Aspects of the Monster Tower Construction: Geometric, Combinatorial, Mechanical, Enumerative

Susan Colley Gary Kennedy

November 10, 2022

These are notes for a 4-lecture series we delivered at the Workshop on Nash Blow-up and Semple Tower, II, June 3–7, 2019, at KU Leuven in Belgium. We also prepared four individual sets of lecture slides. Related lectures were delivered by Ferran Dachs Cadefau and Alejandro Bravo-Doddoli.

Introduction

In these lectures we survey the construction and use of the monster tower (also known as the Semple tower) in three distinct areas of mathematics.

Lecture 1: The Monster Tower (Kennedy) — This lecture will explain how three seemingly different situations lead to the same construction:

1. Compactifying curvilinear data (algebraic geometry)
2. Studying Goursat distributions (differential geometry)
3. Analyzing a truck with trailers (mechanics and control theory)

Lecture 2: Combinatorial Aspects (Colley) — We explain a natural system of coordinate charts on the monster space. We show how to lift (prolong) a curve in the base into the tower. We explain a natural coarse stratification of the monster, catalogued by a simple system of code words.

Lecture 3: Mechanical Aspects (Kennedy) — A version of the monster tower construction creates the natural configuration space for a truck with trailers. We explain the model and survey some important features, including Lie brackets of its basic vector fields, its singular configurations, and its dynamics.

Lecture 4: Enumerative Aspects (Colley) — We begin with an introduction (via examples) to the subjects of enumerative geometry and intersection theory. Specializing to the enumeration of contacts between plane curves, we present a strategy for counting such contacts and illustrate it with a quadruple contact formula we once proved. The ideas behind this formula lead to a discussion of the orbits of the monster space, and to the idea of appropriately lifting a family of curves.

Lecture 1: The monster tower

This lecture will explain how three seemingly different situations lead to the same construction:

1. Compactifying curvilinear data (algebraic geometry)
2. Studying Goursat distributions (differential geometry)
3. Analyzing a truck with trailers (mechanics and control theory)

Compactifying curvilinear data (algebraic geometry)

Let's start with a smooth manifold or complex manifold or smooth algebraic variety M of dimension $m > 1$. Suppose we have two smooth curves C_1 and C_2 passing through a point, and that we have a system of local coordinates $x_1, \ldots, x_m$ based there, so that, for each curve, dx_1 doesn't vanish at the point. We say that the curves *have the same curvilinear data* up to order k at the point if the values of all derivatives $d^j x_i/(dx_1)^j$ agree up to order k.

Figure 1 shows the parabola $y = \frac{1}{2}x^2 - 1$ and the unit circle. Working at the point $(0, -1)$, for the parabola we see that

$$\left(y, y', y'', y^{(3)}, y^{(4)}\right) = (-1, 0, 1, 0, 0)\,,$$

whereas for the circle we have

$$\left(y, y', y'', y^{(3)}, y^{(4)}\right) = (-1, 0, 1, 0, 3)\,.$$

Thus the two curves have the same curvilinear data up to order 3. Rotating both curves and now looking at the point $\left(\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}\right)$, we find that

$$\left(y, y', y'', y^{(3)}, y^{(4)}\right) = \left(\frac{1}{2}\sqrt{2},\ -1, -2\sqrt{2},\ -12,\ -60\sqrt{2}\right)$$

for the parabola, and

$$\left(y, y', y'', y^{(3)}, y^{(4)}\right) = \left(\frac{1}{2}\sqrt{2},\ -1,\ -2\sqrt{2},\ -12,\ -72\sqrt{2}\right)$$

for the circle.

Agreeing in curvilinear data up to order k is $(m-1)k$ conditions, and one can check that they are independent of local coordinates. One can see that a *curvilinear datum* is a point in a manifold (or smooth variety) of dimension $m + (m-1)k$, and that in fact we have a tower of such manifolds, with M as the base, with fiber a projective space $\mathbb{P}^{m-1}$ at the first stage, and then affine space fibers $\mathbb{A}^{m-1}$ thereafter.

Furthermore, it makes sense to *lift* our curves up through the tower. In other words, to each point p on the smooth curve C we can associate the point

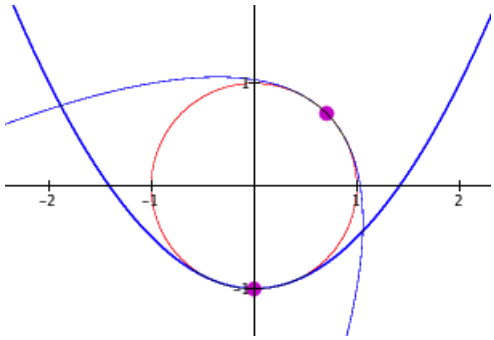


Figure 1: A parabola and its osculating circle, before and after rotation.

recording its curvilinear data (up to whatever order we like), and the locus of these points is again a smooth curve.

