Algebraic Visualization Design for Pedagogy

Gordon Kindlmann∗
University of Chicago

Carlos Scheidegger†
University of Arizona

ABSTRACT
We report on successes and challenges in using our recently-proposed Algebraic Visualization Design as the basis for teaching visualization. The experiences suggest that the principles in that paper can serve as unifying framework for introducing the fundamentals of data visualization, as long as the overtly mathematical flavor of the framework is not a disincentive. We find that the better we can relate the framework to ideas to which students have already been exposed, the better they understand it.

1 INTRODUCTION

1.1 Summary of Algebraic Visualization Design
In an InfoVis 2014 paper [9], we proposed an abstract mathematical way to describe problems and properties of visualizations, called Algebraic Visualization Design (AVD), reviewed here.

Our model describes relationships between three elements of visualization: the space \( D \) of “data” to be visualized, the space \( R \) of data representations, and the space \( V \) of visualizations. While “data” normally refers to the information input to a visualization algorithm, we distinguish between the underlying object of interest (e.g. a set of numbers), and its actual representation on a computer (e.g. an array in memory). The mappings between these spaces are captured in a single equation, shown with its commutative diagram:

\[
\begin{align*}
\alpha & \colon D \to R \\
\omega & \colon R \to V \\
\alpha \circ \omega & \colon D \to V
\end{align*}
\]

Lowercase \( r \) maps from data \( D \) to representation \( R \), and lowercase \( v \) is the visualization method mapping from \( R \) to visual stimulus \( V \). The mappings from \( D \) back to \( D \) (or mappings on \( D \)), are termed data symmetries, denoted \( \alpha \). The visualization symmetries, denoted \( \omega \), are mappings on \( V \). The identity maps on \( D \) and \( V \), which send each element to itself, are \( 1_D \) and \( 1_V \), respectively.

Using alpha and omega emphasizes that we are interested in the relationship between the very beginning and the very end of the visualization pipeline, while seeking to abstract away as much as possible of the internal details. In a commutative diagram, when two nodes are connected by two paths, the composition of functions along either path must be the same. The equality in (1) is between two possible paths from the upper-left \( D \) to the lower-right \( V \). Going down then right is \( D \xrightarrow{\alpha} R \xrightarrow{r_2} V \), or in terms of function composition (read right to left) \( v \circ r_2 \circ \alpha \), and going right then down is \( D \xrightarrow{\alpha} R \xrightarrow{r_1} V \) (or \( \omega \circ \alpha \circ r_1 \)).

AVD says that in successful visualizations, important changes in the data \( \alpha \) are well-matched, via (1), with obvious visual changes \( \omega \). The \( \alpha \) are answers to the question, “If the world had been different in some interesting and important way, how would the data have been different?” If one has data for which visualization may provide insight, the significance of that insight hinges on knowing, at a general level, what phenomena or relationships may be captured by the data. The \( \alpha \) embody the low-level tasks that the visualization is designed for. These \( \alpha \) describe which structures from the data/operation abstraction layer will be reflected in the encoding/interaction technique layer [12]. The \( \alpha \) are not a property of the data itself; they instead capture the kinds of questions about the data that we want the visualization to answer. For the same data, different users may have different \( \alpha \).

Obvious visual changes \( \omega \) are aligned with elementary perceptual tasks, such as judgments of position along a common scale, or of length and direction [3]. When possible, the \( \omega \) generated by the chosen \( \alpha \) should coincide with perceptually pre-attentive channels [7]. However, applying AVD assumes more than just identifying perceptual channels: it connects to the mathematical structure of those channels. In color perception, for example, opponent color theory [15][17] endows the space of colors with a kind of negation (green is “negative” red, orange is “negative” blue). Thus, when designing color scales for data in which negation is a meaningful data change \( \alpha \), one can evaluate a color scale by seeing whether the corresponding visual change \( \omega \) aligns with finding opponent hues.

