OPTIMAL TRANSPORTATION, DOLD-THOM'S THEOREM, AND SWEEPOUTS

J. H. MARTEL

This article aims to develop connections between Optimal Transportation of Monge-Kantorovich and Algebraic Topology. Such applications were developed in our thesis [ref], wherein we studied the possibility of representing souls/spines of spaces as the singularities of certain optimal semicoupling programs. Here we return to some foundational issues in Algebraic Topology, and especially Dold-Thom's theorem on representing homology cycles and sweepouts. In the author's experience, these connections are invaluable computational tools leading to constructive applications of algebraic topology.

1. Brouwer and Kantorovich

Around 1910 A.D. the foundations of algebraic topology were rigorously established by L.E.J. Brouwer in a series of articles using the singular homology functors. His fundamental theorems include the following:

(A) *Brouwer's Invariant Domain Theorem*: If U is an open subset of $\mathbb{R}^n$, and $f : U \to \mathbb{R}^n$ is a continuous injective map, then f is an *open* map and the image $f(U)$ is an open subset of $\mathbb{R}^n$.

(B) *Brouwer's Fixed Point Theorem*: If $f : B^n \to B^n$ is a continuous self-mapping of the n-dimensional unit ball B, then f admits at least one fixed point, i.e. there exists $x \in B$ such that $f(x) = x$.

(C) *Brouwer's No Retract Theorem*: Let X be a oriented manifold with boundary ∂X. Then there does not exist continuous deformation retracts $X \rightsquigarrow \partial X$. In particular, there does not exist any continuous deformation retract from the closed ball B to its boundary sphere $S = \partial B$.

In later years Brouwer came to reject his earlier work, and there developed a controversy between Brouwer's so-called Intuitionism and Hilbert's Gottingen Formalism. Which stands today?

2. Configuration Spaces and Dold-Thom

2.1. **Basepoints.** The purpose of this article is to describe the greatest theorem you've never heard of, namely Dold-Thom's Theorem. In principle, there are several

Date: January 7, 2022.

variations on DT's original theorem [ref]. The variations arise in the different types of configuration spaces which can be defined on a given space. We remark that Almgren's thesis [ref] contains an important generalization of DT's original theorem, and is the basis for the formal definitions of "sweepout" in geometric topology, c.f. [Guth-Width Inequality]. Further variants of DT's theorem can be found in [Segal, McDuff, ...]. However, from the author's perspective, all these variations suffer a severe defect, namely they are all *basepoint dependant*. This article was motivated by the author's desiring a basepoint-free version of DT's theorem for which the long exact sequence of relative homology [ref] is naturally isomorphic to the long exact sequence in homotopy [ref]. Indeed, reduced singular homology $\tilde{H}_*$ is basepoint independant. However fundamental groups and homotopy groups are intrinsically basepoint dependant. We remark that the choice of basepoint on a space X is as complicated as Poincaré's fundamental group $\pi_1(X, pt)$. The fundamental fact of analysis situs is that canonical basepoints do not exist when π_1 is nontrivial.

2.2. **Brief Survey of Configuration Spaces.** The basic models are:

(1) Hard disks (where collisions of point particles is prohibited);

(2) Hard disks with basepoint (where collisions are prohibited except with an annihilating basepoint);

(3) Water droplets (c.f. Gromov's category of finite probability spaces [ref], where collisions are encouraged and additive);

(4) Electrostatic droplets (e.g. signed Borel-Radon measures on X, possibly with indivisible quanta, e.g. Millikan's electron).

In example (1) the configuration space forms a monoid without identity, and the disks are not allowed to collide or intersect or join, just as hard disks cannot be joined together.

In example (2), a basepoint pt can be chosen on X and this basepoint serves as zero element, thus rendering the monoid into an additive topological group. However, as we commented above, a choice of basepoint is noncanonical when $\pi_1 \neq 0$.

The formal construction of earlier configuration spaces in the literature involves symmetric products with the diagonal removed, i.e. a limit of $X^N - \Delta$ as $N \to +\infty$. The conventional presentations of Dold-Thom's theorem, including the original [DT58], view configuration spaces in the hard-disk model, and arbitrary basepoint on X must be selected. This basepoint excludes the possibility of relating the conventional Dold-Thom theorem to relative homology.

Firstly, we remark that braid groups [ref] are indeed the fundamental groups of configuration spaces of hard disks on a background disk. [ref: Birman?]

