Sturm und Train

Danny Calegari

When I'm seized by emotional turmoil arising from a rejection of neoclassical literary norms, I like to go for a run. I run along the shore of Lake Michigan, but I'm out of shape, so I need a psychological crutch to keep running: the markers they have every mile by which I can measure my progress. When I'm done I can round up to estimate my running distance. Yesterday I ran two miles!

Using discrete numbers to measure a continuous variable is nothing new. I round up my height (five foot nine) and I round down my age and weight (forty-something and none of your business). And we sometimes go back from discrete to continuous. In dynamics there's the *rotation number* of an orientation-preserving homeomorphism of a circle. Let's let $\phi : S^1 \rightarrow S^1$ be a homeomorphism. Since I'm a topologist, I'm going to identify my circle with the perimeter of Lake Michigan, and my apartment will correspond to some base point p in S^1 . I start at p and run (in the positive direction) from p to $\phi(p)$. Then I run (also in the positive direction) from $\phi(p)$ to $\phi^2(p)$, then from $\phi^2(p)$ to $\phi^3(p)$, and so on. I can't tell exactly how far I've run at each stage, but I can count the number of laps—i.e., complete circuits around the lake or, equivalently, the circle—by how many times I pass by my apartment building. The rotation number is the limit

$$\text{rot}(\phi) := \lim_{m \rightarrow \infty} \frac{1}{m} \cdot \text{number of laps from } p \text{ to } \phi^m(p).$$

Danny Calegari is a professor of mathematics at the University of Chicago. His email address is dannyc@math.uchicago.edu.

For permission to reprint this article, please contact:
reprint-permission@ams.org.

DOI: <https://doi.org/10.1090/noti1955>

Running laps around Lake Michigan is just like finding successive zeros of solutions of a differential equation. No, really. Let's take a homogeneous second-order linear differential equation

$$f''(t) + a(t)f'(t) + b(t)f(t) = 0$$

and let's suppose $f_1(t)$ and $f_2(t)$ are a pair of linearly independent solutions. The *Sturm Comparison Theorem* says that between any two successive zeros t_0 and t_1 of f_1 there is *exactly* one zero of f_2 . Here's the proof. Since the equation is second order, a solution f is uniquely determined by the values of $f(t)$ and $f'(t)$ at any fixed t . In particular for a nontrivial solution like f_1 , the pair $(f_1(t), f_1'(t))$ is a *nonzero* vector in $\mathbb{R}^2$ for all t , and so is $(f_2(t), f_2'(t))$. The slopes of these vectors $\sigma_1(t)$, $\sigma_2(t)$ at time t give a pair of points on the "circle" $\mathbb{RP}^1$, and t is a zero of f_j when the slope $\sigma_j(t)$ is infinity. As t increases, the slopes $\sigma_1(t)$ and $\sigma_2(t)$ move around the circle. They can change direction sometimes, but the slope can only cross through infinity in the clockwise direction. The *Wronskian* $W(f_1, f_2)(t)$ is the determinant $f_1(t)f_2'(t) - f_1'(t)f_2(t)$. It satisfies the equation $W'(t) = -a(t)W(t)$, so it can never change sign. This means that $\sigma_1(t)$ and $\sigma_2(t)$ are never equal. As t increases from t_0 to t_1 , the slope σ_1 runs once around the circle, and therefore in this time interval σ_2 passes my apartment building exactly once, qed.

As with most things in life, this story turns out to be a manifestation of symplectic geometry. Let's change variables by an integrating factor $g(t) := p(t)f(t)$. Then $g' = p'f + pf'$ and

$$g'' = p''f + 2p'f' + pf'' = (p'' - bp)f + (2p' - ap)f'.$$

Setting $p(t) = e^{\int a(t)/2}$ gives $p' = ap/2$, so

$$g''(t) + c(t)g(t) = 0$$

with $c := (bp - p'')/p$.

Notice that f is zero precisely when g is. The space of solutions to our new differential equation for g is a real vector space of dimension 2. As remarked above, the values of a solution $(g(t), g'(t))$ at any fixed t give coordinates on this vector space. Transforming the coordinates at time 0 to the coordinates at time t gives a 1-parameter family of matrices $M(t)$ satisfying

$$\begin{pmatrix} g(t) \\ g'(t) \end{pmatrix} = M(t) \cdot \begin{pmatrix} g(0) \\ g'(0) \end{pmatrix}$$

for all solutions g . In particular, $M'(t) = \begin{pmatrix} 0 & 1 \\ -c(t) & 0 \end{pmatrix} M(t)$. Since $M'M^{-1}$ has trace zero it follows that M always has determinant 1: it preserves the *area form* on $\mathbb{R}^2$ (this is equivalent to saying that the Wronskian of two solutions is constant). This suggests we should think of $\mathbb{R}^2$ as $T^*\mathbb{R}$ with its tautological symplectic form and the spans of $(g(t), dg(t))$ as a 1-parameter family of Lagrangian subspaces (i.e., lines) moving by the symplectic (i.e., area-preserving) 2×2 matrices $M(t)$.

