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 $\Gamma(H)$:

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

- 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.
- H is a fiber bundle over S representing visual space
- K and S is a triangulizable topological space or CW complex encoding the connectivity of the points in E

All the fiber bundles mentioned in this work is assumed to be locally trivial [7, 16].

1.1 Data Fiber Bundles

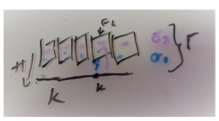


Figure 1: write up some words here

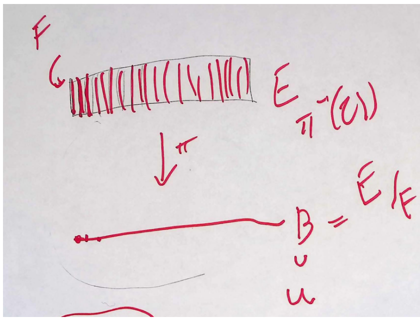
As proposed by Butler [3, 4], we model data as a fiber bundle (E, K, π, F) with $\pi : E \to K$ where K which can be thought of as a set of keys k. A section $\tau : K \to E$ is an instance of the data that lies in E and is discussed in section **??**.

We provide a brief description of fiber bundles because we model data, visual transformations, and a prerendered visual graphic as fiber bundles. Fiber bundles are product spaces of toplological spaces, which are a set of points with a set of neighborhoods for each point[7, 16]. A fiber bundle is a structure (E, B, π, F) consisting of topological total space E, base space K, fiber space F and the map from total space to base space:

$$F \hookrightarrow E \xrightarrow{\pi} B \tag{2}$$

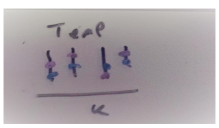
where there is a bijection from F to every fiber F_b over point $b \in B$ in E and the function $\pi : E \to B$ is the map into the B quotient space of E. By defintion of a fiber bundle, π is always a mapping from total space to base space, independent of the points $p \in E$, and therefore we call this mapping π for all the fiber bundles in the model.

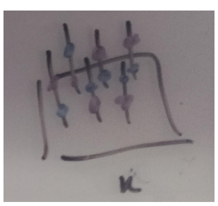
1.1.1 Base space



B is the quotient space of E, meaning it is the set of equivalence classes of elements p in E defined via the map $\pi : E \to B$ that sends each $p \in E$ to its equivalence class in $[p] \in B$ [14].

As shown in figure **??**, the fibers F divide E into smaller spaces consisisting of F and an open set neighborhood [7, 16] around F.





Example 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 [13].

1.1.2 Fiber

$$\begin{tikzcd} \pi^{-1}(U) \arrow[r, "\varphi"] \arrow[d, "\pi"'] & U \times F \arrow[dl, "\mathrm{proj}_U"] \\ U & \end{tikzcd} \tag{3}$$

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

$F_b = \pi^{-1}(b \in B)$ is homomorphic to the fiber F.

Example

1.1.3 Section

The section f is the mapping from base space to total space $f : B \to E$

$$\begin{tikzcd} F \arrow[r, hook] & E \arrow[d, "\pi"'] \\ & B \arrow[u, bend right, "f"'] \end{tikzcd} \tag{4}$$

such that f is the right inverse of π

$$\pi(f(b)) = b \text{ for all } b \in B \tag{5}$$

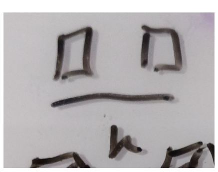
In a locally trivial fiber bundle, $E = B \times F$ [7, 16]:

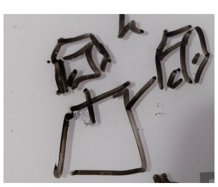
$$f(b) = (b, g(b)) \tag{6}$$

where the domain of $g(b)$ is F_b and returns a point p in F_b. The space of all possible sections f of E is $\Gamma(E)$. All sections $f \in \Gamma(E)$ have the same fibers F and connectivity B.

Example 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 $\tau : K \to \Gamma(\sigma)$. This is equivalent to $\tau : k \to E$ since $F = \Gamma(\sigma)$ and F is the embedding in E on which the records r lie.

1.1.4 Example





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 $C = \{\text{Temp}, \text{Time}\}$:

$$F = [temp_{min}, temp_{max}] \times [time_{min}, time_{max}] \tag{7}$$

and the function τ that retrieves records from F is

$$\begin{aligned} \tau(k) = (k, (r : \text{Temp} \rightarrow & temp, r : \text{Time} \rightarrow time)) && (8) \\ & temp \in [temp_{min}, temp_{max}], time \in [time_{min}, time_{max}]) && (9) \end{aligned}$$

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

1.1.5 Sheaf and Stalk

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

$$\begin{tikzcd} \iota^* E \arrow[d, "\pi"'] & E \arrow[l, "\iota^*"'] \arrow[d, "\pi"'] \\ U \arrow[u, bend right, "\iota^* f"'] & B \arrow[l, "\iota"] \arrow[u, bend right, "f"'] \end{tikzcd} \tag{10}$$

The localized section of fibers $\iota^* f : U \to \iota^* E$ is the sheaf $\mathcal{O}(E)$ with germ of $\xi^* f$ which is the localized section. The neighborhood of points the sheaf lies over is the stalk $\mathcal{F}_b$ [18, 19] which is a .... 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.

