1 Statistical concepts on shape space

If we wish to define statistical concepts like mean and variance on our shape space - the manifold of unparametrized curves - we are faced with some challenges. To begin with our manifold fold is infinite-dimensional, and all theory developed so far in this exposition has been on finite-dimensional manifolds. This is an important part of the generalizations of mean and variance, since we can quite easily define a measure on M. If a point $x \in M$ can be written in finitely many local coordinates, a measure can be defined via the use of the metric on the induced (finite) basis of $T_x M$. It is not at all obvious how this method of constructing a measure on M generalizes to the infinite-dimensional case.

For example, the prime example of a infinite-dimensional shape space is a function space, e.g., the space of all continuous function. As is well-known, simply constructing a random variable on such a space is non-trivial; a fundamental construction is the Brownian motion, the existence of which relies on a deep result of Kolmogorov. Our space Imm is an example of a restriction of a space of continuous functions (to immersions), and the space $\mathcal{I}$ is even more involved. Thus, simply defining the random points to work with on the shape space promises some challenges.

Secondly, all results regarding existence and uniqueness of mean points have relied on either a global assumption or local assumptions on the Riemannian curvature of the manifold. The behaviour of the curvature of our shape space is not at all well understood, especially since the choice of metric is non-trivial to begin with. One might think that the curvature of $\mathcal{I}$ equipped with the L^2-metric is the least difficult to examine, but since the distance function induced by this metric is vanishing, the resulting variance, $\sigma_X^2(y)$ would be 0 for all random variables X and shapes y (if it is even possible to generalize the variance-formulas in section **??** to the infinite dimensional case.). Thus every point would be a mean point of X which is not a very informative statement.

To make the difficulties stand out explicitly, we write our Definition **??** when naively replacing the finite-dimensional manifold M with our shape space Imm and using an almost local metric. Then we have that y is some fixed shape and X is a random shape (still not defined):

$$\sigma_X^2(y) = \int_{\text{Imm}} \inf_{c \in \mathcal{C}_y^z} \{L_{G^\Phi}(c)\} \, p_X(z) \, \mathrm{d}z,$$

where we let $\mathcal{C}_y^z$ denote all path starting in y and ending in z. First of all, it is not clear that this integral over a function space is well-defined. Secondly, to calculate an explicit expression for the variance at y (and later on the mean), we would most likely need to be able to find a useful expression for the distance L_{G^Φ}, which is not trivial. Thirdly, if we succeed in making sense of the above

expression, the next step would be to consider how to minimize it, i.e., what we are actually interested in is

$$\operatorname*{argmin}_{y \in \mathrm{Imm}} \int_{\mathrm{Imm}} \inf_{c \in \mathcal{C}_y^z} \{L_{G^\Phi}(c)\} \, p_X(z) \, \mathrm{d}z.$$

This is optimization problem over a function space, so we could maybe use some tools from the calculus of variations to solve it.

All in all, at each step a lot of rather non-trivial problems need to be clarified, we imagine that one could proceed along the way sketched above.