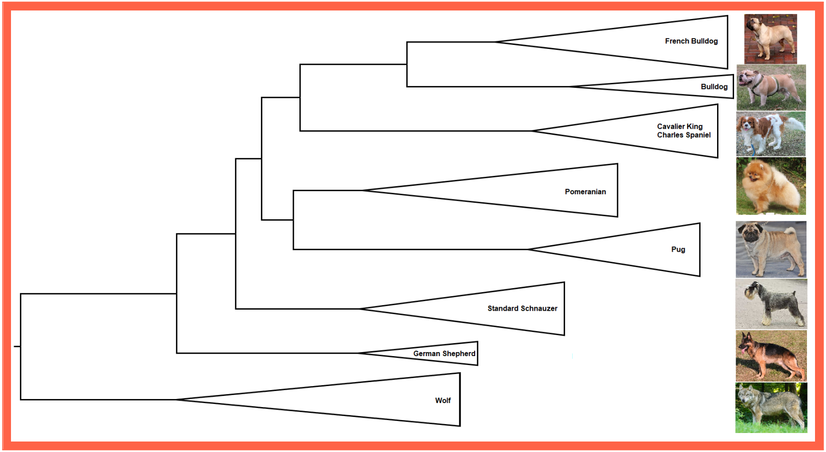
What is...tropical geometry - part 22?

Or: Tropical applications 2 - Phylogenetic trees

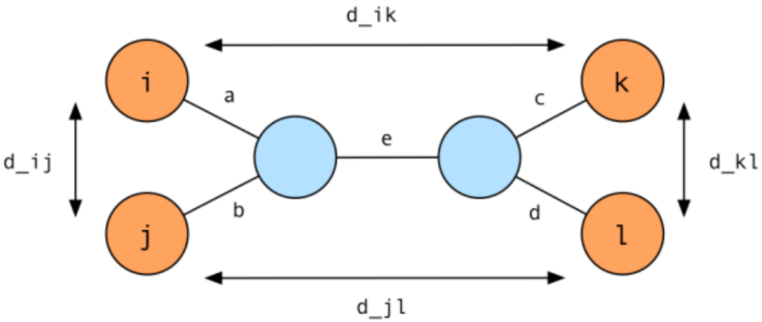
Phylogenetics: math formulation



- **Goal** Reconstruct evolutionary trees from DNA data
- **Input** A distance matrix between species
- **Output** A metric tree with leaves labeled by species

Tree metrics

$$d_{ij} + d_{kl} \leq d_{ik} + d_{jl} = d_{il} + d_{jk}$$

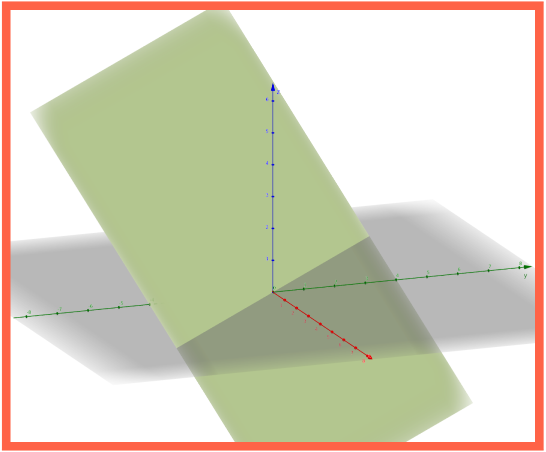


Picture from https://www.kuniga.me/blog/2019/05/10

- A tree metric $d(i, j)$ = path length between leaves i, j
- Characterized by the four-point condition (above)
- Space of tree metrics = all distance matrices satisfying this condition

Enter tropical geometry

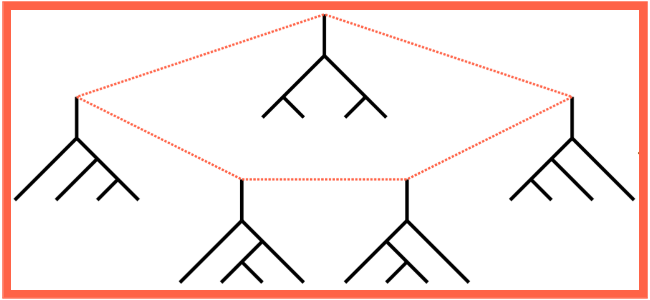
An element of $G(2,3)$:



- $G(2,n)$ = classical Grassmannian of 2-planes in $\mathbb{C}^n$
- $\mathrm{Trop}(G(2,n))$ = its tropicalization, a polyhedral fan
- This variety turns out to encode tree space

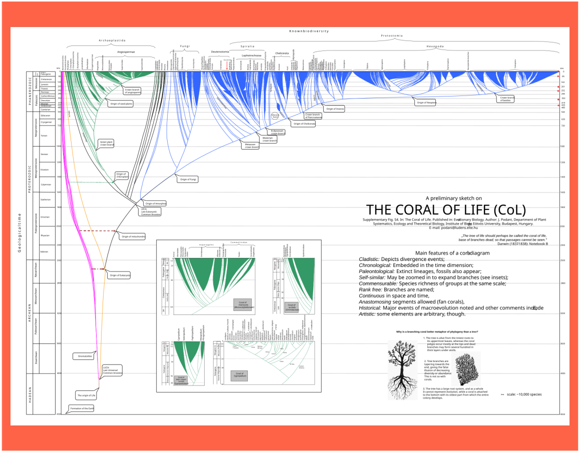
For completeness: A formal statement

$\mathrm{Trop}(G(2,n)) = $ space of phylogenetic trees on n leaves



- Cones of $\mathrm{Trop}(Gr(2,n)) \leftrightsquigarrow$ different tree topologies
- Coordinates in each cone = edge lengths
- Reference: Speyer–Sturmfels (2004), *The Tropical Grassmannian*

Applications



- Compute distance matrix from DNA/protein data
- Project onto the tropical Grassmannian (tree space)
- Recover the phylogenetic tree efficiently and robustly

Thank you for your attention!

I hope that was of some help.