What's the analog in higher dimensions? Let's suppose we have a vector equation $g''(t) = -C(t)g(t)$ for $g : \mathbb{R} \rightarrow \mathbb{R}^n$ where $C(t)$ is a symmetric matrix. This could arise e.g. as a system of Newton's equations where $g(t) \in \mathbb{R}^n$ is the "position" of the system at time t , and $\frac{1}{2}\langle C(t)g(t), g(t) \rangle$ is the associated potential energy. The pair $(g(t), dg(t))$ is a point in the symplectic vector space $T^*\mathbb{R}^n$ and evolves in time as the orbit of the action of a family of symplectic $2n \times 2n$ matrices $M(t)$.

The right thing to consider in this context are not lines but *Lagrangian subspaces*, i.e., linear subspaces of real dimension n on which the symplectic form (a nondegenerate antisymmetric quadratic form) vanishes identically. The analog of "slope infinity" is a Lagrangian subspace σ that contains a vertical vector, i.e., a nonzero vector with zero projection to $\mathbb{R}^n$. By abuse of notation we call such Lagrangian subspaces *vertical*.

If q is a smooth real-valued function on $\mathbb{R}^n$, then dq , thought of as a section of $T^*\mathbb{R}^n$, has tangents everywhere parallel to a Lagrangian subspace. Every Lagrangian subspace of $T^*\mathbb{R}^n$ without a vertical vector arises this way as the graph of dq for some homogeneous quadratic form $q : \mathbb{R}^n \rightarrow \mathbb{R}$.

Another way to define Lagrangian subspaces is as follows. On a complex vector space the imaginary part of a Hermitian inner product is a symplectic form. So if we identify $\mathbb{C}^n$ with symplectic $\mathbb{R}^{2n}$, the Lagrangian subspaces of $\mathbb{R}^{2n}$ correspond to the maximal real subspaces of $\mathbb{C}^n$ on which the restriction of the Hermitian inner product takes only real values.

The analog of the Sturm Comparison Theorem is the following: if, under the evolution of $M(t)$ some Lagrangian subspace becomes vertical $n + 1$ times, every other Lagrangian subspace becomes vertical at least once. And in fact the absolute value of the difference of the number of times two subspaces become vertical on any time interval $[t_0, t_1]$ is at most n .

There's a geometric proof very similar in substance to the proof of the ordinary Sturm Theorem given above. The trick is to find the higher dimensional analog of the perimeter of Lake Michigan and of the concept of "passing by my apartment building."

Let Λ_n denote the space of Lagrangian subspaces of symplectic $\mathbb{R}^{2n}$. This is a homogeneous space; thinking of a symplectic form as the imaginary part of a Hermitian inner product identifies Λ_n with the coset space $U(n)/O(n)$. So Λ_1 is $S^1/\pm 1 = \mathbb{RP}^1$ from before. The space Λ_2 is more complicated: it's homeomorphic to the twisted S^2 bundle over S^1 , where the monodromy around the circle takes S^2 to itself by the antipodal map. In general Λ_n is a closed manifold of real dimension $n(n+1)/2$. Thinking of it as a coset space, taking determinants gives Λ_n the structure of a fibration over the circle

$$\det : U(n)/O(n) \rightarrow S^1/\pm 1,$$

and we can try to figure out what it means to say that a family of Lagrangian subspaces "winds around the circle." But when we run around an honest topological circle the order of two runners makes topological sense; what's the analog of this order structure for two Lagrangian subspaces of $\mathbb{R}^{2n}$?

If $\sigma \subset \mathbb{R}^{2n}$ is a Lagrangian subspace, Arnol'd defines the *train* of σ , denoted $\mathcal{V}(\sigma)$, to be the variety of Lagrangian subspaces τ that intersect σ in a subspace of positive dimension. Since Lagrangian subspaces each have dimension n , the train is exactly the complement of the open subset of Lagrangian subspaces τ that intersect σ in general position. If we identify $\mathbb{R}^{2n}$ with $T^*\sigma$, then in a neighborhood of σ , nearby Lagrangian subspaces are all equal to the graph of dq for some homogeneous quadratic form q on σ . The train $\mathcal{V}(\sigma)$ cuts up the neighborhood of σ into "chambers" distinguished from one another by the number of positive and negative eigenvalues of q . A family of Lagrangians $\sigma(t)$ winds "positively" in Λ_n if $\sigma'(t)$ always points into the chamber of positive definite forms on $\sigma(t)$. This notion of positivity is invariant under symplectic transformations of Λ_n . If we identify $\mathbb{R}^{2n}$ with $T^*\mathbb{R}^n$, then the "completely vertical" fiber $V := T_0^*\mathbb{R}^n$ is a Lagrangian subspace "dual" to the zero section, and a Lagrangian subspace is vertical in the sense we defined above if and only if it is in $\mathcal{V}(V)$, the train of V . From now on we'll just write the vertical train as $\mathcal{V}$.