We have traced this idea back as far as the work of Halphen. In his doctoral thesis [20] and in a 94-page paper [19], he wrote about what we would now call the parameter space for curvilinear data in the plane.

Here are two things we'd like to do:

1. Give a coordinate-free version of the construction.

2. Compactify the spaces in the tower, in some natural way.

The first-order datum is giving us a tangent direction, so at the first stage of the construction we have $M(1) = \mathbb{P}TM$, the total space of the projectivized tangent bundle. The lift of a curve is its *Nash blowup*.

Naively, one might guess that one just continues in this way: define $M(2) = \mathbb{P}TM(1)$, etc., but this can't be right, because it doesn't give us a manifold of the right dimension. The dimension of $M(2)$ ought to be $3m - 2$, but the dimension of $\mathbb{P}TM(1)$ is $4m - 3$. To say this another way, most points of the proposed space are inaccessible: they can't be reached by lifting (twice) a curve from the base. There must be something special about the tangent lines to (once) lifted curves — what is it?

Semple [35] supplied the answer. To explain what he did, we'll work in a somewhat larger context, and explain what we call the *basic construction*.

The basic construction

Suppose that M is a smooth manifold or nonsingular algebraic variety over an algebraically closed field of characteristic 0. Suppose that $\mathcal{B}$ is a rank b subbundle of its tangent bundle TM. Let $\widetilde{M} = \mathbb{P}\mathcal{B}$, the total space of the projectivization of the bundle, and let $\pi : \widetilde{M} \to M$ be the projection. A point $\widetilde{p}$ of $\widetilde{M} = \mathbb{P}\mathcal{B}$ over $p \in M$ represents a line inside the fiber of $\mathcal{B}$ at p, and since $\mathcal{B}$ is a subbundle of TM, this is a *tangent direction* to M at p. Let

$$d\pi : T\widetilde{M} \to \pi^* TM$$

denote the derivative map of π. A tangent vector to $\widetilde{M}$ at $\widetilde{p}$ is said to be a *focal vector* if it is mapped by $d\pi$ to a tangent vector at p in the direction represented by $\widetilde{p}$; in particular a vector mapping to the zero vector (called a *vertical vector*) is considered to be a focal vector. The subspace of focal vectors is called the *focal space*. (Figure 2 illustrates the construction, assuming that $b = 2$.) The set of all focal vectors forms a subbundle $\widetilde{\mathcal{B}}$ of $T\widetilde{M}$, called the *focal bundle*; its rank is again b. Thus we can iterate this construction to obtain a tower of spaces (i.e., smooth manifolds or nonsingular algebraic varieties) together with their associated bundles.

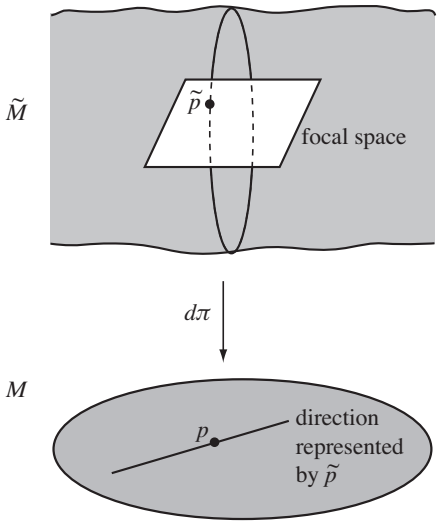


Figure 2: The focal space.

If we begin the construction by taking $\mathcal{B}$ to be the tangent bundle TM itself, then the resulting tower

$$\cdots \to M(k) \xrightarrow{\pi_k} M(k-1) \xrightarrow{\pi_{k-1}} \cdots \to M(2) \to M(1) \to M(0) = M$$

is called the *Semple tower* over the base M. The space $M(k)$ is said to be at *level* k. Observe that it is the total space of a $\mathbb{P}^{m-1}$-bundle over $M(k-1)$; in particular, $M(1)$ is the total space of the projectivized tangent bundle $\mathbb{P}TM$. The bundle constructed at step k of the construction is called the kth *focal bundle* and denoted Δ_k; it is a subbundle of the tangent bundle $TM(k)$.