From (1), AVD derives three principles, and gives concrete names to ways in which a visualization may fail the principles, summarized in Table 1. Calling these “principles” may suggest that these are tools of judgment, but we intend them to be more descriptive than prescriptive. The Principle of Representation Invariance says that from the same dataset, different visualizations will not arise simply by changing the representation of the same underlying data. If this is not the case, there is some mapping on representations, called the hallucinator, with a non-trivial effect on the visualization. This notion of invariance is rooted in Stevens’s statement of invariantive statistics, which gave specific mathematical meaning to categorical, ordinal, interval, and ratio data scales [13]. Conversely, the Unambiguous Data Depiction Principle says that an interesting \( \alpha \) applied to the data should induce a non-trivial \( \omega \). If this is not the case, then there is an injectivity failure [20], and we give a name to the \( \alpha \) for which \( \omega = 1_V \): a confuser. Visualization design requires trade-offs; designers may accept some confusers (data changes that are invisible) when they do not relate to the intended analysis tasks, to ensure that the \( \alpha \) are shown clearly. Finally, assuming both Invariance and Unambiguity hold, the Correspondence Principle is satisfied when neither \( \alpha \) nor \( \omega \) is the identity, and they solve (1) in a particular way, notated “\( \alpha \equiv \omega \)” to suggest congruence [15] between the data and visualization symmetries. The Correspondence Principle says that \( \omega \) somehow makes sense, given \( \alpha \). The visualization viewer would then be able to infer the \( \alpha \) from \( \omega \), or, the visualization makes the \( \alpha \) legible. In the setting of metric spaces on both data and visual perception, Correspondence is closely related to the notion of visual embedding [7]. When an important \( \alpha \) is not legible because the associated \( \omega \) is hard to understand, then the visualization as a jumbler. Conversely, when the apparent structure of the visualization suggests some obvious visual symmetries \( \alpha \), those should correspond to important data symmetries \( \alpha \). If not, the visualization has a misleader. Diverging color scales for ratio data, for example, tend to satisfy the Correspondence Principle by mapping negation of data to negation of color hue. Considering visualizations not just as functions mapping from data to visual stimuli, but as functors mapping from changes in data to changes in the visual, and then capturing properties of the visualization that may be problematic in

∗e-mail: glk@uchicago.edu
†e-mail: cscheid@email.arizona.edu
terms of those functors (Table 1), is the basis for the “algebraic” in algebraic visualization design.

Our hope, however, is that no prior understanding of functors or algebra is necessary to use AVD for visualization practice or pedagogy. General rules of thumb like Tufte’s “show data variation, not design variation” [14] become more actionable when, faced with a new or unsatisfactory visualization, one has a concrete course of action for evaluating or improving the visualization. AVD empowers students to ask pointed questions about what a visualization does or does not show, and AVD supplies a terminology for the problems revealed, which helps focus subsequent discussion and redesign. If the data were different, would the visualization be different (Unambiguous), and, different in an informative way (Correspondence)? If it is ambiguous: what are the data changes we are blind to? (Confuser)? If it is not informative: how else can we lay out or encode the data to create a better correspondence (removing Jumblers)? Are there apparent properties in the visualization that are not actually in the data (Misleaders)? Could the visualization have ended up looking different, due only to changes (Hallucinators) in the computational/numerical representation that should be inconsequential, but are not (Invariance)?

1.2 “It was my understanding there would be no math”

Often, data visualization students will not be computer science majors, and even CS students will have relatively little mathematical training. We personally value the rigor that the underlying theory provides, but it is natural to wonder whether the notation behind AVD is worth the price. In our experience, the most important consequence of using AVD in the classroom is that it forces students to think about the interplay between \( \alpha \) and \( \omega \). The central insight is the notion that every change in the data — every possible different world — corresponds to a possible change in the visualization. If the change in the data does not change the visualization, this is also important information.

In other words, the mathematics are there to ensure us that we can make the notions precise. But no math is needed to use AVD while thinking about design: we imagine possible different worlds where the data were different in an interesting way, and think about what that change does to the visualization.

2 Successes and Challenges

We give some examples of how we have used AVD to successfully explain some principles of visualization design in our classes, and highlight things that have repeatedly arisen as points of confusion.

2.1 Banking to 45

The “Banking to 45” rule of thumb [13] for determining the aspect ratio of a 2D function plot is a simple introductory example of what can be directly derived from the Correspondence Principle. We start with the recognition that the purpose of the plot may not be just to show values of a function \( y = f(x) \), but also its derivative \( m = f' = dy/dx \). This leads to considering a particular \( \alpha \) at some point on the graph: \( \alpha(m) = m \pm \Delta m \). This asserts that small changes in slope are interesting and important to see. The corresponding \( \omega \) will be some local change in the slope of the plot. The slope (rise over run) of a line is not an elementary perceptual channel, but line orientation is [3]. “Banking to 45” answers how to map the chosen \( \alpha \) to a \( \omega \) that on average is as clear as possible.