1.1.6 Triangulization

The base space K is a representation of the connectivity of the data, specifically whether the points in E are discrete or sampled from a continous space. The same dataset can be expressed with different K.

In our draft implementation of the data as fiber bundle model, we represent K as a simplacial complex.

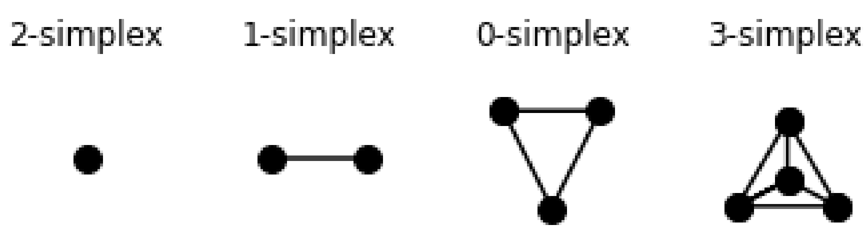


Figure 2: Simplices encode the connectivity of the data, from fully disconnected (0 simplex) observations to all observations are connected to at least 3 other observations. Higher order simplicies are outside the scope of this paper.

One way to represent the topological space K is as a set composed of simplices, such as those shown in figure **??**. Simplices are a way of encoding the connectivity of each observation ($\sigma(k)$) to another:

Example chopping up a torus maybe? talk about how that gets unpacked into triangles and then into vertices

1.2 Prerender Space

We model the prerender space on which lives on ideal version of the visulaization as a fiber bundle (H, S, π, D). H is the predisplay space, with a fiber D dependent on the target physical display and a base space of S.

1.2.1 Base space

K can be considered a subspace of the screen base space S such that $\xi : S \to K$ is a deformation retraction [15]

$$\begin{tikzcd} E \arrow[d, "\pi"'] & H \arrow[d, "\pi"'] \\ K & S \arrow[l, "\xi"'] \end{tikzcd} \tag{11}$$

that goes from a region $s \in S_k$ to its associated point k, such that when $\xi(s) = k$, $\xi^*\tau(s) = \tau(k)$.

1.2.2 Fiber and Section

A section $\rho : S \to H$ is a 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 visual space as defined in D.

Example For a physical screen display, we can consider a predisplay space that is a trivial fiber bundle $H = \mathbb{R}^5 \times S$ such that ρ is

$$\rho(s) = [x(s), y(s), r(s), g(s), b(s)] \tag{12}$$

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

$$s = \rho^{-1}_{XY}(xy) \tag{13}$$

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

$$r = \iint_s \rho_R(s) ds^2 \tag{14}$$

$$g = \iint_s \rho_G(s) ds^2 \tag{15}$$

$$b = \iint_s \rho_B(s) ds^2 \tag{16}$$

Here we assume a single opaque 2D image such that the z and $alpha$ fields can be omitted. To support overplotting and transparency, we consider $D = R^7$

1.2.3 Example

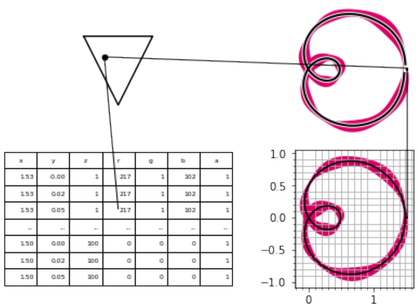


Figure 3

As illustrated in figure 3, words.

1.3 Artist

In this section we will define the artist as a mapping from a sheaf $\mathcal{O}(E)$ to $\mathcal{O}(H)$.

$$A : \mathcal{O}(E) \to \mathcal{O}(H) \tag{17}$$

The artist decomposes to mapping data to visual $\nu : E \to V$, then compositing V pulled back along ξ to ξ^*V to a visual mark in prerender space $Q : \xi^*V \to H$.

$$\begin{tikzcd}
E \arrow[r, "\nu"] \arrow[dr, "\pi"'] & V \arrow[d, "\pi"] & \xi^*V \arrow[l, "\xi^*"'] \arrow[r, "Q"] \arrow[d, "\xi^*\pi"'] & H \arrow[dl, "\pi"] \\
& K & S \arrow[l, "\xi"'] &
\end{tikzcd} \tag{18}$$

The visual fiber bundle $(V,\ K,\ \pi,\ P)$ has section $\mu : V \to K$ that resolves to a visual variable [2, 12] in fiber P. The visual transformer ν is a set of functions each targeting a different μ

$$\{\nu_0, \dots, \nu_n\} : \{\tau_0, \dots, \tau_n\} \mapsto \{\mu_0, \dots, \mu_n\} \tag{19}$$

where μ_i are the visual parameters in the assembly function $Q(\mu_0, \dots, \mu_n)(s) = \rho(s)$.