Example (3) arises from Gromov's category of finite probability spaces [ref], where the objects are so-called reductions $f : \mu \to \nu$ between finite probability spaces μ, ν.

The category of (4) consists of finite electroneutral configurations, where the objects are again reductions $f : \mu \to \nu$. The proper formalization of (4) into a topological abelian group is as follows.

Definition 1. Let X be a topological space. Let $\mathbb{Z}(X)$ be the group of finitely-supported $\mathbb{Z}$-valued distributions on X, i.e.

$$\mathbb{Z}(X) := \{\sum n_x.x \mid \text{only finitely many nonzero } n_x \in \mathbb{Z}\}.$$

Let $\epsilon_X : \mathbb{Z}(X) \to \mathbb{Z}$, $\epsilon_X(\sum n_x x) = \sum n_x$ be the canonical augmentation map. Let $AG_0(X) := \ker(\epsilon_X)$ be the kernel of the augmentation map.

We make some remarks:

If G is an abelian group, we define $G(X)$ and $AG_0(X; G)$ in the obvious way. However the cases of $G = \mathbb{Z}$ and $G = \mathbb{Z}/2\mathbb{Z}$ appear to be cases of most interest, and henceforth we suppress the 'G' from notation.

Morever because $\mathbb{Z}$ is discrete the point charges (+), (-) are not infinitely divisible. This fails to be true if we replace $\mathbb{Z}$-valued distributions with $\mathbb{R}$-valued distributions, i.e. there exists unit charge quanta if and only G is discrete.

We observe that $\mathbb{Z}(X)$ is a topological abelian group, with zero element $\emptyset := \sum 0x = 0$ corresponding to the zero distribution, and whose zero element is essentially the zero element 0 of $\mathbb{Z}$. One might consider $\emptyset$ the "vacuum state on X". It follows that $AG_0(X)$ is a topological subgroup, consisting of all distributions with zero net charge. One naturally sees AG_0 as a discrete version of all signed Borel-Radon measures μ on X which integrate to zero $\int_X 1.d\mu(x) = 0$, i.e. which are orthogonal to all constant functions on X.

The essence of DT theorems is to represent homology groups in terms of homotopy groups on configuration spaces. In otherwords to identify the homology functors as homotopy functors. The basic idea of this article is that the vacuum state $\emptyset$ serves as a type of "canonical basepoint on X", and this enables a natural equivalence between the long exact sequences of relative homology and homotopy.

It's important to establish the relative version of DT. Let Y be a closed subset of X. Our goal is to define a relative configuration space $AG_0(X/Y)$, and the key identification will be the canonical isomorphism

$$AG_0(X/Y) = AG_0(X)/AG_0(Y).$$

In terms of net zero charged particle configurations, the idea is to view Y as a *reservoir* where excess charges can "ground out". That is, spheres in the relative DT group $AG_(X/Y)$ basepointed at $\emptyset$ are recombinations which either neutralize away from Y, or neutralize at Y.

Here is the formal definition. If Y is a closed subset of X, then $\mathbb{Z}(Y)$ and $AG_0(Y)$ is a closed subgroup of $\mathbb{Z}(X)$ and there is a canonical quotient $\mathbb{Z}(X) \to \mathbb{Z}(X)/AG_0(Y)$.

Morever the augmentation map ϵ_X canonically descends to the quotient as a type of augmentation map

$$\epsilon_{X/Y} : \mathbb{Z}(X)/AG_0(Y) \to \mathbb{Z},$$

and we identify

$$AG_0(X/Y) := \ker(\epsilon_{X/Y}).$$

We follow the original approach of [DT58], which theorems are stated in terms of natural quasifibrations between the configuration spaces. Recall the definition of quasifibration:

Definition 2. A continuous map $f : X \to Y$ between topological spaces X, Y is a *quasifibration* if the canonical inclusion of fibres $f^{-1}(y)$ into the homotopy fibre of f is a weak homotopy equivalence for every $y \in Y$.