Prior to any visualization, the tangent to \( y = f(x) \) has slope \( m = \Delta y/\Delta x \). Once plotted on the page with vertical \( v \) and horizontal \( h \) coordinates, the slope is \( \Delta v/\Delta h = Rm \), where \( R \) is determined by the aspect ratio of the plot. Viewers perceive the line orientation \( \theta = \tan^{-1}(Rm) \). The Correspondence Principle tells us to maximize the change \( \omega \) in the orientation for a given change \( \alpha \) in slope, i.e., find the \( R \) that maximizes \( d\theta/d\alpha = d/d\alpha \tan^{-1}(Rm) = \frac{r}{1 + mr^2} \). Using calculus, students solve \( d/d\alpha \tan^{-1}(Rm) = 0 \) and find that \( R = \pm \frac{1}{m} \Rightarrow \theta = \tan^{-1}(\pm 1) = \pm 45^\circ \). While this solves Correspondence for a single point, it gives students a starting point for learning about considering average plot slope, or even multi-scale slope [8].

2.2 Including Zero in Plot Scale

AVD may help reduce confusion about the necessity of including zero in the vertical scale of a 2D function plot. A blog post by Andy Cotgreave [1] shows an example of when including zero is unhelpful, summarized in Fig. 1, which plots the world record times for men’s 100m dash over the last century. Cotgreave discounts Fig. 1(a), which includes zero in the vertical axis as having four problems: “It doesn’t really expose the change of the record over time”, “It especially doesn’t highlight the impact Usain Bolt had on the record”, “It doesn’t make great use of the space – there’s lots of dead space.”, and “It’s boring.”. Through a class discussion, students were encouraged to give mathematical “teeth” to these statements, using AVD. This in turn requires students to isolate the data changes \( \alpha \) that are important to show: the changes in the data that would arise in some interesting but plausible alternative universe.

Being a time interval, the data contains ratio values (for which zero is an intrinsically meaningless), but the interesting \( \alpha \) are not ratio changes \( \alpha(x) = bx \). The noteworthy changes have the form \( \alpha(x) = x - d \), an example of how ratio-valued data may be effectively interval-valued (like degrees F or C). AVD answers the question of “should I include zero?” with another question requiring some awareness of the purpose of the visualization: “do the relevant \( \alpha \) depend on zero?”. For the race times, \( \alpha(x) = x - d \) do not involve zero

Table 1: Algebraic visualization design principles for evaluating a visualization method \( \nu \), expressed in terms of \( \nu \).

<table>
<thead>
<tr>
<th>Principle Name</th>
<th>Precondition</th>
<th>Requirement</th>
<th>Name for failure</th>
<th>Failure definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representation Invariance</td>
<td>( \alpha = 1_D )</td>
<td>( \omega = 1_V )</td>
<td>Hallucinator</td>
<td>( H(\nu) = { h</td>
</tr>
<tr>
<td>Unambiguous Data Depiction</td>
<td>( \omega = 1_V )</td>
<td>( \alpha = 1_D )</td>
<td>Confuser</td>
<td>( C(\nu) = { \alpha</td>
</tr>
<tr>
<td>Visual-Data Correspondence</td>
<td>( \alpha \neq 1_D, \omega \neq 1_V )</td>
<td>( \alpha \cong \omega )</td>
<td>Jumblers and Misleaders</td>
<td>(see Sec. 1 text)</td>
</tr>
</tbody>
</table>

Figure 1: History of world record times in men’s 100m dash, plotted with (a) or without (b) including zero in the vertical scale.

[^1]: http://grayvaneudote.com/uncategorized/mythbusters-should-you-start-your-axes-at-zero
because no plausible record time will ever be close to zero, and the
importance of Usain Bolt’s record is from \( d \), not \( -\frac{2}{d} \). Correspond-
dence suggests using available space to show all such \( \omega \) as clearly
as possible, disregarding zero as a special value. Cotgreave suggests
that not including zero is “bending or breaking” a “rule”. We suggest
AVD provides a simpler consistent framework.

