ARTICLES

Tropical Mathematics

DAVID SPEYER
Massachusetts Institute of Technology
Cambridge, MA 02139
speyer@math.mit.edu

BERND STURMFELS
University of California at Berkeley
Berkeley, CA 94720
bernd@math.berkeley.edu

This article is based on the Clay Mathematics Senior Scholar Lecture that was delivered by Bernd Sturmfels in Park City, Utah, on July 22, 2004. The topic of this lecture was the *tropical approach* in mathematics. This approach was in its infancy at that time, but it has since matured and is now an integral part of geometric combinatorics and algebraic geometry. It has also expanded into mathematical physics, number theory, symplectic geometry, computational biology, and beyond. We offer an elementary introduction to this subject, touching upon arithmetic, polynomials, curves, phylogenetics, and linear spaces. Each section ends with a suggestion for further research. The proposed problems are particularly well suited for undergraduate students. The bibliography contains numerous references for further reading in this field.

The adjective *tropical* was coined by French mathematicians, including Jean-Eric Pin [**16**], in honor of their Brazilian colleague Imre Simon [**19**], who was one of the pioneers in what could also be called *min-plus algebra*. There is no deeper meaning in the adjective *tropical*. It simply stands for the French view of Brazil.

Arithmetic

Our basic object of study is the *tropical semiring* $(\mathbb{R} \cup \{\infty\}, \oplus, \odot)$. As a set this is just the real numbers $\mathbb{R}$, together with an extra element ∞ that represents infinity. However, we redefine the basic arithmetic operations of addition and multiplication of real numbers as follows:

$$x \oplus y := \min(x, y) \qquad \text{and} \qquad x \odot y := x + y.$$

In words, the *tropical sum* of two numbers is their minimum, and the *tropical product* of two numbers is their sum. Here are some examples of how to do arithmetic in this strange number system. The tropical sum of 3 and 7 is 3. The tropical product of 3 and 7 equals 10. We write these as

$$3 \oplus 7 = 3 \qquad \text{and} \qquad 3 \odot 7 = 10.$$

Many of the familiar axioms of arithmetic remain valid in tropical mathematics. For instance, both addition and multiplication are *commutative*:

$$x \oplus y = y \oplus x \qquad \text{and} \qquad x \odot y = y \odot x.$$

The *distributive law* holds for tropical multiplication over tropical addition:

$$x \odot (y \oplus z) = x \odot y \oplus x \odot z,$$

where no parentheses are needed on the right, provided we respect the usual order of operations: Tropical products must be completed before tropical sums. Here is a numerical example to illustrate:

$$3 \odot (7 \oplus 11) = 3 \odot 7 = 10,$$
$$3 \odot 7 \oplus 3 \odot 11 = 10 \oplus 14 = 10.$$

Both arithmetic operations have a neutral element. Infinity is the *neutral element* for addition and zero is the *neutral element* for multiplication:

$$x \oplus \infty = x \qquad \text{and} \qquad x \odot 0 = x.$$

Elementary school students tend to prefer tropical arithmetic because the multiplication table is easier to memorize, and even long division becomes easy. Here are the tropical *addition table* and the tropical *multiplication table*:

$\oplus$	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	1	1	1	1	1	1	1
2	1	2	2	2	2	2	2
3	1	2	3	3	3	3	3
4	1	2	3	4	4	4	4
5	1	2	3	4	5	5	5
6	1	2	3	4	5	6	6
7	1	2	3	4	5	6	7

$\odot$	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	2	3	4	5	6	7	8
2	3	4	5	6	7	8	9
3	4	5	6	7	8	9	10
4	5	6	7	8	9	10	11
5	6	7	8	9	10	11	12
6	7	8	9	10	11	12	13
7	8	9	10	11	12	13	14

But watch out: tropical arithmetic is tricky when it comes to subtraction. There is no x to call "10 minus 3" because the equation $3 \oplus x = 10$ has no solution x at all. To stay on safe ground, we content ourselves with using addition $\oplus$ and multiplication $\odot$ only.