Dold-Thom Theorem - Relative Version 3. Let Y be closed subspace of X. Then the short exact sequence of topological abelian groups

$$0 \to AG_0(Y) \to AG_0(X) \to AG_0(X/Y) \to 0,$$

is a quasifibration inducing a long exact sequence of $\emptyset$-pointed homotopy groups

$$\cdots \to \pi_{*+1}(AG_0(Y), \emptyset) \to \pi_{*+1}(AG_0(X), \emptyset) \to \pi_{*+1}(AG_0(X/Y), \emptyset) \to \pi_*(AG_0(Y), \emptyset) \to \cdots$$

which is naturally equivalent to the long-exact sequence of relative homology groups

$$\cdots \to \tilde{H}_{*+1}(Y) \to \tilde{H}_{*+1}(X) \to H_{*+1}(X, Y) \to \tilde{H}_*(Y) \to \cdots .$$

Proof of (3). The proof of Dold–Thom applies. One needs verify that the functors $X \mapsto \pi_* AG_0(X)$ satisfy the Eilenberg–Steenrod axioms, from which it follows that the functors $\pi_* AG_0$ and $\tilde{H}_*$ are naturally equivalent. □

We make some remarks:

In terms of category theory, DT says that if G is an abelian group, then the functor $X \mapsto \pi_*(AG_0(X; G), \emptyset)$ is naturally equivalent to the reduced singular homology functor $X \mapsto \tilde{H}_*(X; G)$ in the category TOP of basepoint-free topological spaces. Again, we emphasize that (3) is basepoint independant, with the vacuum state $\emptyset$ serving as "canonical basepoint" on X.

Formally the homotopy groups $\pi_q(AG_0(X), \emptyset)$ consist of homotopy classes of pointed continuous maps $f : (S^q, pt) \to (AG_0(X), \emptyset)$. Identifying pt with the point at-infinity, we can thus model homotopy classes as compactly supported q-parameter family of distributions on $\mathbb{R}^q$, where "compactly supported" means the distribution is equal to vacuum state outside a compact subset.

The topology on $AG_0(X)$ implies recombinations of $\emptyset$ satisfy local conservation of charge [ref: Feynman]. Thus point charges (+) and (-) must be born from the same spatial position on X. Therefore the recombinations require the point charges to continuously move on X, and without "teleportation". This is analogous to the distinction between the standard L^1 and L^2 optimal transport.

One immediate consequence of DT is that we find new topological models for the $K(G,n)$ spaces. For example, we find $K(\mathbb{Z},2)$ is modelled by $AG_0(S^2)$, as opposed to the usual $\mathbb{C}P^\infty$ model. Identifying S^2 with the Riemann sphere $\mathbb{C}P^1$, we can identify $AG_0(S^2)$ with the set of rational functions on $\mathbb{C}P^1$, i.e. the space of meromorphic functions on the Riemann sphere. The correspondance from meromorphic functions to $AG_0(S^2)$ is given by assigning to every meromorphic function f it's divisors, i.e. poles and zeros of f. More generally we find $K(G,n)$ modelled by $AG_0(S^n;G)$ for every integer $n \geq 1$ and abelian group G.

As consequence of (3) we have following

Lemma 4. *If Y is closed subset of X, then we have canonical isomorphism*

$$H_*(X,Y)/image(H_*(X)) = \ker(H_{*-1}(Y) \to H_{*-1}(X)),$$

where $image(H_*(X))$ is the image of $H_*(X)$ in the relative homology group $H_*(X,Y)$, and the morphism $H_{*-1}(Y) \to H_{*-1}(X)$ is induced by the inclusion $Y \hookrightarrow X$.

Proof. Long exactness in (3) implies

$$image(H_*(X)) = \ker(\delta)$$

and

$$image(\delta) = \ker(H_{*-1}(Y) \to H_{*-1}(X)).$$

But we have canonical isomorphism

$$H_*(X,Y)/\ker(\delta) = image(\delta),$$

and the result follows. □

While the isomorphism $H_*(X,Y) = H_{*-1}(Y)$ is canonical between homology groups, the morphisms are noncanonical on the level of singular chains. However depending on the relative geometry of X, Y, there are circumstances where the correspondance can be made between singular chains.

