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 $\Gamma(H)$:

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

- 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.
- K is a triangulizable space encoding the connectivity of the points in E
- H is a fiber bundle over S representing visual space
- S is a simplacial complex encoding the visualization

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[4, 10].

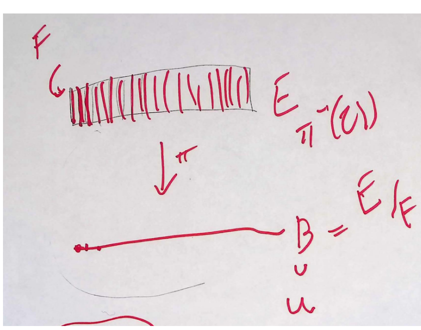
1.1 Fiber Bundles

We provide a brief description of fiber bundles because we model data, visual transformations, and a prerendered visual graphic as fiber bundles. 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 $\pi : E \rightarrow B$ 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 $[p] \in B$ [8].

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 $\mathcal{T}$

$$\mathcal{T}_b = \{U \subseteqq B : \{p \in E : [p] \in U\} \in \mathcal{T}_E\} \tag{3}$$

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

1.1.2 Fiber

As shown in equation!**??**, every point $b \in B$ has a local open set neighborhood U [4, 10]

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

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

1.1.3 Section

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

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

such that f is the right inverse of π

$$\pi(f(b)) = b \text{ for all } b \in B \tag{6}$$

In a locally trivial fiber bundle, $E = B \times F$ [4, 10]:

$$f(b) = (b, g(b)) \tag{7}$$

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.

1.1.4 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.

$$\begin{tikzcd} \iota^* E \arrow[d, "\pi"'] & E \arrow[l, "\iota^*"'] \arrow[d, "\pi"'] \\ U \arrow[u, bend right, "\iota^* f"'] & B \arrow[l, "\iota"'] \arrow[u, bend right, "f"'] \end{tikzcd} \tag{8}$$

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

$$\iota^{-1}\mathcal{F}(\{b\}) = \varinjlim_{b \subseteq U} \mathcal{F}(U) = \varinjlim_{b \in U} = \mathcal{F}_b \tag{9}$$

which through ι gets the data in E at and near to b. 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 Data Model

As proposed by Butler [2, 3], 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 1.2.4.

1.2.1 Example

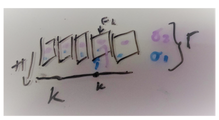


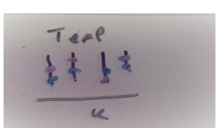
Figure 1: write up some words here

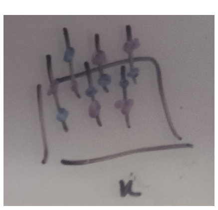
As illustrated by figure **??**, the vertical lines F are the range of possible temperature values embedded in the total space E. The base space K of the fiber bundle is a line because the data points r in E are on a space that is continous in one dimension.

1.2.2 Base Space

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.

1.2.3 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 [7].

1.2.4 Fiber and Sections

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

$$\pi : U \to DT \tag{10}$$

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

$$\sigma : C \to DT \tag{11}$$

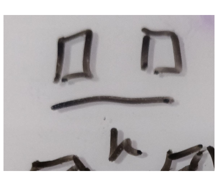
where C is the finite set of names of columuns, which we generalize to data fields in E. The set of all values restricted to the datatypes in DT is U_σ

$$\begin{tikzcd} U_\sigma \arrow[r] \arrow[d, "\pi_\sigma"'] & U \arrow[d, "\pi"] \\ C \arrow[r, "\sigma"] & DT \end{tikzcd} \tag{12}$$

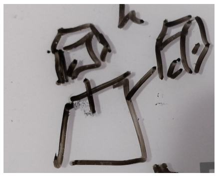
The pullback $U_\sigma := \sigma^{-1}(U)$ restricts U to the datatypes of the fields in C such that U_σ is the fiber product $U \times_{DT} C$, and the pullback $\pi_\sigma : U_\sigma \to C$ specifies the domain bundle U_σ over C induced by σ. The fiber F is the cartesian product of all sets in the disjoint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

1.2.5 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 $C = \{\mathrm{Temp}, \mathrm{Time}\}$:



$$F = [temp_{min}, temp_{max}] \times [time_{min}, time_{max}] \tag{13}$$

and the function τ that retrieves records from F is

$$\tau(k) = (k, (r : \text{Temp} \rightarrow temp, r : \text{Time} \rightarrow time)) \tag{14}$$

$$temp \in [temp_{min}, temp_{max}], time \in [time_{min}, time_{max}]) \tag{15}$$

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

1.3 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.

1.3.1 Base space

K can be considered a subspace of the screen base space S such that $\xi : S \rightarrow K$ is a deformation retraction [9]

$$\begin{tikzcd} E \arrow[d, "\pi"'] & H \arrow[d, "\pi"'] \\ K & S \arrow[l, "\xi"'] \end{tikzcd} \tag{16}$$

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

1.3.2 Fiber and Section

A section $\rho : S \rightarrow H$ 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. For a physical screen display, the predisplay space is a trivial fiber bundle $H = \mathbb{R}^7 \times S$ such that ρ is

$$\rho(s) = \{x, y, z, r, g, b, a\} \tag{17}$$

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

$$s = \rho^{-1}_{XY}(xy) \tag{18}$$

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

$$r = \iint_s \rho_R(s) ds^2 \tag{19}$$

$$g = \iint_s \rho_G(s) ds^2 \tag{20}$$

$$b = \iint_s \rho_B(s) ds^2 \tag{21}$$

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

1.3.3 Example

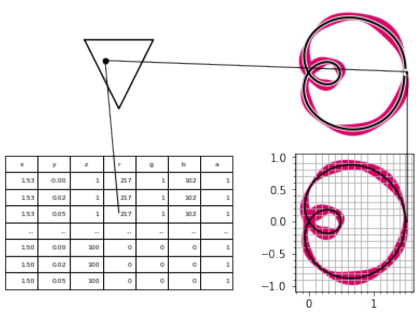


Figure 2

As illustrated in figure 2, words.

