve introduced the idea of interactive models as justificatory tools, there remains much more work to do on this very complex topic, particularly in establishing informal proofs reliant on interactive models as capable of epistemic justification.

## References


[1] De Toffoli, Silvia, and Valeria Giardino. “Forms and Roles of Diagrams in Knot Theory.” Erkenntnis, vol. 79, no. 4, 2013, pp. 829–842., https://doi.org/10.1007/s10670-013-9568-7.

[2] De Toffoli, Silvia. ”Who’s Afraid of Mathematical Diagrams?” Draft.

[3] Livingston, Charles, et al. “A Survey of Classical Knot Concordance.” Handbook of Knot Theory, Elsevier Science, Burlington, 2005.

[4] Doyle, Tim, et al. “Proofs without Words and Beyond.” Convergence, 2014, https://doi.org/10.4169/convergence20140801.

[5] Alsina, Claudi, and Roger B. Nelsen. “6.3: Viviani’s Theorem.” Charming Proofs: A Journey into Elegant Mathematics, Mathematical Association of America, Washington, 2014.

[6] Kawasaki, Ken-Ichiroh. “Proof without Words: Viviani’s Theorem.” Mathematics Magazine, vol. 78, no. 3, 2005, pp. 213–213., https://doi.org/10.1080/0025570x.2005.11953328.