Incidentally, this symplectic winding number of a fam-

ity of Lagrangian subspaces is related to something called the *Maslov index*. An n -dimensional submanifold Y of $\mathbb{R}^{2n}$ is Lagrangian if its tangent space is everywhere parallel to a Lagrangian subspace. There is a “Gauss map” $Y \rightarrow \Lambda_n$. The Maslov index $\mu \in H^1(Y; \mathbb{Z})$ is the pullback of the generator of $H^1(\mathbb{R}\mathbb{P}^1; \mathbb{Z}) = \mathbb{Z}$ under the composition $Y \rightarrow \Lambda_n \rightarrow \mathbb{R}\mathbb{P}^1$.

There’s an analog of rotation number too. If ϕ is a symplectic matrix we can define the *symplectic rotation number* of ϕ by picking a Lagrangian subspace $\sigma \in \Lambda_n$ and a positively winding path γ from σ to $\phi(\sigma)$ and for each m defining γ_m to be the path from σ to $\phi^m(\sigma)$ obtained by concatenating the paths $\gamma, \phi(\gamma), \dots, \phi^{m-1}(\gamma)$. Then we can define $\text{rot}(\phi)$ to be the limit

$$\text{rot}(\phi) := \lim_{m \rightarrow \infty} \frac{1}{m} \cdot \text{number of times the path } \gamma_m \text{ from } \sigma \text{ to } \phi^m(\sigma) \text{ crosses } \mathcal{V}.$$

Now let’s return to the proof of the higher dimensional Sturm Comparison Theorem. Let $\sigma_j(t)$ for $j = 1, 2$ be two families of Lagrangian subspaces of the form $M(t)\sigma_j(0)$ with $M(t)$ symplectic arising as above. For $j = 1, 2$ and for any t define

$$n_j(t) := \sum_{0 \leq s \leq t} \text{dimension of vertical subspace of } \sigma_j(s).$$

A vector $(g(t), g'(t))$ in $\sigma_j(t)$ is vertical when $g(t) = 0$. As t increases, $g(t)$ becomes positive or negative according to the sign of $g'(t)$. Thus when $\sigma_j(t)$ passes through the train of V this intersection is nondegenerate and always has positive intersection number (although Λ_n is not typically orientable, every train is “coorientable” because we can compute the number of positive eigenvalues of a quadratic form on a vector space without choosing an orientation). Winding once around Λ_n induces n intersections with the train $\mathcal{V}$, counting with multiplicity as in the definition of $n_j(t)$. This is because we pass once through each of the $n + 1$ chambers of Λ_n cut out by $\mathcal{V}$.

For any two initial values $\sigma_1(0), \sigma_2(0)$ we can find a loop γ containing both of them that winds exactly once around Λ_n and for which γ' is always “positive.” The loop γ breaks up into two positive arcs: α from $\sigma_1(0)$ to $\sigma_2(0)$ and β from $\sigma_2(0)$ to $\sigma_1(0)$. Applying $M(t)$ gives us a 1-parameter family of positive loops γ_t containing $\sigma_1(t)$ and $\sigma_2(t)$, which decompose into positive arcs α_t and β_t . For any t the two arcs

$$(\alpha_0 \text{ composed with } \sigma_2[0, t]) \quad \text{and} \\ (\sigma_1[0, t] \text{ composed with } \beta_t)$$

are homotopic rel. endpoints, so their algebraic intersections with the train are equal. On the other hand, since every intersection of each arc with the train is positive, their *geometric* intersections with the train are equal. Since α_0 is

contained in γ_0 and β_t is contained in γ_t , they each have at most n intersections with the train, all positive. So

$$|\alpha_0 \cap \text{train}| + n_2(t) = |\beta_t \cap \text{train}| + n_1(t),$$

and $|n_1(t) - n_2(t)| \leq n$.

AUTHOR’S NOTE. This generalization of the Sturm Comparison Theorem and its proof and the point of view that proves it are all due to Vladimir Arnol’d and explained in his paper “Sturm theorems and symplectic geometry,” *Func. Anal. Appl.* 19 (1985), no. 4, 251–259. MR0820079

Danny Calegari

Credits

Author photo is courtesy of the author.

CIMAT

MATHEMATICAL SCIENCES SEMESTERS
IN GUANAJUATO, MEXICO

An innovative study abroad program
combining **mathematics, statistics, data science, and computer science.**

The Spring Semester in Mathematical Tools for Data Science offers courses on:

- Design of Algorithms
- Discrete Mathematics
- Probability and Statistics
- Numerical Optimization and Machine Learning
- Geometric and Graphing Tools

Application
NOW OPEN
mathsciencesgto.cimat.mx