The earliest instance of the construction seems to occur in Gherardelli's 1941 paper [18], which constructed $\mathbb{P}^2(2)$ over the projective plane $\mathbb{P}^2$. The full tower was explained by Semple in [35]. We learned about the construction from remarks of Collino in [11], and used it to study problems of enumerative geometry in [7], [8], and [9]. It was treated in greater generality by Lejeune-Jalabert [25]. Demailly used it to study positivity questions for hyperbolic varieties in [12].

Studying Goursat distributions (differential geometry)

Now let's make a fresh start. (Our main source for this exposition is [28], and we will adopt its notations.)

A *distribution* $\mathcal{D}$ on a smooth manifold is a subbundle of its tangent bundle. We want to consider *germs* of distributions at a point; thus we may work at the origin in Euclidean space $\mathbb{R}^n$. Two distributions represent the same germ if they agree in some neighborhood of the origin. Let d be the rank of $\mathcal{D}$; its *type* is the pair (d, n); the number $s := n - d$ is called the *corank*. (To match these notations with those in our basic construction, replace (d, n) by (b, m).)

There is a natural topology on the set of germs. There is also a natural notion of equivalence of germs: local diffeomorphism taking one distribution to the other.

How can you tell when two germs are equivalent? Let's look at the well-known theorem of Frobenius:

A distribution is completely integrable if and only if it is involutive.

(See, e.g. [1, pp. 95–96].)

We need to explain the terminology in this statement. The distribution is said to be *completely integrable* if through each point we can find a submanifold whose tangent space at that is the fiber of $\mathcal{D}$. Equivalently, we can find local coordinates $x_1, x_2, \ldots, x_n$ so that at each point $\mathcal{D}$ is spanned by $\partial/\partial x_1, \partial/\partial x_2, \ldots, \partial/\partial x_d$. To understand the meaning of "involutive," we need the notion of *Lie bracket of vector fields*. Suppose that X and Y are vector fields. We can think of them as operators on functions, and thus compose them. We define

$$[X, Y] = XY - YX,$$

again an operator on functions, and one can prove that it is again a vector field. The distribution $\mathcal{D}$ is said to be *involutive* if the Lie bracket of two arbitrary vector fields in $\mathcal{D}$ is again in $\mathcal{D}$.

Now differential geometers and control theorists are interested in noninvolutive distributions. (In Lecture 3 we'll say something about why.) It is therefore natural to consider the *Lie square*

$$\mathcal{D}^2 = \mathcal{D} + [\mathcal{D}, \mathcal{D}]$$

where the second term consists of all possible Lie brackets of vector fields in $\mathcal{D}$. In general this won't be a distribution, since the rank may vary from point to point, but let's suppose that it is, and let's iterate the construction to get a sequence of subbundles (as we assume) of the tangent bundle

$$\mathcal{D} = \mathcal{D}_1 \subset \mathcal{D}_2 \subset \mathcal{D}_3 \subset \cdots$$

where each bundle is the Lie square of the previous bundle. Call this the *Lie squares sequence*.

The bundle is said to be *Goursat* if we have this mild growth: the rank increases by one at each step, until we reach the tangent bundle.

Engel [14] studied Goursat distributions of type $(2,4)$. Here the sequence of ranks is $2,3,4$. He discovered that all such distributions are equivalent. We now call this sole distribution the *Engel distribution*.

Building on Engel's work, Cartan [3] showed how to *prolong* the Engel distribution to obtain Goursat distributions, and he showed that for a generic prolongation there is a single normal form, i.e., a single equivalence class of Goursat distributions. In [28], Montgomery and Zhitomirskii write: "Some researchers believe that Cartan missed the singularities [the nongeneric prolongations] in the problem of classifying Goursat distributions. It would be more accurate to say that he was not interested in them."

So what is this prolongation procedure? In the case $d = 2$, we've already seen it above. If we have a rank 2 Goursat distribution $\mathcal{D}$ on a manifold M, then we're in the situation of the basic construction, and thus we obtain $\widetilde{M} := \mathbb{P}\mathcal{D}$, a $\mathbb{P}^1$-bundle over M, together with a new rank 2 distribution $\widetilde{\mathcal{D}}$. (Montgomery and Zhitomirskii use the notations PM for $\widetilde{M}$ and $\mathcal{E}$ for $\widetilde{\mathcal{D}}$.) The inverse images of the bundles in the Lie squares sequence all increase in rank by one, and $\widetilde{\mathcal{D}}$ fits in at the beginning:

$$\widetilde{\mathcal{D}} \subset \pi^{-1}\mathcal{D} \subset \pi^{-1}\mathcal{D}_2 \subset \pi^{-1}\mathcal{D}_3 \subset \cdots .$$

Thus one observes that again $\widetilde{\mathcal{D}}$ is Goursat. By iterating the construction, one obtains a tower of $\mathbb{P}^1$-bundles over M. Montgomery and Zhitomirskii called it the *monster tower*.

