# Geodesic and parallel models for leaf shape

Stephan F. Huckemann* and Thomas Hotz

Institute for Mathematical Stochastics, Georg-August Universität Göttingen

## 1 Introduction

Since more than a decade, many models for the evolution of shape under biological growth have been developed employing classical statistical methodology to tangent space projections of shape data called *Procrustes residuals*. Among them, are the *geodesic hypothesis* stating that

*biological growth evolves mainly along a geodesic in Kendall's shape spaces,*

by (Le and Kume(2000)) and the *parallel hypothesis* by (Morris et al.(2000)Morris, Kent, Mardia, and Aykroyd) stating that Procrustes residuals of similar biological entities follow curves parallel in a Euclidean geometry of the tangent space. In order to relate to the geodesic hypothesis, we consider here only curves that are images of geodesics. In collaboration with the Institute for Forest Biometry and Informatics at the University of Göttingen we consider in this study leaf shape data during a growing period collected non-destructively from two clones and a reference tree of black Canadian poplars at an experimental site, cf. Figure 1. The question investigated is:

*exploiting the geodesic and parallel hypotheses, how do genes affect leaf shape growth on trees? can it be modelled by two shape parameters, offset and direction?*

The answer is no, if Procrustes residuals are used. In particular the geodesic hypothesis has to be rejected for leaf shape growth of clones as well as the parallel hypothesis for leaf shape growth of clones versus a reference tree. Introducing curvature and more refined tests, however, both hypotheses can be accepted allowing for simple geodesic and parallel growth models.

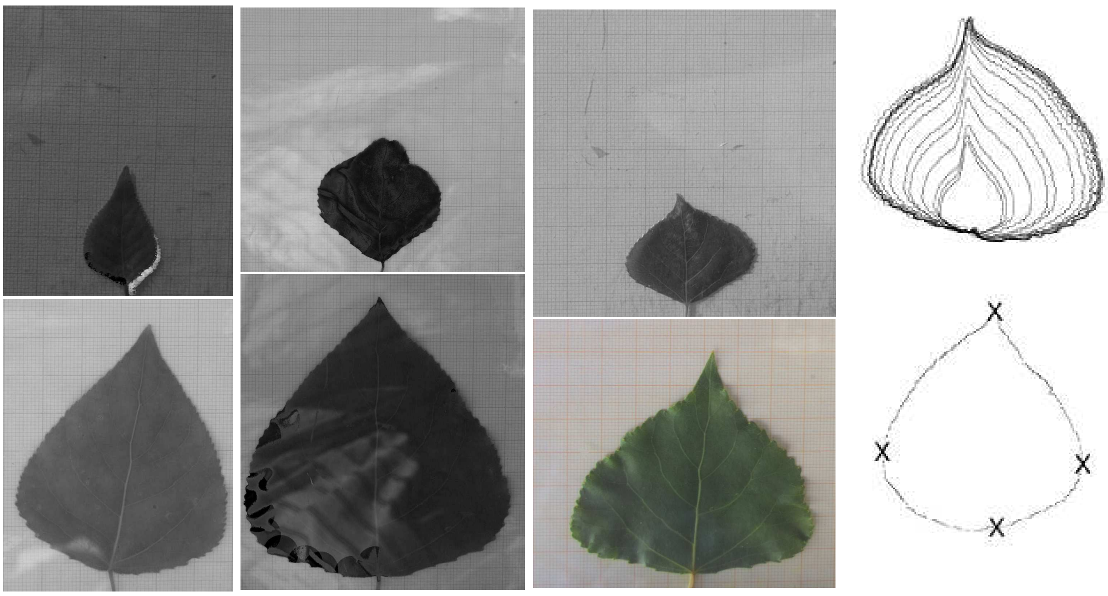

*Figure 1: Columns 1–3: young leaves (top row), old leaves (bottom row) of clone 1 (column 1), clone 2 (column 2) and a reference tree (column 3). Column 4: contours non-destructively extracted during growth (top) and a quadrangular configuration (bottom) with landmarks at petiole, tip and largest orthogonal extensions defining a shape in Kendall's shape space $\Sigma_2^4$.*

## 2 Tests for Shape Dynamics

For every leaf considered a *first geodesic principal component* (GPC – a generalization to manifolds of a first principal component direction) is computed either from its first two shapes (then the GPC is just the geodesic connecting the two) or from the rest of the shapes over a growing period (for an algorithm, see (Huckemann and Hotz(2009))). All of the following procedures produce data of two groups in a Euclidean space. The corresponding test then tests for a common mean via the classical Hotelling $T^2$-test. It is well known that the corresponding statistic is robust to some extend under nonnormality, one condition being finite higher order moments, a condition clearly met on a compact space. There is also asymptotic robustness under unequal covariances if the ratio of sample sizes tends to 1, e.g. (Lehmann(1997), p. 462).

