1 Notation & Definitions

In this section we introduce a mathematical description of the visualization pipeline where artist A functions transform data of type $\Gamma(E)$ to an intermediate representation in prerendered display space of type $\Gamma(H)$:

$$A : O(E) \to O(H) \tag{1}$$

$$A : \tau \to \rho \tag{2}$$

- A is the function that converts an instance of data $\Gamma(E)$ to an instance of a visual representation $\Gamma(H)$
- E is a locally trivial fiber bundle over K representing data space.
- K is a triangulizable space encoding the connectivity of the observations in the data.
- H is a fiber bundle over S representing visual space
- S is a simplacial complex of triangles encoding the connectivity of the visualization of the data in E
- $\tau : K \to E$ is the data being visualized
- $\rho : S \to H$ is the render map

When E is a trivial fiber bundle $E = F \times K$, it can be assumed that all fibers F_k over $k \in K$ are equal. Fiber bundles are product spaces of toplological spaces, which are a set of points with a set of neighborhoods for each point[5, 10].

1.1 Data Model

We use a fiber bundle model to represent the data, as proposed by Butler [2, 3]. A fiber bundle is a structure (E, K, π, F) consisting of topological spaces E, K, F and the map from total space to base space:

$$\begin{tikzcd} F \arrow[r, hook] & E \arrow[d, dashed, "\pi"'] \\ & K \arrow[u, bend right] \end{tikzcd} \tag{3}$$

where there is a bijection from F to every fiber F_k over point $k \in K$ in E and the function $\pi : E \to K$ is the map into the K quotient space of E. Every point in the base space $k \in K$ has a local open set neighborhood U [5, 10]

$$\begin{tikzcd} \pi^{-1}(U) \arrow[r, dashed, "\varphi"] \arrow[d, dashed, "\pi"'] & U \times F \arrow[dl, "\mathrm{proj}_U"] \\ U & \end{tikzcd} \tag{4}$$

such that $\varphi : \pi^{-1}(U) \to U \times F$ is a homeomorphism where π and proj_U both map to U and the fiber over k $F_k = \pi^{-1}(k \in K)$ is homomorphic to the fiber F.

The section τ is the mapping $\tau : K \to E$

$$\begin{tikzcd} F \arrow[r, hook] & E \arrow[d, bend right, "\pi"'] \\ & K \arrow[u, bend right, "\tau"'] \end{tikzcd} \tag{5}$$

such that it is the right inverse of π

$$\pi(\tau(k)) = k \text{ for all } k \in K \tag{6}$$

In a locally trivial fiber bundle, $E = K \times F$ [5, 10]:

$$\tau(k) = (k, g(k)) \tag{7}$$

where the domain of $g(k)$ is F_k and returns a data point r. The space of all possible sections τ of E is $\Gamma(E)$. All datasets $\tau \in \Gamma(E)$ have the same variables F and connectivity K but can have different values such that $\tau_i \neq \tau_j$.

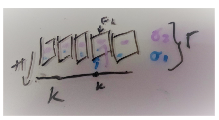
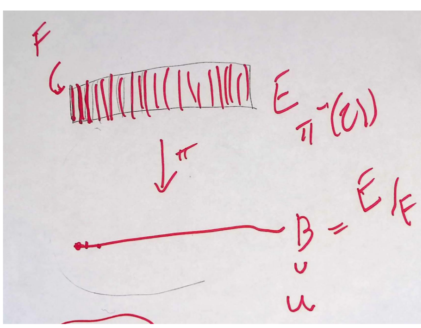


Figure 1: write up some words here

As illustrated by figure 1, the vertical lines F are the range of possible temperature values embedded in the total space E. The base space K of the fiber bundle is a line because the data points r in E are on a space that is continous in one dimension.