It is extremely important to remember that 0 is the multiplicative identity element. For instance, the tropical *Pascal's triangle*, whose rows are the coefficients appearing in a binomial expansion, looks like this:

$$\begin{array}{ccccccccc} & & & & 0 & & & & \\ & & & 0 & & 0 & & & \\ & & 0 & & 0 & & 0 & & \\ & 0 & & 0 & & 0 & & 0 & \\ 0 & & 0 & & 0 & & 0 & & 0 \\ \cdots & \cdots & & \cdots & & \cdots & & \cdots & \cdots \end{array}$$

For example, the fourth row in the triangle represents the identity

$$\begin{aligned} (x \oplus y)^3 &= (x \oplus y) \odot (x \oplus y) \odot (x \oplus y) \\ &= 0 \odot x^3 \oplus 0 \odot x^2 y \oplus 0 \odot x y^2 \oplus 0 \odot y^3. \end{aligned}$$

Of course, the zero coefficients can be dropped in this identity:

$$(x \oplus y)^3 = x^3 \oplus x^2 y \oplus x y^2 \oplus y^3.$$

Moreover, the *Freshman's Dream* holds for all powers in tropical arithmetic:

$$(x \oplus y)^3 = x^3 \oplus y^3.$$

The three displayed identities are easily verified by noting that the following equations hold in classical arithmetic for all $x, y \in \mathbb{R}$:

$$3 \cdot \min\{x, y\} = \min\{3x, 2x + y, x + 2y, 3y\} = \min\{3x, 3y\}.$$

Research problem The tropical semiring generalizes to higher dimensions: The set of convex polyhedra in $\mathbb{R}^n$ can be made into a semiring by taking $\odot$ as "Minkowski sum" and $\oplus$ as "convex hull of the union." A natural subalgebra is the set of all polyhedra that have a fixed *recession cone* C. If $n = 1$ and $C = \mathbb{R}_{\geq 0}$, this is the tropical semiring. Develop linear algebra and algebraic geometry over these semirings, and implement efficient software for doing arithmetic with polyhedra when $n \geq 2$.

Polynomials

Let $x_1, \ldots, x_n$ be variables that represent elements in the tropical semiring $(\mathbb{R} \cup \{\infty\}, \oplus, \odot)$. A *monomial* is any product of these variables, where repetition is allowed. By commutativity and associativity, we can sort the product and write monomials in the usual notation, with the variables raised to exponents,

$$x_2 \odot x_1 \odot x_3 \odot x_1 \odot x_4 \odot x_2 \odot x_3 \odot x_2 = x_1^2 x_2^3 x_3^2 x_4,$$

as long as we know from context that x_1^2 means $x_1 \odot x_1$ and not $x_1 \cdot x_1$. A monomial represents a function from $\mathbb{R}^n$ to $\mathbb{R}$. When evaluating this function in classical arithmetic, what we get is a linear function:

$$x_2 + x_1 + x_3 + x_1 + x_4 + x_2 + x_3 + x_2 = 2x_1 + 3x_2 + 2x_3 + x_4.$$

Although our examples used positive exponents, there is no need for such a restriction, so we allow negative integer exponents, so that every linear function with integer coefficients arises in this manner.

FACT 1. *Tropical monomials are the linear functions with integer coefficients.*

A *tropical polynomial* is a finite linear combination of tropical monomials:

$$p(x_1, \ldots, x_n) = a \odot x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} \oplus b \odot x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n} \oplus \cdots$$

Here the coefficients $a, b, \ldots$ are real numbers and the exponents $i_1, j_1, \ldots$ are integers. Every tropical polynomial represents a function $\mathbb{R}^n \to \mathbb{R}$. When evaluating this function in classical arithmetic, what we get is the minimum of a finite collection of linear functions, namely,

$$p(x_1, \ldots, x_n) = \min\bigl(a + i_1 x_1 + \cdots + i_n x_n,\ b + j_1 x_1 + \cdots + j_n x_n, \ldots\bigr).$$

This function $p : \mathbb{R}^n \to \mathbb{R}$ has the following three important properties:

- p is continuous,
- p is piecewise-linear, where the number of pieces is finite, and
- p is concave, that is, $p(\frac{\mathbf{x}+\mathbf{y}}{2}) \geq \frac{1}{2}(p(\mathbf{x}) + p(\mathbf{y}))$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

It is known that every function that satisfies these three properties can be represented as the minimum of a finite set of linear functions. We conclude:

FACT 2. *The tropical polynomials in n variables $x_1, \ldots, x_n$ are precisely the piecewise-linear concave functions on $\mathbb{R}^n$ with integer coefficients.*

As a first example consider the general cubic polynomial in one variable x,

$$p(x) = a \odot x^3 \oplus b \odot x^2 \oplus c \odot x \oplus d. \tag{1}$$

To graph this function we draw four lines in the (x, y) plane: $y = 3x + a$, $y = 2x + b$, $y = x + c$, and the horizontal line $y = d$. The value of $p(x)$ is the smallest y-value such that (x, y) is on one of these four lines, that is, the graph of $p(x)$ is the lower envelope of the lines. All four lines actually contribute if

$$b - a \le c - b \le d - c. \tag{2}$$

These three values of x are the breakpoints where $p(x)$ fails to be linear, and the cubic has a corresponding factorization into three linear factors:

$$p(x) = a \odot (x \oplus (b - a)) \odot (x \oplus (c - b)) \odot (x \oplus (d - c)). \tag{3}$$

See FIGURE 1 for the graph and the roots of the cubic polynomial $p(x)$.