**Tests for common means.** Following the classical scheme (e.g. (Dryden and Mardia(1998), Chapter 7)), all shapes of the two groups considered are projected to the tangent space of their overall Procrustes mean giving Procrustes residuals with the null hypothesis,

*Procrustes residuals of shape growth for every leaf have the same Euclidean mean.*

**Tests for common directions.** Following (Morris et al.(2000)Morris, Kent, Mardia, and Aykroyd), compute the Euclidean first principal component (PC) unit vector of each set of Procrustes residuals corresponding to the shapes of a single leaf's evolution pointing into the direction of growth. The residual tangent space coordinates of these directions at their common residual mean closer to the data are projected orthogonally to the Euclidean space of suitable dimension. The null hypothesis is then,

*Procrustes residuals of shape growth for every leaf share the same first PC.*

Even for common mean static shape, the null hypothesis is affected by curvature, cf. Figure 2.

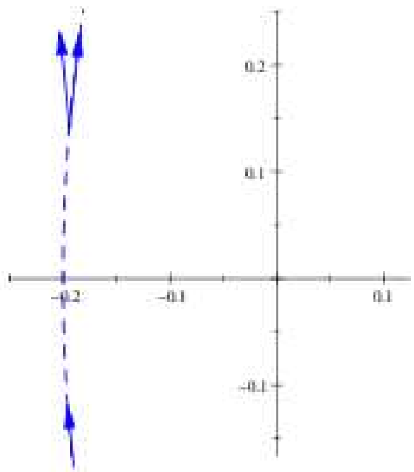


*Figure 2: Tangent space of the two-dimensional $\Sigma_2^3$ (obtained from $\Sigma_2^4$ by leaving out one landmark) under the inverse Riemann exponential at the intrinsic mean corresponding to the extent of the overall data (clones 1 + 2 and reference tree). The left top vector is the affine parallel transport of the bottom vector in the Euclidean tangent space, the right top vector is its intrinsic parallel transplant along a geodesic (dashed).*

**The geodesic test.** Every first GPC computed above determines a unique element in the space of unit speed geodesics of Kendall's space of planar shapes. Using the embedding of the space of pre-geodesics $O_2^H(2,3)$ into $\mathbb{C}^3 \times \mathbb{C}^3$ as detailed in (Huckemann(2010)), these elements are orthogonally projected to the tangent space of the two group's Ziezold mean thus giving data in a Euclidean space. The corresponding null hypothesis is then

*the temporal evolution of shape for every group follows a common geodesic.*

The test for common geodesics relies on a central limit theorem derived in (Huckemann(2010)).

**The test for parallelity.** On a two-sphere, any two different geodesics intersect but also pass parallel through the equator with respect to the intersection points. In general, parallel transport is not transitive for non flat spaces. Hence, the parallel hypothesis has no global version. Still, one may aim at *local* concepts of parallelity depending on specific offsets, however.

**Definition.** *Call $n$ oriented geodesic segments* parallel at offsets $x_1, \ldots, x_n$ with respect to another offset $x_0$, *if the parallel transplants to the tangent space of $x_0$ of their unit speed directions $v_1, \ldots, v_n$ in the tangent spaces at $x_1, \ldots, x_n$, respectively, are uniquely defined and identical.*

Here is a computationally simple null hypothesis,

> *all geodesic segments are parallel at their point nearest to the intrinsic mean over all intrinsic means with respect to that overall intrinsic mean,*

the version of the *parallel hypothesis* for the following. For the parallel transplants of the corresponding individual GPC's directions at the individual means which lie on a unit sphere around the origin of the tangent space at the overall intrinsic mean, residual tangent space coordinates as in the test for common directions can be chosen.

# 3 Data Analysis, Geodesic and Parallel Model

All of the above tests were applied to compare the initial leaf growth on one tree with the later leaf growth on a tree with different genes, see Table 1. Initial and subsequent leaf growth