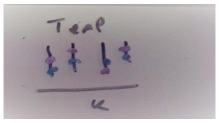
1.1.1 Base Space K

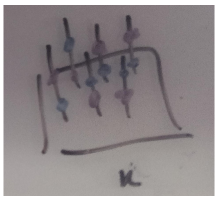


K is the quotient space of E, meaning it is the set of equivalence classes of elements r in E defined via the map $\pi : E \to K$ that sends each $r \in E$ to its equivalence class in $[r] \in K$ [9].

As shown in figure **??**, the fibers F divide E into smaller spaces consisisting of F and an open set neighborhood around F. This subdivision is projected down to the toplology $\mathcal{T}$

$$\mathcal{T}_k = \{U \subseteq K : \{r \in E : [r] \in U\} \in \mathcal{T}_E\} \tag{8}$$





where $[r] \in U$ is the point $k \in K$ with an open set surrounding it that has an open preimage in E under the surjective map $\pi : r \to [r]$.

in figure **??**, temperature is the only one data field in r but the K base spaces are different. subfig[1] is a timeseries, so the temperature in r at time t is dependent on the temperature in r_{t-1} and the temperature in r_{t+1} is dependent on r_t; this connectivity is expressed as a one dimensional K where K is the number line. In the case of the map, every temperature in r is dependent on its nearest neighbors on the plane, and one way to express this is by encoding K as a plane. K does not know the time or latitude or longitude of the point as those are metadata variables describing the k rather than the value of k. The mapping $\tau : K \to E$ provides the binding between the key $k \in K$ and the value r in E [7].

1.1.2 Fiber Space F

We use Spivak's formalization of data base schemas as the basis of our fiber space F [12]. He defines the type specification

$$\pi : U \to DT \tag{9}$$

where DT is the set of data types (as identified by their names) and U is the disjoint set of all possible objects x of all types in DT. This means that for each type $T \in DT$, the preimage $\pi^{-1}(T) \subset U$ is the domain of T, and $x \in \pi^1(T) \subset U$ is an object of type T. Spivak then defines a schema (C, σ) of type π, where π is the universe of all types, such that

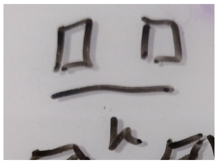
$$\sigma : C \to DT \tag{10}$$

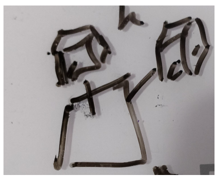
where C is the finite set of names of columuns, which we generalize to data fields in E. The set of all values restricted to the datatypes in DT is U_σ

$$\begin{tikzcd} U_\sigma \arrow[r] \arrow[d, "\pi_\sigma"] & U \arrow[d, "\pi"] \\ C \arrow[r, "\sigma"] & DT \end{tikzcd} \tag{11}$$

The pullback $U_\sigma \coloneqq \sigma^{-1}(U)$ restricts U to the datatypes of the fields in C such that U_σ is the fiber product $U \times_{DT} C$, and the pullback $\pi_\sigma : U_\sigma \to C$ specifies the domain bundle U_σ over C induced by σ. The fiber F is the cartesian product of all sets in the disjoint union U_σ.

For each field $c \in C$, the record function $r : C \to U_\sigma$ returns an object of type $\sigma(c) \in DT$. The set of all records $\Gamma(\sigma)$ is the set of all sections on U_σ. Spivak defines the τ mapping from an index of databases K to records $\Gamma(\sigma)$ as $\tau : K \rightarrow \Gamma(\sigma)$. This is equivalent to $\tau : k \rightarrow E$ since $F = \Gamma(\sigma)$ and F is the embedding in E on which the records r lie.





The fiber in figure **??** is the space of possible temperature values in degrees celsius, such that $F = [temp_{min}, temp_{max}]$ and is named Temp. In figure **??** time is encoded as a second dimension. This means that the set of possible values F with $C = \{\text{Temp}, \text{Time}\}$:

$$F = [temp_{min}, temp_{max}] \times [time_{min}, time_{max}] \tag{12}$$

and the function τ that retrieves records from F is

$$\tau(k) = (k, (r : \text{Temp} \rightarrow temp, r : \text{Time} \rightarrow time)) \tag{13}$$

$$temp \in [temp_{min}, temp_{max}], time \in [time_{min}, time_{max}]) \tag{14}$$

Since $\tau(k) = (k, r)$, *temp* is bound to a named data field and *sigma* binds *temp* to a temperature data type.