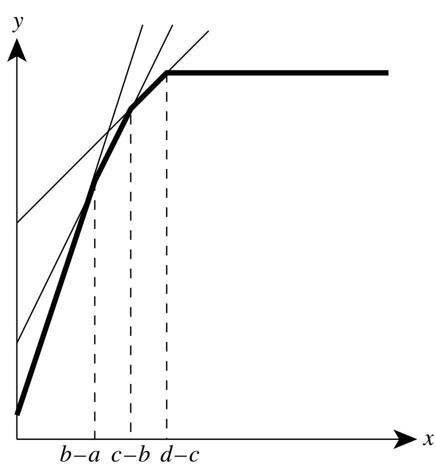

Figure 1 The graph of a cubic polynomial and its roots

Every tropical polynomial function can be written uniquely as a tropical product of tropical linear functions (in other words, the *Fundamental Theorem of Algebra* holds tropically). In this statement we must emphasize the word *function*. Distinct polynomials can represent the same function. We are not claiming that every polynomial factors as a product of linear polynomials. What we are claiming is that every polynomial can be replaced by an equivalent polynomial, representing the same function, that can be factored into linear factors. For example, the following polynomials represent the same function:

$$x^2 \oplus 17 \odot x \oplus 2 = x^2 \oplus 1 \odot x \oplus 2 = (x \oplus 1)^2.$$

Unique factorization of polynomials no longer holds in two or more variables. Here the situation is more interesting. Understanding it is our next problem.

Research problem The factorization of multivariate tropical polynomials into irreducible tropical polynomials is not unique. Here is a simple example:

$$\begin{aligned}&(0 \odot x \oplus 0) \odot (0 \odot y \oplus 0) \odot (0 \odot x \odot y \oplus 0)\\ &\quad = (0 \odot x \odot y \oplus 0 \odot x \oplus 0) \odot (0 \odot x \odot y \oplus 0 \odot y \oplus 0).\end{aligned}$$

Develop an algorithm (with implementation and complexity analysis) for computing all the irreducible factorizations of a given tropical polynomial. Gao and Lauder [**8**] have shown the importance of tropical factorization for the problem of factoring multivariate polynomials in the classical sense.

Curves

A tropical polynomial function $p : \mathbb{R}^n \to \mathbb{R}$ is given as the minimum of a finite set of linear functions. We define the *hypersurface* $\mathcal{H}(p)$ to be the set of all points $\mathbf{x} \in \mathbb{R}^n$ at which this minimum is attained at least twice. Equivalently, a point $\mathbf{x} \in \mathbb{R}^n$ lies in $\mathcal{H}(p)$ if and only if p is not linear at $\mathbf{x}$. For example, if $n = 1$ and p is the cubic in (1) with the assumption (2), then

$$\mathcal{H}(p) = \big\{b - a, c - b, d - c\big\}.$$

Thus the hypersurface $\mathcal{H}(p)$ is the set of "roots" of the polynomial $p(x)$.

In this section we consider the case of a polynomial in two variables:

$$p(x, y) = \bigoplus_{(i,j)} c_{ij} \odot x^i \odot y^j.$$

FACT 3. *For a polynomial in two variables, p, the tropical curve $\mathcal{H}(p)$ is a finite graph embedded in the plane $\mathbb{R}^2$. It has both bounded and unbounded edges, all of whose slopes are rational, and the graph satisfies a zero tension condition around each node, as follows:*

Consider any node (x, y) of the graph, which we may as well take to be the origin, $(0, 0)$. Then the edges adjacent to this node lie on lines with rational slopes. On each such ray emanating from the origin consider the smallest nonzero lattice vector. *Zero tension* at (x, y) means that the sum of these vectors is zero.

Our first example is a *line* in the plane. It is defined by a polynomial:

$$p(x, y) = a \odot x \oplus b \odot y \oplus c \qquad \text{where } a, b, c \in \mathbb{R}.$$

The curve $\mathcal{H}(p)$ consists of all points (x, y) where the function

$$p : \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto \min\big(a + x, b + y, c\big)$$

is not linear. It consists of three half-rays emanating from the point $(x, y) = (c - a, c - b)$ into northern, eastern, and southwestern directions. The zero tension condition amounts to $(1, 0) + (0, 1) + (-1, -1) = (0, 0)$.