Our interpretation of (4) is that relative cycles of X (mod Y) are represented either by chains w which are already cycles $\partial w = 0$ on X, or chains w whose chain boundaries ∂w are nontrivial on Y but null homologous in X. In terms of DT, this says that if Y is a ground reservoir for charges on X, then recombinations of $\emptyset$ will have parts which are recombinations in X separated from Y and other parts representing those charges grounding out at Y. For example, any excess charge on a conductive plate will either ground out at the boundary, or recombine and neutralize with some local charge. For the relative homology groups, we can identify a component of $AG_0(X/Y)$ which is represented by "spontaneous" transports from $\emptyset$

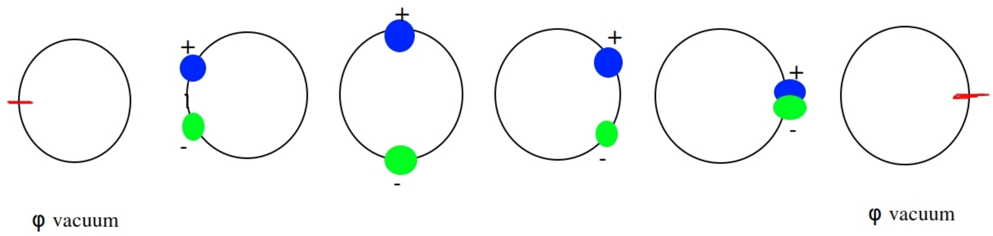

FIGURE 1. Basic 1-cycle on the one-dimensional circle S^1

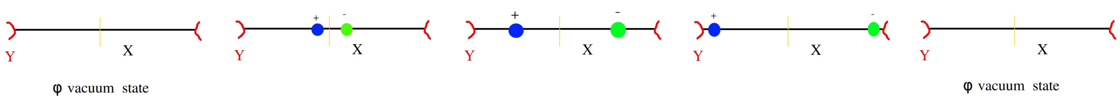

FIGURE 2. Nontrivial relative 1-cycle on $X = [0,1]$ with boundary $Y = \partial X = \{0,1\}$

to the reservoir Y, see (2.2) for basic example on one-dimensional disk with boundary reservoir.

3.

Now we begin to develop the connections between Dold-Thom's theorem (DT) and optimal transportation (OT). Let (X, d, σ) be a source space, say an oriented complete compact Riemannian manifold with $\sigma = vol_X$. The topologist faces several questions:

> *How to generate nontrivial homological k-cycles in X? How to generate nontrivial k-dimensional sweepouts of X?*

These questions have long history in topology, and we review the classical argument developed by Steenrod, Eilenberg-Maclane as described in [RobKirby]. If G is an abelian group, then Eilenberg-Maclane defined $K(G,n)$ to be those topological spaces satisfying $\pi_*(K(G,n),pt) = G$ for $* = n$ and $\pi_*(K(G,n),pt) = 0$ otherwise. Moreover if X is any topological space, then there exists natural equivalence between the set of homotopy classes $[X, K(G,n)]$ of continuous maps $f : X \to K(G,n)$ and the singular cohomology groups $H^n(X;G)$. The Hurewicz theorems [ref] imply the singular cohomology of $K(G,n)$ is concentrated in dimension n with a single spherical generator ι. If $f : X \to K(G, n)$ is a smooth and transverse curve, and homotopically nontrivial, then $f*(\iota) \neq 0$ is a nontrivial cohomology class in $H^n(X)$. Moreover if $y \in K(G, n)$ is a regular value of f, then $f^{-1}(y)$ represents a homologically nontrivial $(N-n)$-cycle Poincaré dual to $f^*(\iota)$, where $N = \dim(X)$. Thus we can use the points $y \in Y \approx K(G, n)$ to sweepout cycles on X via the fibres $f^{-1}(y) \hookleftarrow X$.

In practice, the transversality hypothesis is generic but yet also very difficult to obtain. Geometry is not generic in the smooth sense. Moreover the models $K(G, n)$ are not generally well known (although compare our earlier remarks). Thom[ref] reasoned that only the lower skeleton of any $K(G, n)$ model is necessary, depending on the cohomological dimension of X.

For example, constructing 1-cycles on a two dimensional surface S requires constructing nontrivial smooth maps $f : S \to AG_0(S^1)$. Evenmore as we vary the fibres $f^{-1}(y)$, over all the values $y \in image(f)$, we get a sweepout of S into *disjoint* cycles, with possibly some blowup at the critical points in S. Our goal is to generate smooth maps from the regularity theory of optimal transport.

