Title: When is the past in the past?: Modelling the effect of demographic history on sample allele spectra via polynomial diffusions

Authors: Lynette Caitlin Mikula, Claus Vogl

Analytically tractable solutions to Wright-Fisher/Moran diffusion densities can be obtained via spectral decomposition of the associated diffusion generators. Coupling such forward diffusion results with polynomial expansions of sampling schemes at the extant time enables exact calculation of expected sample allele frequency distributions. Our starting point is a computationally efficient bi-allelic diffusion model with low scaled overall mutation rates: Mutations enter at the monomorphic boundaries, and traverse the polymorphic region through drift. Changing population sizes can be modelled by re-scaling the temporal component of the spectral decomposition. This modulates the trajectories of the polymorphic allele frequencies after a mutation has occurred by the corresponding eigenvalues. The allele frequencies themselves determine the temporally unaffected spatial component, constituted by eigenvectors of weighted orthogonal polynomials. Piecewise deterministic and stochastic changes in population size can be incorporated into the model by this re-scaling. Calculating expected allele frequencies for changes of different magnitude, strength, and age help answer questions such as: Why are the effective population sizes of species with boom-bust life cycles, e.g. Drosophila, difficult to estimate? What is the effect of minor, stochastic fluctuations in population size? More generally, are sample allele frequency spectra in equilibrium when we expect them to be?—Overall, we propose that orthogonal polynomial diffusion approaches offer a unique perspective on temporally changing allele frequency distributions. Under simplifying assumptions, statistically optimal inference of past scaled mutation rates is even possible although the information in the allele frequency spectra is limited.