Here is a general method for drawing a tropical curve $\mathcal{H}(p)$ in the plane. Consider any term $\gamma \odot x^i \odot y^j$ appearing in the polynomial p. We represent this term by the point (γ, i, j) in $\mathbb{R}^3$, and we compute the convex hull of these points in $\mathbb{R}^3$. Now project the lower envelope of that convex hull into the plane under the map $\mathbb{R}^3 \to \mathbb{R}^2$, $(\gamma, i, j) \mapsto (i, j)$. The image is a planar convex polygon together with a distinguished subdivision Δ into smaller polygons. The tropical curve $\mathcal{H}(p)$ (actually its negative) is the *dual graph* to this subdivision. Recall that the dual to a planar graph is another planar graph whose vertices are the regions of the primal graph and whose edges represent adjacent regions.

As an example we consider the general quadratic polynomial

$$p(x, y) = a \odot x^2 \oplus b \odot xy \oplus c \odot y^2 \oplus d \odot x \oplus e \odot y \oplus f.$$

Then Δ is a subdivision of the triangle with vertices $(0, 0)$, $(0, 2)$, and $(2, 0)$. The lattice points $(0, 1)$, $(1, 0)$, $(1, 1)$ can be used as vertices in these subdivisions. Assuming that $a, b, c, d, e, f \in \mathbb{R}$ satisfy the conditions

$$2b \le a + c, 2d \le a + f, 2e \le c + f,$$

the subdivision Δ consists of four triangles, three interior edges, and six boundary edges. The curve $\mathcal{H}(p)$ has four vertices, three bounded edges, and six half-rays (two northern, two eastern, and two southwestern). In FIGURE 2, we show the negative of the quadratic curve $\mathcal{H}(p)$ in bold with arrows. It is the dual graph to the subdivision Δ which is shown in thin lines.

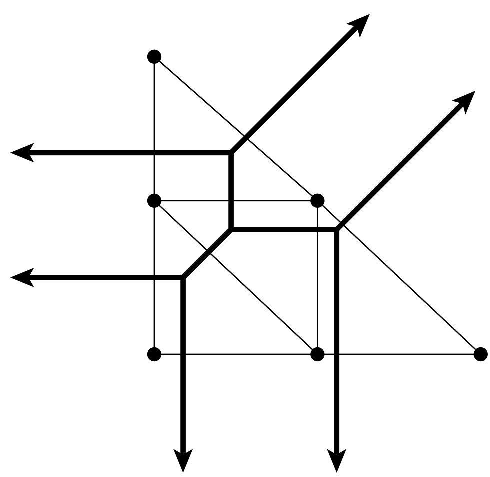

Figure 2 The subdivision Δ and the tropical curve

FACT 4. *Tropical curves intersect and interpolate like algebraic curves do.*

1. *Two general lines meet in one point, a line and a quadric meet in two points, two quadrics meet in four points, etc.*
2. *Two general points lie on a unique line, five general points lie on a unique quadric, etc.*

For a general discussion of *Bézout's Theorem* in tropical algebraic geometry, illustrated on the MAGAZINE cover, we refer to the article [**17**].

Research problem Classify all combinatorial types of *tropical curves in* 3*-space* of degree d. Such a curve is a finite embedded graph of the form

$$C = \mathcal{H}(p_1) \cap \mathcal{H}(p_2) \cap \cdots \cap \mathcal{H}(p_r) \subset \mathbb{R}^3,$$

where the p_i are tropical polynomials, C has d unbounded parallel halfrays in each of the four coordinate directions, and all other edges of C are bounded.

Phylogenetics

An important problem in computational biology is to construct a *phylogenetic tree* from distance data involving n leaves. In the language of biologists, the labels of the leaves are called *taxa*. These taxa might be organisms or genes, each represented by a

DNA sequence. For an introduction to phylogenetics we recommend books by Felsenstein [**7**] and Semple and Steele [**18**]. Here is an example, for $n = 4$, to illustrate how such data might arise. Consider an alignment of four genomes:

$$\begin{array}{rl} \text{Human:} & ACAATGTCATTAGCGAT\ldots \\ \text{Mouse:} & ACGTTGTCAATAGAGAT\ldots \\ \text{Rat:} & ACGTAGTCATTACACAT\ldots \\ \text{Chicken:} & GCACAGTCAGTAGAGCT\ldots \end{array}$$

From such sequence data, computational biologists infer the distance between any two taxa. There are various algorithms for carrying out this inference. They are based on statistical models of evolution. For our discussion, we may think of the distance between any two strings as a refined version of the Hamming distance (= the proportion of characters where they differ). In our (Human, Mouse, Rat, Chicken) example, the inferred distance matrix might be the following symmetric 4×4-matrix:

$$\begin{array}{ccccc} & H & M & R & C \\ H & 0 & 1.1 & 1.0 & 1.4 \\ M & 1.1 & 0 & 0.3 & 1.3 \\ R & 1.0 & 0.3 & 0 & 1.2 \\ C & 1.4 & 1.3 & 1.2 & 0 \end{array}$$