1.3.1 Example: Matplotlib Visual Fiber

For example, for Matplotlib [8], some of the possible types in P are:

1.3.2 Visual Channels

$\nu : E \to V$ is an equivariant map such that there is a homomorphism from left monoid actions on E_i to left monoid actions on V_i where i identifies a field in the fiber. E_i and V_i each contain a set of values as defined in F and P respectively. A validly constructed ν is

$\boldsymbol{\nu_i}$	$\boldsymbol{\mu_i}$	$\boldsymbol{codomain(\nu_i)}$
position	x, y, z, theta, r	$\mathbb{R}$
size	linewidth, markersize	$\mathbb{R}^+$
shape	markerstyle	$\{f_0, \ldots, f_n\}$
color	color, facecolor, markerfacecolor, edgecolor	$\mathbb{R}^4$
texture	hatch	$\mathbb{N}^{10}$
	linestyle	$\{f_0, \ldots, f_n\} \times (\mathbb{R}, \mathbb{R}^{+n, n\%2=0})$

one where the diagram

$$\begin{array}{ccc} E_i & \xrightarrow{\nu_i} & V_i \\ {\scriptstyle m_e}\downarrow & & \downarrow{\scriptstyle m_v} \\ E_i & \xrightarrow{\nu_i} & V_i \end{array} \tag{20}$$

commutes such that $\nu_i(m_e(E_i)) = m_v(\nu_i(E_i))$.

Definition A monoid[11] M is a set that is closed under an associative binary operator $*$ and has an identity element $e \in M$ such that $e * a = a * e = a$ for all $a \in M$. A left monoid action [1, 17] of M is a set X with an action $\bullet$ with the properties:

closure $\bullet : M \times X \to X$,

associativity for all $m, t \in M$ and $x \in X, m \bullet (t \bullet x) = (m \bullet t) \bullet x$

identity for all $x \in X, e \in M, e \bullet x = x$

Example: Partial Order To preserve ordering of elements in E_i, ν must be a monotonic function such that given $e_1, e_2 \in E_i$

$$\text{if } e_1 \leq e_2 \text{ then } \nu(e_1) \leq \nu(e_2) \tag{21}$$

Example: Translation **fairly certain I lost the thread here** According to Stevens, interval data is a set with general linear group actions [10, 20]. Position is a visual variable that can support translation [2, 9, 12].

$$\nu(x + c) = \nu(x) + \nu(c) \tag{22}$$

Example: Invalid ν Given a transform $t(x) = x + 2$, we construct a ν that always takes data to .5:

$$\begin{array}{ccc} E_1 & \xrightarrow{\lambda : e \mapsto .5} & V_i \\ {\scriptstyle 2e}\downarrow & & \downarrow{\scriptstyle 2v} \\ E_1 & \xrightarrow{\lambda} & V_1 \end{array} \tag{23}$$

This ν is invalid because the graph does not commute for t:

$$\nu(t(e)) \stackrel{?}{=} t(\nu(e)) \tag{24}$$

$$.5 \stackrel{?}{=} t(.5) \tag{25}$$

$$.5 \neq 2 * .5 \tag{26}$$

To consutruct a valid ν, the diagram must commute for all monoid actions on the sets E_i, V_i.

1.3.3 Assembling Marks

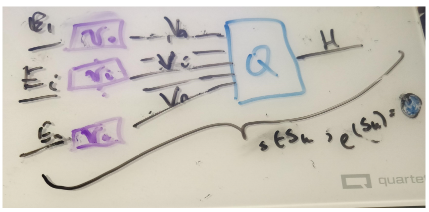


Figure 4: The ν functions convert data E to visual V. Q assembles the different types of visual parameters V_i into a graphic in H. $Q \circ \mu(\xi^{-1} J)$ forms a visual mark by applying Q to a region mapped to connected components $J \subset K$.

The assembly function $Q : \Gamma(V) \rightarrow \Gamma(H)$ composites the visual variables in V into an element in H. Given a monoid action on a set in V, there should be a monoid action on the corresponding $Q(\Gamma(V))$. While $\Gamma(V)$ holds for all cases, we can specialize to the bundle V or the sheaf (V) depending on the specific visualization.

Proposition If $\forall g \in M$ and $\forall x_1, x_2 \in \Gamma(V)$ then $Q(x_1) = Q(x_2)$ $implies Q(g \circ x_1) = Q(g \circ x_2)$; therefore we can define a group action on $Q(\Gamma(V)) = Y$ as $g \circ y = y'$ where $y' = Q(g \circ x)$ with $x \in f^{-1}(y)$

For each region s in the display space H, the mark [2, 5] 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 path connected component $J \subset K$.

$$H \underset{\rho(\xi^{-1}(J))}{\overset{}{\rightleftarrows}} S \underset{\xi^{-1}(J)}{\overset{\xi(s)}{\rightleftarrows}} J_k = \{j \in K \text{ s. t. } \exists \gamma \text{ s.t. } \gamma(0) = k \text{ and } \gamma(1) = j\} \tag{27}$$

where the path[6] γ from k to j is a continous function from the interval [0,1]. To get back to the display space H from J, 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 such that $Q \circ \xi^* \mu(\xi^{-1}(J))$ generates a graphical mark.

1.3.4 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