2.3 Misleading Plots

Even when a visualization is widely recognized as being misleading,
AVD provides a mathematical way of describing what exactly is
misleading about it. Representative Jason Chaffetz (R Utah) used
Fig. 3(a) to question Planned Parenthood funding. The graph title
“Abortions up – Life-saving procedures down” suggests a negative
correlation between the two. However, the red and pink lines have
different vertical scales (a dual-axis plot), preventing visual com-
parison of their slopes. This plot has been called misleading but
students were pressed to articulate precisely why. The pedagogical
value of discovering such a mathematical basis is that it facilitates
generalization and application to other examples.

With some guided class discussion, students considered an \( \omega \)
that mathematically expresses the relationship implied by the “Abor-
tions up – Life-saving procedures down” title: what if the plots are
reflected across a horizontal axis? This negates the trends by effect-
ively swapping the two plots, while still interpreting each relative
to its original vertical scale. The simple X-shaped visual structure
of plot supports this clear \( \omega \). The Correspondence Principle then
asks whether the \( \alpha \) corresponding to \( \omega \) is an interesting or important
data change. Figure 2(c) shows the \( \alpha \) by a 2D parametric plot over
time. Because the vertical scale for abortions was so exaggerated in
the original plot, there is very little change along the abortion axis.
This is not an especially interesting or significant \( \alpha \). On the other
hand, with a single vertical axis and the X-shaped plot structure, the
same \( \omega \) would lead to a very different \( \alpha \) shown in Fig. 2(d). Here
\( \alpha \) preserves a significant negative correlation. The absence of a
significant \( \alpha \) corresponding to a clear \( \omega \) is one way of articulating
why Fig. 3(a) is misleading.

As a teaching example, the challenge is getting students to recog-
nize and state the visibly apparent \( \omega \). This requires a kind of thinking
that may arise from modern interactive visualization tools, but which
we find to be nonetheless uncommon. Students must shift their think-
ing from “from this visualization, what can I learn about the (single)
dataset it came from?” to “how could I manipulate the visualization
itself, and what changes in the data could give rise to that?” (exclud-
ing parameter changes affecting appearance). This approach may be
fostered with a human-computer interface perspective that analyzes
the visualizations in terms of their affordances: apparent opportuni-
ties for direct manipulation.

2.4 Kernel Choice by Representation Invariance

When teaching “scientific visualization” methods such as volume
rendering or streamlines, students must reconstruct a continuous
field from a discrete grid of regular samples, by convolving the data
with a reconstruction kernel. The literature on kernel choice can
be somewhat daunting, but AVD provides a way of introducing it
in a concrete way. Müller et al. (and others previously) suggest
that kernel accuracy can be understood not just in the traditional
terms of pass-bands in frequency space, but in terms of Taylor series,
which expands a function according to a series of derivatives. More accurate kernels are those that push the introduction of error
terms to higher-order terms in the Taylor series. This means that for
some low-order polynomial, a kernel may exactly reconstruct the
function from discrete samples.

Figure 3 shows this with two different samplings of the same func-
tion \( f(x) = -1.5x + x^2 \). In this case, the “data” is the underlying
function \( f(x) \), and the “representations” are different possible regu-
lar samplings of the function, producing some array of values. The
visualization \( v_{cont} \) with linear interpolation creates a plot (Fig. 3(a))
that reveals the sampling specifics. A change in representation \( \hat{h} \) (a change in how the function is sampled) is a hallucinator because
it caused a visible change \( \omega \neq 1_v \) in the visualization. Visualizing
\( v_{cont} \) with more accurate kernel, such as Catmull-Rom, the quadratic
function in question can be reconstructed exactly, removing the hal-
lucinator. The plot is the same regardless of sampling. Implementing
1D convolution, and testing it by plotting data for which the correct
answer is known a priori (similar to Fig. 3) was the first assignment
in a scientific data visualization class.

2.5 Are Confusers Bad?

Students have frequently noted that a wide range of visualization
tools have confusers, from dimensionality reduction methods
that necessarily project out many degrees of freedom in the data
to things as simple as summary statistics (e.g. mean, co-
variance), which may still have a role in a visualization, even though
Anscombe’s quartet famously illustrates their confusers. Students
then ask, “So how can you possibly avoid all confusers?” We believe
this highlights two things. First, we have to date failed to sufficiently
emphasize that these principles, and the descriptive terms for their
failures (like confusers) are only descriptive. The point is not to
simply eliminate confusers: any sufficiently interesting visualization
must have confusers. The point is to encourage students to think
critically and concretely about the space of possible data changes \( \alpha \),
and to then make an informed distinction between the important \( \alpha \)
that should have clearly visible \( \omega \), and the non-essential \( \alpha \) that may
be confusers (with faint or invisible or \( \omega \)).