The problem of phylogenetics is to construct a tree with edge lengths that represent this distance matrix, provided such a tree exists. In our example, a tree does exist, as depicted in FIGURE 3, where the number next to the each edge is its length. The distance between two leaves is the sum of the lengths of the edges on the unique path between the two leaves. For instance, the distance in the tree between "Human" and "Mouse" is $0.6 + 0.3 + 0.2 = 1.1$, which is the corresponding entry in the 4×4-matrix.

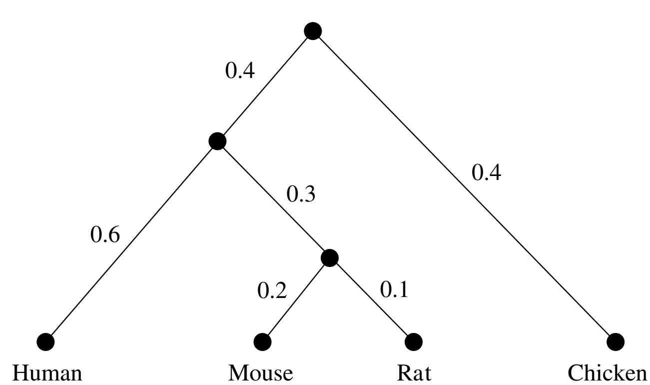

Figure 3 A phylogenetic tree

In general, considering n taxa, the *distance* between taxon i and taxon j is a positive real number d_{ij} which has been determined by some bio-statistical method. So, what we are given is a real symmetric $n \times n$-matrix

$$D = \begin{pmatrix} 0 & d_{12} & d_{13} & \cdots & d_{1n} \\ d_{12} & 0 & d_{23} & \cdots & d_{2n} \\ d_{13} & d_{23} & 0 & \cdots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{1n} & d_{2n} & d_{3n} & \cdots & 0 \end{pmatrix}.$$

We may assume that D is a *metric*, meaning that the triangle inequalities $d_{ik} \leq d_{ij} + d_{jk}$ hold for all i, j, k. This can be expressed by matrix multiplication:

FACT 5. *The matrix D represents a metric if and only if $D \odot D = D$.*

We say that a metric D on $\{1, 2, \ldots, n\}$ is a *tree metric* if there exists a tree T with n leaves, labeled $1, 2, \ldots, n$, and a positive length for each edge of T, such that the distance from leaf i to leaf j is d_{ij} for all i, j. Tree metrics occur naturally in biology because they model an evolutionary process that led to the n taxa.

Most metrics D are not tree metrics. If we are given a metric D that arises from some biological data then it is reasonable to assume that there exists a tree metric D_T that is close to D. Biologists use a variety of algorithms (for example, "neighbor joining") to construct such a nearby tree T from the given data D. In what follows we state a tropical characterization of tree metrics.

Let $X = (X_{ij})$ be a symmetric matrix with zeros on the diagonal whose $\binom{n}{2}$ distinct off-diagonal entries are unknowns. For each quadruple $\{i, j, k, l\} \subset \{1, 2, \ldots, n\}$ we consider the following tropical polynomial of degree two:

$$p_{ijkl} = X_{ij} \odot X_{kl} \oplus X_{ik} \odot X_{jl} \oplus X_{il} \odot X_{jk}. \tag{4}$$

This polynomial is the *tropical Grassmann-Plücker relation*, and it is simply the tropical version of the classical Grassmann-Plücker relation among the 2×2-subdeterminants of a 2×4-matrix [**14**, Theorem 3.20].

It defines a hypersurface $\mathcal{H}(p_{ijkl})$ in the space $\mathbb{R}^{\binom{n}{2}}$. The *tropical Grassmannian* is the intersection of these $\binom{n}{4}$ hypersurfaces. It is denoted

$$Gr_{2,n} = \bigcap_{1 \leq i < j < k < l \leq n} \mathcal{H}(p_{ijkl}).$$

This subset of $\mathbb{R}^{\binom{n}{2}}$ has the structure of a *polyhedral fan*, which means that it is the union of finitely many convex polyhedral cones that fit together nicely.