| dataset 1 | dataset 2 | geodesics | parallelity | directions | means |
|---|---|---|---|---|---|
| **Comparison for different genes** | | | | | |
| clone 1 young (21) | reference young (12) | $0.001^{**}$ | 0.463 | 0.649 | $< 0.001^{**}$ |
| clone 2 young (11) | reference young (12) | $0.004^{**}$ | 0.492 | 0.789 | $0.002^{**}$ |
| clone 1 young (21) | reference old (9) | $0.007^{**}$ | 0.729 | $0.008^{**}$ | $< 0.001^{**}$ |
| clone 2 young (11) | reference old (9) | $0.026^{*}$ | 0.663 | $0.023^{*}$ | $< 0.001^{**}$ |
| clone 1 old (20) | reference young (12) | $< 0.001^{**}$ | 0.185 | $0.001^{**}$ | $< 0.001^{**}$ |
| clone 2 old (11) | reference young (12) | $0.009^{**}$ | $0.013^{*}$ | $0.014^{*}$ | $0.005^{**}$ |
| clone 1 old (20) | reference old (9) | 0.087 | 0.529 | $0.023^{*}$ | $0.001^{**}$ |
| clone 2 old (11) | reference old (9) | $0.021^{*}$ | 0.178 | $0.018^{*}$ | $0.007^{**}$ |
| **Comparison with identical genes** | | | | | |
| clone 1 young (21) | clone 2 old (11) | 0.753 | $0.035^{*}$ | $< 0.001^{**}$ | $< 0.001^{**}$ |
| clone 1 young (21) | clone 2 young (11) | 0.973 | 0.693 | 0.824 | 0.715 |
| clone 2 young (11) | clone 1 old (20) | 0.066 | $0.026^{*}$ | $0.002^{**}$ | $< 0.001^{**}$ |
| clone 1 old (20) | clone 2 old (11) | 0.171 | 0.078 | 0.210 | 0.713 |
| **Comparison with random samples** | | | | | |
| clone 1 young (21) | one-day sample (13) | $0.025^{*}$ | $< 0.001^{**}$ | 0.530 | $< 0.001^{**}$ |
| clone 2 young (11) | one-day sample (13) | 0.079 | $0.011^{*}$ | 0.573 | $< 0.001^{**}$ |
| reference 1 young (12) | one-day sample (13) | $< 0.001^{**}$ | $< 0.001^{**}$ | 0.310 | $< 0.001^{**}$ |
| clone 1 old (20) | one-day sample (13) | $0.002^{**}$ | $< 0.001^{**}$ | $0.029^{*}$ | $< 0.001^{**}$ |
| clone 2 old (11) | one-day sample (13) | $< 0.001^{**}$ | $< 0.001^{**}$ | 0.060 | $< 0.001^{**}$ |
| reference 1 old (9) | one-day sample (13) | $0.006^{**}$ | $0.006^{**}$ | 0.133 | $< 0.001^{**}$ |

*Table 1: $p$-values for several tests for the discrimination of clones from the reference tree via leaf growth ("young" denotes the dataset comprising of the first initial two observations and "old" the dataset comprising the remaining observations). The sample size (number of different leaves followed over their growing period) of the corresponding data set is reported in parentheses.*

on trees with the same gene material provided negative controls for the geodesic hypothesis. To obtain positive controls also for the parallel hypothesis indicating the tests' discrimantive power, leaves from a poplar tree, genetically different from the clones and the references tree, were collected during one day (no growth) and randomly assigned to 13 samples (each used in the test like showing a leaf being followed over time). From these results, even when taking Bonferroni adjustments for multiple testing into account, we conclude the following:

(a) Clone and reference tree can be discriminated by partial observations of leaf shape growth not necessarily covering the same interval of the growing period via the test for common geodesics. This is not possible via a test for common means (due to temporal change of shape) or common directions (due to curvature).

(b) Only the geodesic test is appropriate for and validates the geodesic hypothesis.

(c) The test for common directions is unsuitable to investigate the parallel hypothesis. Using the test for parallelity, parallel leaf shape growth is not rejected for genetically different trees.

(d) To identify typical leaf shape growth, two initial observations per leaf suffice.

This gives rise to the *geodesic model* for the shape of a typical leaf of tree $i$ at time $t$:

$$[z_i(t)] = [x_i \cos(f_i(t)) + v_i \sin(f_i(t)) + \epsilon_i(t)]$$

with unit vectors $v_i$ horizontal at pre-shapes $x_i$, functions $f_i$ relating time to size and isotropic error $\epsilon_i(t)$ with zero mean and finite variance. With a common unit vector $v$ horizontal at a common pre-shape $z$ and the horizontal lift of the shape space's parallel transport $\theta_{z,x_i}$ from the horizontal space at $z$ to the horizontal space at $x_i$ (cf. (Huckemann et al.(2010)Huckemann, Hotz, and Munk)), the geodesic model turns into the *parallel model* upon setting $v_i = \theta_{z,x_i}(v)$.

# Acknowledgment

The authors would like to thank their colleagues from the Institute for Forest Biometry and Informatics at the University of Göttingen whose interest in leaf growth modeling prompted this research, in particular Michael Henke for supplying with the leaf data. Also, the authors gratefully acknowledge support by DFG grant MU 1230/10-1 and DFG SFB 803.

# References

Dryden, I. L., Mardia, K. V., 1998. Statistical Shape Analysis. Wiley, Chichester.

Huckemann, S., 2010. Intrinsic inference on the mean geodesic and tree discrimination by leaf growth. Submitted.

Huckemann, S., Hotz, T., 2009. Principal components geodesics for planar shape spaces. Journal of Multivariate Analysis 100, 699–714.

Huckemann, S., Hotz, T., Munk, A., 2010. Intrinsic MANOVA for Riemannian manifolds with an application to Kendall's space of planar shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 32 (4), 593–603.

Le, H., Kume, A., 2000. Detection of shape changes in biological features. Journal of Microscopy 200 (2), 140–147.

Lehmann, E. L., 1997. Testing Statistical Hypotheses (Springer Texts in Statistics). Springer.

Morris, R., Kent, J. T., Mardia, K. V., Aykroyd, R. G., 2000. A parallel growth model for shape. In: Arridge, S., Todd-Pokropek, A. (Eds.), Proceedings in Medical Imaging Understanding and Analysis. Bristol: BMVA, pp. 171–174.