Second, we sense that the culture around data visualization, at
least as represented by students’ ideas about why they should learn
visualization, is oriented around an ideal of “quality” or “excellence”
that is an objective and intrinsic property of a visualization, which
implies that a role of design is to judge a bad visualization from a
good one. The quality of a visualization, however, is much more
dependent on context, purpose, and audience: an infographic that
effectively engages a reader in a journalistic setting satisfies very dif-
ferent needs than a figure in a scientific research paper. The reflexive
assertions about the necessity of including zero in the vertical scale
can be a specific example of a judgment that can be made more
helpful with consideration of the \( \alpha \) that actually matter, given the
culture context, as noted above. AVD provides a vocabulary to teach
visualization design as a certain kind of informed comparison and
promise, but it is not a deterministic machinery to either create
visualizations or increase their quality.

2.6 Nomenclature

The nomenclature of AVD has benefits and drawbacks. Students
seem to like the terms “hallucinators” and “confusers”, because
they name specific things that can be tested and compared. The
difference between “misleaders” and “jumblers” is harder to convey,
consistent with our original concern about slippage between
these terms. Still, having some way of distinguishing Invariance,
Unambiguity and Correspondence failures is appreciated because
it is more fine-grained than existing terms like “expressive” and
“effective”. A visualization can fail to be expressive, for example,
either by suggesting non-existent facts about the data (against
Invariance, with a hallucinator) or by failing to show something
important about the data, in which case either be because some
important \( \alpha \) is invisible (against Unambiguity, with a confuser) or
because it is shown in a confusing way (again Correspondence).
For future classes, we may adopt the term “non-correspondence” or
“incongruity” to refer collectively to jumblers and misleaders.
A more fundamental point of confusion arises with “data” versus “representation”. Distinguishing them risks of a tautology relative to the hallucinator, but we could say that representation is where changes should not have an effect (e.g. representing one set with two different orderings of elements in an array), whereas data is where changes should have an effect (e.g. two consequentially different orderings of elements in an list). Ambiguity between hallucinators and confusers is possible. It is also disorienting for students to hear that the input to the visualization method is not data, but a representation, given the ubiquitous use of “data” to mean whatever is stored in files and processed by programs. Noting that a tool to visual a spreadsheet should give the same result if the rows were re-ordered is one possibly helpful example. Appealing to standard computer science pedagogy, we can also say that the difference between data and representation is analogous to the difference between an abstract data structure and its implementation, respectively: the interface to a data structure should be independent of the implementation, and it is a bug to have the interface expose specifics of the implementation. More philosophically, especially with a statistical consideration of “data” as underlying distributions or functions (e.g. Fig. 3) and “representation” as some specific sampling of the distributions, students can reasonably ask, “So is the purpose of visualization to look at the numbers we have, or is to infer something about whatever generated the numbers?” The answer is of course “yes; it depends”, which can lead into a discussion about the range of possible tasks in visualization, from mundane data quality checks, to discovering general data trends.

3 Discussion

Our experience with teaching visualization via A VD is mainly limited to young computer science undergraduates, who typically imagine that assigned problems have solutions, and that solutions are either correct or incorrect (though those with some software engineering experience may recognize a more flexible space of possible answers). Using A VD, students must answer two essential questions that do not have obviously correct answers (from Sec. 1): (1) “If the world had been different in some interesting way, how would the data have been different?” (what are the important \( \alpha \)), and (2) “How then would the visualization look different, and would that make visual sense?” (what are the resulting \( \omega \), and does \( \omega \approx \omega \).

The first question forces students to expand their thinking from data visualization to the larger context of the system or investigation that generated the data. The importance of a given \( \alpha \) is not just a property of the data, but is derived from the original motivations to measure, organize, and understand the data. The second question forms a linkage into perceptual psychology and human-computer interface design. Especially when teaching visualization with tools like D3 [2] that facilitate interaction, students can use buttons to trigger specific \( \alpha \), and observe the resulting \( \omega \). Based on our limited experience, guiding students to consider specific \((\alpha, \omega)\) pairs in terms of Unambiguity and Correspondence is more actionable than trying to adhere to general guidelines (e.g. “maximize the data-ink ratio” [4]), and it provides a concrete path towards understanding visualization design as a mode of critical thinking.

REFERENCES