FACT 6. *A metric D on $\{1, 2, \ldots, n\}$ is a tree metric if and only if its negative $X = -D$ is a point in the tropical Grassmannian $Gr_{2,n}$.*

The statement is a reformulation of the *Four Point Condition* in phylogenetics, which states that D is a tree metric if and only if, for all $1 \leq i < j < k < l \leq n$, the *maximum* of the three numbers $D_{ij} + D_{kl}$, $D_{ik} + D_{jl}$, and $D_{il} + D_{jk}$ is attained at least twice. For $X = -D$, this means that the *minimum* of the three numbers $X_{ij} + X_{kl}$, $X_{ik} + X_{jl}$, and $X_{il} + X_{jk}$ is attained at least twice, or, equivalently, $X \in \mathcal{H}(p_{ijkl})$. The tropical Grassmannian $Gr_{2,n}$ is also known as the *space of phylogenetic trees* [**3, 14, 20**]. The combinatorial structure of this beautiful space is well studied and well understood.

Often, instead of measuring the pairwise distances between the various taxa, it can be statistically more accurate to consider all r-tuples of taxa and jointly measure the dissimilarity within each r-tuple. For example, in the above tree, the joint dissimilarity of the triple {Human, Mouse, Rat} is 1.2, the sum of the lengths of all edges in the subtree containing the mouse, human, and rat. Lior Pachter and the first author showed in [**15**] that it is possible to reconstruct the tree from the data for all r-tuples, as long as $n \geq 2r - 1$.

At this point in the original 2004 lecture notes, we had posed a research problem: to characterize the image of the given embedding of $Gr_{2,n}$ into $\mathbb{R}^{\binom{n}{r}}$, particularly in the case $r = 3$. Since then, Christiano Bocci and Filip Cools [**4**] have solved the problem for $r = 3$, and they made significant progress on the problem for higher r. While there is still work to be done, we now suggest the following less studied problem, borrowed from the end of [**14**, Chapter 3].

Research problem We say that a metric D has *phylogenetic rank* $\leq k$ if there exist k tree metrics $D^{(1)}, D^{(2)}, \ldots, D^{(k)}$ such that

$$D_{ij} = \max\left(D_{ij}^{(1)}, D_{ij}^{(2)}, \ldots, D_{ij}^{(k)}\right) \qquad \text{for all } 1 \leq i, j \leq n.$$

Equivalently, the matrix $X = -D$ is the sum of the matrices $X^{(i)} = -D^{(i)}$:

$$X = X^{(1)} \oplus X^{(2)} \oplus \cdots \oplus X^{(k)}.$$

The aim of the notion of phylogenetic rank is to model distance data that is a mixture of k different evolutionary histories. The set of metrics of phylogenetic rank $\leq k$ is a polyhedral fan in $\mathbb{R}^{\binom{n}{2}}$. Compute this fan, and explore its combinatorial, geometric, and topological properties, especially for $k = 2$.

Tropical linear spaces

Generalizing our notion of a line, we define a *tropical hyperplane* to be a subset of $\mathbb{R}^n$ of the form $\mathcal{H}(\ell)$, where ℓ is a tropical linear function in n unknowns:

$$\ell(x) = a_1 \odot x_1 \oplus a_2 \odot x_2 \oplus \cdots \oplus a_n \odot x_n.$$

Here $a_1, \ldots, a_n$ are arbitrary real constants. Solving linear equations in tropical mathematics means computing the intersection of finitely many hyperplanes $\mathcal{H}(\ell)$. It is tempting to define tropical linear spaces simply as intersections of tropical hyperplanes. However, this would not be a good definition because such arbitrary intersections can have mixed dimension, and they do not behave the way linear spaces do in classical geometry.

A better notion of tropical linear space is derived by allowing only those intersections of hyperplanes that are "sufficiently complete," in a sense we explain later. The definition we offer directly generalizes our discussion about phylogenetics. The idea is that phylogenetic trees are lines in tropical projective space, whose Plücker coordinates X_{ij} are the negated pairwise distances d_{ij}.

We consider the $\binom{n}{d}$-dimensional space $\mathbb{R}^{\binom{n}{d}}$ whose coordinates $X_{i_1 \cdots i_d}$ are indexed by d-element subsets $\{i_1, \ldots, i_d\}$ of $\{1, 2, \ldots, n\}$. Let S be any $(d-2)$-element subset of $\{1, 2, \ldots, n\}$ and let i, j, k, and l be any four distinct indices in $\{1, \ldots, n\} \backslash S$. The corresponding *three-term Grassmann Plücker relation* $p_{S,ijkl}$ is the following tropical polynomial of degree two:

$$p_{S,ijkl} = X_{Sij} \odot X_{Skl} \oplus X_{Sik} \odot X_{Sjl} \oplus X_{Sil} \odot X_{Sjk}. \tag{5}$$

We define the *Dressian* to be the intersection

$$Dr_{d.n} = \bigcap_{S,i,j,k,l} \mathcal{H}(p_{S,ijkl}) \subset \mathbb{R}^{\binom{n}{d}},$$

where the intersection is over all S, i, j, k, l as above. The term *Dressian* refers to Andreas Dress, an algebraist who now works in computational biology. For relevant references to his work and further details see [**11**].