1.4 Artist

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

$$A : \mathcal{O}(E) \to \mathcal{O}(H) \tag{22}$$

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 $Q : \xi^*V \to H$.

$$\begin{tikzcd} E \arrow[r, "\nu"] \arrow[dr, "\pi"'] & V \arrow[d, "\pi"] & \xi^*V \arrow[r, "Q"] \arrow[d, "\xi^*\pi"'] & H \arrow[dl, "\pi"] \\ & K & S \arrow[l, "\xi"] & \end{tikzcd} \tag{23}$$

The visual fiber bundle $(V,\ K,\ \pi,\ P)$ has section $\mu : V \to K$ that resolves to a visual variable [1, 6] in fiber P. The visual transformer ν is a set of functions each targeting a different μ

$$\nu : \{\nu_0, \ldots, \nu_n\} \to \{\mu_0, \ldots, \mu_n\} \tag{24}$$

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

1.4.1 Example: Matplotlib Visual Fiber

For example, for Matplotlib [5], some of the possible types in P are:

ν_i	μ_i	$dom(\nu_i)$
position	x, y, z	$\mathbb{R}$
size	linewidth, markersize	$\mathbb{R}$
shape	markerstyle	$((\mathbb{R}, \mathbb{R}), x \in \{0, 1, 2, 3, 4, 79\})^{n\dagger}$
color	color, colors, linecolor, facecolor, markerfacecolor, edgecolor	$\mathbb{R}^4$
texture	hatch	$/^* \backslash^* -^* +^* x^* o^* O^* .^* *^* \in \Sigma^+$ $\Sigma = \{/, \backslash, \vert, -, +, x, o, O, ., *\}^\ddagger$
	linestyle	$(\mathbb{Z}, (\mathbb{Z}^n))^\dagger$

Table 1:
$\dagger n > 0$
$\ddagger \Sigma^+ = \cup_{n \in \mathbb{Z}^+} \Sigma^n$

1.4.2 Visual Channels

The function $\nu : E \to V$ is a mapping from the data bundle to the visual bundle

$$\begin{tikzcd} E_1 \arrow[r, "\nu"] \arrow[d, "f"'] & V_1 \arrow[d, "f"] \\ E_2 \arrow[r, "\nu"] & V_2 \end{tikzcd} \tag{25}$$

such that a visual section $\mu : K \to V$ is equivalent to the visual transformation of the data $\mu(k) = \nu \circ \tau(k)$

+2 on data side means +2 on visual side

.2 = E1 - lambda e -¿ .2 V1 = .2 +.2 +.2 .4 = E2 - lambda e -¿ .2 V2 = .2

.2 = E1 - lambda e-¿ .2 V1 = .2

V2 = .2

1 = E1 - lambda e-¿e/5 = .25

+ 1 +.25

2 = E2 - lambda e-¿e/5 = .5

preserve unique mapping? does ν need to be invertable?

shifts in either space need to be mirrored in the other space, this is not valid here's why: $\lambda e : .2$

$$\begin{tikzcd} E_1 \arrow[r, "\lambda : e \mapsto .5"] \arrow[d, "e+2"'] & V_1 \arrow[d, "h+2"] \\ E_2 \arrow[r, "\lambda"] & V_2 \end{tikzcd} \tag{26}$$

x,y = 1,2 y,x = 2,1 y,x = 1,2
The constraint on $\nu : \tau \to \mu$

Allowable Maps For example, stevens: preserve structure .... Some examples of allowable maps are described in Bertin/Munzner: nus are constrained to preserve..., expect functorial nus - transform on input side to transform on output side (E1-E1, V1-V1, nu) ¡- condition that's how we're constraining nu. functorial in ν component

1.4.3 Constructing Marks

The visual fiber bundle V gets pulled back over S via ξ such that

$$\begin{tikzcd} \xi^*V \arrow[r, "Q"] \arrow[dr, "\xi^*\pi"'] & H \arrow[d, "\pi"] \\ & S \end{tikzcd} \tag{27}$$

the function $Q : \xi^*V \to H$ composites points in V into $\mathbb{R}^{\ntriangleright}$ tuples. The section μ is pulled over s

$$\begin{tikzcd} \xi^*V \arrow[r, "Q"] & H \\ & S \arrow[u, "\rho"'] \arrow[ul, "\xi^*\mu"] \end{tikzcd} \tag{28}$$

such that the composition Q of μ is equivalent to the render $\rho(s) = Q \circ \mu(s)$. (In practice $\mu \approxeq \xi^*\mu$) Q is the constructor for the mark, one Q per artist, A = sum of artists Q as there's default line, point, etc that's stretched and tweaked and whatever else by the μ visual variable + structure preserving, Q must have map that commutes such that a change in μ carries over somehow to ρ, functorial in its arguments - commutative diagram - into rho