Our thesis [Mar] introduced a method for applying optimal transportation to algebraic topology. The idea is that singularities of optimal transports are well-defined topological subvarieties whose geometry is described by the geometry of the cost. Given a source (X, σ), we introduce target spaces (Y, τ) and costs $c : X \times Y \to \mathbb{R}$. When the cost c satisfies certain assumptions, we find optimal transport solutions and the solutions to Kantorovich's dual program are unique and have sufficiently regular topology. In the dual program, the key definition is the c-subdifferential $\partial^c\psi(y) \hookrightarrow X$. Using the c-subdifferential we define *Kantorovich's contravariant functor* $Z : 2^Y \to 2^X$ by $Z(Y_I) = \cap_{y \in Y_I} \partial^c\psi(y)$ for every closed subset $Y_I \hookrightarrow Y$. We view Z as defining a (Y, τ)-parameter family of cells $\partial^c\psi(y)$ in X. The cells need not be disjoint. There are standard assumptions on the cost c such that almost every cell $Z(y) = \partial^c\psi(y)$ is an $(n-k)$-dimensional submanifold of X, and otherwise singular subvarieties. Our main idea is that every pair of target (Y, τ) and cost c determines a unique Y-parameter family of subsets $Z(y) = \partial^c\psi(y)$ of X.

These definitions lead to the following problems:

- Find general conditions on Y, τ, c such that c-optimal transports π generate continuous maps $f = f_\pi : X \to AG_0(Y)$, and such that c-optimal transports π general nontrivial sweepouts of X;
- Find general conditions such that the functor Z is continuous in the Almgren-Guth topology of "flat chains". [See Guth, S.1; Gromov text]. Explicitly, this requires proving that if y_0, y_1 are sufficiently close in Y, then $Z(y_0)$ and $Z(y_1)$ bound a chain C of arbitrarily small area with $\partial C = Z(y_0) - Z(y_1)$.

For the purposes of generating sweepouts, the geometric homology spheres S^d form a natural class of target spaces. But given a source X, can we readily generate costs $c : X \times S^d \to \mathbb{R}$? Already the case of constructing interesting costs $S \times S^1 \to \mathbb{R}$, or better yet $S \times AG_0(S^1) \to \mathbb{R}$, appears nontrivial problem. Here S denotes a compact oriented surface.

4. Regularity of Optimal Transports

The regularity of optimal transportation is well-developed thanks to the work of McCann, et. al.. Typically the cost assumptions are symmetric in the source and target variables. However in our applications, we specifically distinguish between source points (X, σ) and target points (Y, τ). The key assumptions require the cost c to be C^2 on $X \times Y$, proper, and satisfying a (Twist) condition. The assumptions imply every c-concave function $\psi : Y \to \mathbb{R} \cup \{+\infty\}$ is Lipschitz and semiconcave, and differentiable τ-almost everywhere on Y. Furthermore, $D^2\psi$ is defined τ-almost everywhere on $dom(D\psi)$. [ref]

We follow McCann-Pass and define the following two subsets:

$$X_1(y) := \{x \in X \mid \nabla_y c(x, y) = \nabla_y \psi(y)\},$$

and

$$X_2(y) := \{x \in X_1(y) \mid \nabla^2_{yy} c(x, y) - \nabla^2_{yy}\psi(y) \geq 0\}.$$

By definition we see

$$\partial^c\psi(y) \subset X_2(y) \subset X_1(y).$$

The definition of c-concavity implies $x \in \partial^c\psi(y)$ if and only if equality is attained in $c(x, y) \leq -\phi(x) + \psi(y)$, where $\phi = \psi^c$. But then $x \in X_1$ if $y \in domD\psi$, and $x \in X_2(y)$ if $y \in domD^2\psi$. Moreover X_2 is a closed subset of X_1. If $D^2_{xy}c(x, y)$ is nonsingular for every $x \in X_1(y)$, then the Inverse Function Theorem implies $X_1(y)$ is a smooth submanifold of X.

Example. We now illustrate following the approach of [McCann-Pass]. Transporting (X, σ) to a one-dimensional target measure $(\mathbb{R}, \tau)$. Suppose $c : X \times \mathbb{R}) \to \mathbb{R}$ is a cost. If $\nabla^2_{xy}c(\cdot, y)$ is nonvanishing, i.e. $\frac{\partial c}{\partial y}(\cdot, y)$ has only regular values, then all the fibres of $x \mapsto \frac{\partial c}{\partial y}(\cdot, y)$ define smooth hypersurfaces in X. The c-optimal transport map from σ to τ then has the following form. For given $y \in \mathbb{R}$, define $k = k(y) \in \mathbb{R}$ such that

$$\int_{\{x \in X \mid \frac{\partial c}{\partial y}(x,y) \leq k\}} d\sigma(x) = \int_{-\infty}^{k} d\tau(y).$$