1.1.3 Sheaf and Stalk

As described in equation 4, there is a local space $U \subset K$ around every k. The inclusion map $\iota : U \rightarrow K$ can be pulled back such that $\iota^* E$ is the space of E restricted over U.

$$\begin{tikzcd} \xi^* E \arrow[d, "\pi"'] & E \arrow[l, "\xi^*"'] \arrow[d, "\pi"] \\ U \arrow[u, bend right, "\xi^*\tau"'] & K \arrow[l, "\xi"] \arrow[u, bend right, "\tau"'] \end{tikzcd} \tag{15}$$

The localized section of fibers $\iota^*\tau : U \rightarrow \iota^* E$ is the sheaf $O(E)$ with germ $\xi^*\tau$. The neighborhood of points the sheaf lies over is the stalk $\mathscr{F}_k$ [11, 13]

$$\iota^{-1}\mathscr{F}(\{k\}) = \varinjlim_{k \subseteq U} \mathscr{F}(U) = \varinjlim_{k \in U} = \mathscr{F}_k \tag{16}$$

which through ι gets the data in E at and near to k. Restricting the artist to the sheaf means the artist knows the data in F and also has access to derivatives of the data. This property is useful for some visual transformations.

1.2 Prerender Space

Every point $k \in K$ maps to a space $S_k \in S$, which is the topology of the output of the artist A. The space H is a total space representing the predisplay space, with a fiber dependent on the render space and a base space of §:

$$\begin{tikzcd} D \arrow[r, hook] & H \arrow[d, shift right, dashed, "\pi"'] \\ & S \arrow[u, shift right, dashed, "\rho"'] \end{tikzcd} \tag{17}$$

where $\rho : S \to H$ is mapping from a region s on a mathematical encoding of the image to a region xy on the screen that the renderer then maps to pixel space. For a physical screen display, the predisplay space is a trivial fiber bundle $H = \mathbb{R}^7 \times S$ such that ρ is

$$\rho(s) = \{x, y, z, g, b, a\} \tag{18}$$

To draw an image, a region, H is inverse mapped into a region $s \in S$ where

$$s = \rho^{-1}_{XY}(xy) \tag{19}$$

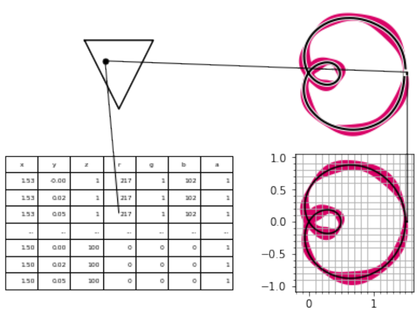
such that the rest of the fields in $\mathbb{R}^7$ are then integrated over s to yield the remaining fields in p

$$R(p) = \oint_s \rho_R(s) ds^2 \tag{20}$$

$$G(p) = \oint_s \rho_G(s) ds^2 \tag{21}$$

$$B(p) = \oint_s \rho_B(s) ds^2 \tag{22}$$

Here we assume a single opaque 2D image such that the z and *alpha* fields can be omitted.



x	y	z	r	g	b	a
1.53	-0.00	1	217	1	102	1
1.53	0.02	1	217	1	102	1
1.53	0.05	1	217	1	102	1
...	...	...	...	...	...	...
1.50	0.00	100	0	0	0	1
1.50	0.02	100	0	0	0	1
1.50	0.05	100	0	0	0	1

Figure 2

As illustrated in figure 2, words.

