

SALZBURG MATHEMATICS COLLOQUIUM

Winter 2017/2018

Ruriko Yoshida (Monterey/USA)

„Principal Component Analysis In The Space Of Phylogenetic Trees“

January 11, 2018

Abstract:

Principal component analysis (PCA) is the most popular approach to rendering two- or three-dimensional representations of the major trends in such multidimensional data. PCA offers a means of quantifying variation and summarizing a collection of phylogenies by dimensional reduction. However, the space of all possible phylogenies on a fixed set of species does not form a Euclidean vector space, so PCA must be reformulated in the geometry of tree-space, which is a CAT(0) geodesic metric space. Here we propose a geometric object which represents a k -th order principal component: the locus of the weighted Fréchet mean of $k + 1$ points in tree-space, where the weights vary over the standard k -dimensional simplex. We establish basic properties of these objects, in particular that locally they generically have dimension k , and we propose an efficient algorithm for projection onto these surfaces. Simulation studies confirm these algorithms perform well, and they are applied to data sets of Apicomplexa gene trees and the African coelacanth genome. This is joint work with T. Nye, X. Tang and G. Weyenberg.

Thursday, 15:00-15:45

Hörsaal 414, 1. Stock