The c-optimal transport map $F : (X, \sigma) \to (Y, \tau)$ is then defined for a given $\bar{x} \in X$ by $F(\bar{x}) = \bar{y}$, where $\bar{y}$ is the unique element satisfying

$$\sigma[\{x \in X \mid \frac{\partial c}{\partial y}(x, \bar{y}) \leq \frac{\partial c}{\partial y}(\bar{x}, \bar{y})\}] = \tau[(-\infty, \frac{\partial c}{\partial y}(\bar{x}, \bar{y}))].$$

5. Comparison with Width Inequalities

Gromov-Guth Width Inequalities: Given X, let $\{z_t\}$ be a sweepout of X into k-dimensional cycles. We require that the sweepout is continuous with respect to Almgren's flat chain topology. [Insert defn] Consider the maximum volume $\max_t vol_k(z_t)$ of the cycles in $\{z_t\}$, and let

$$width_k(X) := \min_{\{z_t\}} \max_t vol_k(z_t),$$

where the minimum ranges over all k-sweepouts of X. Estimates on $width_k$ imply every k-sweepout contains at least one cycle of "large" volume. Now our task is to interpret $width_k$ in terms of optimal transportation.

On a k-dimensional target Y, consider the min-max quantity

$$\alpha_Y := \min_g \max_y \frac{g(y)}{\int_Y g},$$

where the minimum ranges over all positive continuous functions $g : Y \to \mathbb{R}_{\geq 0}$, and the maximum ranges over $y \in Y$. One readily finds the minimum is attained for the uniform distribution on Y, in which case $\alpha_Y = (vol_k(Y))^{-1}$.

Suppose c satisfies the assumptions [A]. Then π describes a measurable decomposition of X into $(n-k)$-subvarieties $\partial^c \psi(y)$. By construction, the coarea formula yields the formula

$$g(y) = \int_{F^{-1}(y)} \frac{1}{JF(x)} \cdot f(x) d\mathscr{H}^{n-k}(x). \tag{1}$$

We review the definition of the Jacobian term $JF(x)$. Consider the derivative DF of F. Then $JF(x)$ is the square root of the sum of squares of all $k \times k$-minors of DF where $k = dim(Y)$. Equivalently one has $JF(x) = \sqrt{det(DF \ . \ {}^t DF)}$.

Our goal is to evaluate the volumes of the cells $\partial^c \psi(y)$, specifically the integral

$$\int_{\partial^c \psi(y)} 1.f(x) d\mathscr{H}^{n-k}(x) = vol_{f\mathscr{H}^{n-k}}[F^{-1}(y)] = vol_{f\mathscr{H}^{n-k}}[\partial^c \psi(y)].$$

Note that we need integrate the $(n-k)$-dimensional Hausdorff measure multiplied by the density $f(x)$. Now the question is how to relate the integrals along the fibres $\partial^c \psi(y)$ when the integrands are identically constant 1 and JF^{-1}. Obviously this is easiest if the Jacobian is constant along the fibres.

For applications to the width inequalities, we study the maximum $\max_y vol_{f\mathscr{H}^{n-k}}[\partial^c\psi(y)]$, and the min-max

$$\beta_{\sigma,c} := \min_{\tau}{}' \max_{y} vol_{f\mathscr{H}^{n-k}}[\partial^c\psi(y)].$$

We include the subscript "c" in β because the subdifferential $\partial^c\psi(y)$ depends on the geometry of the cost c.

But can we compare $width_k(X, g)$ with $\beta_{\sigma,c}$? When the Jacobian is constant along the fibres, then evidently the two quantites are directly related. For example when the dimension Y is zero, then the density $g(y)$ of the target is directly related to the volume of the cells $\partial^c\psi(y)$. [Error?]

References

[DT58] A. Dold and R. Thom. "Quasifaserungen und Unendliche Symmetrische Produkte". In: *Annals of Mathematics* 67.2 (1958), pp. 239–281. URL: http://www.jstor.org/stable/1970005.

[Mar] J.H. Martel. "Applications of Optimal Transport to Algebraic Topology: How to Build Spines from Singularity". PhD thesis. University of Toronto. URL: https://github.com/jhmartel/Thesis2019.