Note that in the special case $d = 2$ we have $S = \emptyset$, the polynomial (5) is the four point condition in (4). In this special case, $Dr_{2,n} = Gr_{2,n}$, and this is precisely the space of phylogenetic trees discussed previously.

We now fix an arbitrary point X with coordinates $(X_{i_1 \cdots i_d})$ in the Dressian $Dr_{d,n}$. For any $(d+1)$-subset $\{j_0, j_1, \ldots, j_d\}$ of $\{1, 2, \ldots, n\}$ we consider the following tropical linear form in the variables $x_1, \ldots, x_n$:

$$\ell^X_{j_0 j_1 \cdots j_d} = \bigoplus_{r=0}^{d} X_{j_0 \cdots \widehat{j_r} \cdots j_d} \odot x_r, \tag{6}$$

where the $\widehat{\ }$ means to omit j_r. The *tropical linear space* associated with the point X is the following set:

$$L_X = \bigcap \mathcal{H}(\ell^X_{j_0 j_1 \cdots j_n}) \subset \mathbb{R}^n.$$

Here the intersection is over all $(d+1)$-subsets $\{j_0, j_1, \ldots, j_d\}$ of $\{1, 2, \ldots, n\}$.

The tropical linear spaces are precisely the sets L_X where X is any point in $Dr_{d,n} \subset \mathbb{R}^{\binom{n}{d}}$. These objects are studied in detail in [**21**] and [**11**]. The "sufficient completeness" referred to in the first paragraph of this section means that we need to solve linear equations using the above formula for L_X, in order for an intersection of hyperplanes actually to be a linear space. The definition of linear space given here is more inclusive than the one used elsewhere [**6, 17, 20**], where L_X was required to come from ordinary algebraic geometry over a field with a suitable valuation.

For example, a 3-dimensional tropical linear subspace of $\mathbb{R}^n$ (a.k.a. a two-dimensional plane in tropical projective $(n-1)$-space) is the intersection of $\binom{n}{4}$ tropical hyperplanes, each of whose defining linear forms has four terms:

$$\ell^X_{j_0 j_1 j_2 j_3} = X_{j_0 j_1 j_2} \odot x_{j_3} \oplus X_{j_0 j_1 j_3} \odot x_{j_2} \oplus X_{j_0 j_2 j_3} \odot x_{j_1} \oplus X_{j_1 j_2 j_3} \odot x_{j_0}.$$

We note that even the very special case when each coordinate of X is either 0 (the multiplicative unit) or ∞ (the additive unit) is really interesting. Here L_X is a polyhedral fan known as the *Bergman fan* of a matroid [**1**].

Tropical linear spaces have many of the properties of ordinary linear spaces. First, they are pure polyhedral complexes of the correct dimension:

FACT 7. *Each maximal cell of the tropical linear space L_X is d-dimensional.*

Every tropical linear space L_X determines its vector of tropical Plücker coordinates X uniquely up to tropical multiplication (= classical addition) by a common scalar. If L and L' are tropical linear spaces of dimensions d and d' with $d + d' \geq n$, then L and L' meet. It is not quite true that two tropical linear spaces intersect in a tropical linear space but it is almost true. If L and L' are tropical linear spaces of dimensions d and d' with $d + d' \geq n$ and v is a generic small vector then $L \cap (L' + v)$ is a tropical linear space of dimension $d + d' - n$. Following [**17**], it makes sense to define the *stable intersection* of L and L' by taking the limit of $L \cap (L' + v)$ as v goes to zero, and this limit will again be a tropical linear space of dimension $d + d' - n$.

It is not true that a d-dimensional tropical linear space can always be written as the intersection of $n - d$ tropical hyperplanes. The definition shows that $\binom{n}{d+1}$ hyperplanes are always enough. At this point in the original 2004 lecture notes, we had asked: What is the minimum number of tropical hyperplanes needed to cut out any tropical linear space of dimension d in n-space? Are n hyperplanes always enough? These questions were answered by Tristram Bogart in [**2**, Theorem 2.10], and a more refined combinatorial analysis was given by Josephine Yu and Debbie Yuster in [**22**]. Instead of posing a new research problem, we end this article with a question.