1.3 Artist

The artist is a mapping from the sheaf $O(E)$ representing the data to a pre-render space $O(H)$.

$$A : O(E) \to O(H) \tag{23}$$

The artist acts on a sheaf because some visual characterstics, such as line thickness, need more information than just the data at a point. The artist through $F : E \to V$ maps from data space to visual variable [1, 4, 8] space

$$\begin{tikzcd}
E \arrow[r, "\nu"] & V \\
K \arrow[u, "\tau"] \arrow[ur, "\mu"'] &
\end{tikzcd} \tag{24}$$

such that $\nu : \tau \mapsto \mu$ is a homorphism where τ and μ are equivariant such that the properties of the field type persist in the visual representation. The map $\xi : S \to K$ goes from a region s to it associated k

$$\begin{tikzcd}
E \arrow[d, "\pi"] & H \arrow[d, "\pi"] \\
K \arrow[u, bend right, "\tau"'] & S \arrow[l, "\xi"] \arrow[ul, "\xi^*\tau"']
\end{tikzcd} \tag{25}$$

and can be pulled back up to E (or V) such that for a region s there is an associated record τ and visual mapping μ. The visual fiber bundle V gets pulled back over S via ξ such that

$$\begin{tikzcd}
\xi^*V \arrow[r, "q"'] & H \\
S \arrow[u, "\xi^*\mu"] \arrow[ur, "\rho"'] &
\end{tikzcd} \tag{26}$$

the composition $q \circ \xi^*\mu$ generates the ρ function described in section 1.2.

1.3.1 Channels

Each ν function maps a field τ_c to a paramterized visual field μ, for example the color or shape visual channels [1, 6]. The set of is a functor from data to visual space such that any ν is an equivariant map if

$$\begin{tikzcd}
E_1 \arrow[r, "F"] \arrow[d] & V_1 \arrow[d, "f"] \\
E_2 \arrow[r, "F"] & V_2
\end{tikzcd} \tag{27}$$

Define a channel in terms of nus This is what we mean by equivariance: step thrpoughh measurment scale groups doesn't matter where computation lives

Here is the thing you need to preserve for X when writing a transform:

compose ν_{param} visual characteristic mappers

$$\nu(\tau) = \mu \tag{28}$$

where τ_{param} is a field in τ bound to the paremeter and μ_param is an intermediate visual representation such that $\nu_shape(\tau_shape)$ returns a μ_shape function that returns a shape for each k.

ν is Hom set? 1) must have identity 2) is composable 3) associativity

homomorphism in land of categories: objects in C, objects in D C arrow D A-B for example stephens group scale

every ν is a functor that

put all the functorial things together is still functorial

1.3.2 Marks

Bertin describes a location on the plane as the signifying characteristic of a point, measurable length as the signifying characteristic of a line, and measurable size as the signifying characteristic of an area and that in display (pixel) space are the point, line, and area marks [1, 4]. For each region s in the display space H, the mark it belongs to can be found by mapping s back to K via the lookup on S then taking $\xi(s)$ back to a point on $k \in K$ which lies on the connected component $J \subset K$.

$$H \underset{\rho(\xi^{-1}(J))}{\overset{}{\rightleftarrows}} S \underset{\xi^{-1}(J)}{\overset{\xi(s)}{\rightleftarrows}} J_k = \{j \in K | \exists \Gamma \text{ s.t. } \Gamma(0) = k \text{ and } \Gamma(1) = j\} \tag{29}$$

To got back to the display space H from the simplacial complex J of the signifier implanted in the mark, the inverse image of $J \in S, \xi^{-1}(J)$ is pushed back to S, and then $\rho(\xi^{-1}(J))$ maps it into R^7.

Can in theory approximate hatching/dashing/etc can be approximated w/ functions and neighborhood of k.

1.3.3 Visual Idioms: Equivalance class of artists

n $O(E)$ of the same type, they output the same type of prerender $O(H)$:

Natural transformation + composition is partial ordering? Back and forth is equivalent