Are there any textbooks on tropical geometry? As of June 2009, there seem to be no introductory texts on tropical geometry, despite the elementary nature of the basic definitions. The only book published so far on tropical algebraic geometry is the volume [**10**] which is based on an Oberwolfach seminar held in 2004 by Ilia Itenberg, Grigory Mikhalkin, and Eugenii Shustin. That book emphasizes connections to topology and real algebraic geometry. Several expository articles offer different points of entry. In addition to [**17**], we especially recommend the expositions by Andreas Gathmann [**9**] and Eric Katz [**12**]. These are aimed at readers who have a background in algebraic geometry. Grigory Mikhalkin is currently writing a research monograph on tropical geometry for the book series of the Clay Mathematical Institute, while Diane Maclagan and the second author have begun a book project titled *Introduction to Tropical Geometry*. Preliminary manuscripts can be downloaded from the authors' homepages. In fall 2009, the Mathematical Sciences Research Institute (MSRI) in Berkeley will hold a special semester on Tropical Geometry.

Acknowledgment. Speyer was supported in this work by a Clay Mathematics Institute Research Fellowship, Sturmfels by the National Science Foundation (DMS-0456960, DMS-0757236).

REFERENCES

1. F. Ardila and C. Klivans, The Bergman complex of a matroid and phylogenetic trees, *J. Combinatorial Theory Ser. B* **96** (2006) 38–49.
2. T. Bogart, A. Jensen, D. Speyer, B. Sturmfels, and R. Thomas, Computing tropical varieties, *J. Symbolic Computation* **42** (2007) 54–73.
3. L. Billera, S. Holmes, and K. Vogtman, Geometry of the space of phylogenetic trees, *Advances in Applied Math.* **27** (2001) 733–767.
4. C. Bocci and F. Cools, A tropical interpretation of m-dissimilarity maps, preprint, `arXiv:0803.2184`.
5. P. Butkovič, Max-algebra: the linear algebra of combinatorics? *Linear Algebra Appl.* **367** (2003) 313–335.
6. M. Develin, F. Santos, and B. Sturmfels, On the rank of a tropical matrix, pages 213–242 in *Combinatorial and Computational Geometry*, Mathematical Sciences Research Institute Publication, Vol. 52, Cambridge Univ. Press, Cambridge, 2005.
7. J. Felsenstein, *Inferring Phylogenies*, Sinauer Associates, Sunderland, MA, 2003.
8. S. Gao and A. Lauder, Decomposition of polytopes and polynomials, *Discrete and Computational Geometry* **26** (2001) 89–104.
9. A. Gathmann, Tropical algebraic geometry, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **108** (2006) 3–32.
10. I. Itenberg, G. Mikhalkin, and E. Shustin, *Tropical Algebraic Geometry*, Oberwolfach Seminars Series, Vol. 35, Birkhäuser, Basel, 2007.
11. S. Hermann, A. Jensen, M. Joswig, and B. Sturmfels, How to draw tropical planes, preprint, `arxiv:0808.2383`.
12. E. Katz, A tropical toolkit, to appear in *Expositiones Mathematicae*, `arXiv:math/0610878`.
13. G. Mikhalkin, Enumerative tropical geometry in $\mathbb{R}^2$, *J. Amer. Math. Soc.* **18** (2005) 313–377.
14. L. Pachter and B. Sturmfels, *Algebraic Statistics for Computational Biology*, Cambridge Univ. Press, Cambridge, 2005.
15. L. Pachter and D. Speyer, Reconstructing trees from subtree weights, *Appl. Math. Lett.* **17** (2004) 615–621.
16. J.-E. Pin, Tropical semirings, *Idempotency* (Bristol, 1994), 50–69, Publ. Newton Inst., Vol. 11, Cambridge Univ. Press, Cambridge, 1998.
17. J. Richter-Gebert, B. Sturmfels, and T. Theobald, First steps in tropical geometry, pages 289–317 in *Idempotent Mathematics and Mathematical Physics*, Contemporary Mathematics, Vol. 377, American Mathematical Society, Providence, RI, 2005.
18. C. Semple and M. Steel, *Phylogenetics*, Oxford University Press, Oxford, 2003.
19. I. Simon, Recognizable sets with multiplicities in the tropical semiring, pages 107–120 in *Mathematical Foundations of Computer Science* (Carlsbad, 1988), Lecture Notes in Computer Science, Vol. 324, Springer, Berlin, 1988.
20. D. Speyer and B. Sturmfels, The tropical Grassmannian, *Advances in Geometry* **4** (2004) 389–411.
21. D. Speyer, Tropical linear spaces, *SIAM J. Disc. Math.* **22** (2008) 1527–1558.
22. J. Yu and D. Yuster, Representing tropical linear spaces by circuits, in *Formal Power Series and Algebraic Combinatorics (FPSAC '07)*, Proceedings, Tianjin, China, 2007.