(Their construction doesn't agree with our basic construction if $d > 2$, but we don't need to talk about that case.)

Now here's what Montgomery and Zhitomirskii did that's really clever: they reversed the construction. Suppose $\mathcal{D}$ is a germ of a rank 2 Goursat distribution of corank $s > 2$. They explain how to *deprolong* it to obtain a rank 2 Goursat distribution. We won't go into the details of this construction, but if you look into their cited paper, the key observation is what they call the Sandwich Lemma. From it, they show how to identify a rank 1 subbundle that ought to be collapsed, so that $\mathcal{D}$ collapses down to a rank 2 distribution of corank $s-1$.

From this one obtains a beautiful understanding of Goursat distributions of rank 2, as follows: given a germ Goursat distribution of rank 2, deprolong repeatedly until you reach corank 2. You now have the Engel distribution! Thus your given germ can be found somewhere in the repeated prolongation of the Engel distribution. In fact one can take this even further, by observing that the Engel distribution can be obtained by twice prolonging the tangent bundle to a surface.

All of this appears in their 2001 paper [28] and the subsequent monograph [29], but it took several years for someone to realize that the two tower constructions — Semple and monster — are the same, in the case where the base is a (real) surface. That person was Montgomery's student Alex Castro, and we think the connection was first written down in his 2010 dissertation at UC Santa Cruz [5]. In light of this discovery, the "monster" terminology has now been adopted in greater generality, as a synonym for "Semple." For example,

the space $M(k)$ encountered earlier is now called the *kth monster*.

You might observe that the construction in algebraic geometry is "bottom-up": in trying to analyze curves on a manifold, we are led to construct a tower over it. By contrast, if you are given a Goursat distribution, you are already somewhere up in the tower, and the Montgomery–Zhitomirskii construction shows you how to move downward.

Analyzing a truck with trailers (dynamics and control theory)

Introductory talks about control theory often mention a truck towing a trailer. The truck is free to move in the plane; it can be driven forward or backward, braked, and steered. We model it as a point with an attached vector, which shows which way it is pointing. (See Figure 3.) The truck can be moved following any differentiable path, subject to the following condition: the velocity vector must point in the direction that the truck is pointing. If the velocity vector is zero, this is interpreted as a vacuous condition (automatically satisfied). It is thus legal to drive the truck simply by turning, i.e., by keeping its position fixed while changing which way it's pointing.

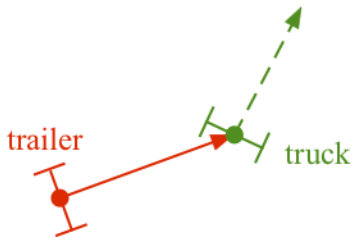


Figure 3: Truck and trailer.

The trailer is also modeled as a point with attached unit vector, and its motion must satisfy two constraints:

1. It must be one unit of distance from the truck.
2. Its velocity must be in the direction of truck. (Again the zero velocity vector always satisfies this condition.)

In the literature of dynamics, control theory, and engineering, the first condition is called a *holonomic constraint*: it's a condition on positions, and it will be incorporated into our choice of configuration space. The second condition, which restricts the allowable velocities, is called a *nonholonomic constraint*.

This does seem to fit the physical situation of a *tractor-trailer* or *big rig*, consisting of a relatively compact cab and a long trailer whose wheels can't be turned; think of the point as the rear of the trailer. It's also a good model of an automobile or a bicycle [15].

The *configuration space* consists of quadruples

$$(x, y, \varphi_0, \varphi_1),$$

where (x, y) is the position of the trailer, φ_0 records the direction of the unit vector pointing from the trailer toward the truck (using the positive x-direction as the reference direction), and φ_1 records the *bending angle* between the direction of the trailer and the direction of the truck. (See Figure 4.) Thus the configuration space is $\mathbb{R}^2 \times S^1 \times S^1$. (Notice that the position of the truck isn't made explicit, but obviously can be inferred.)

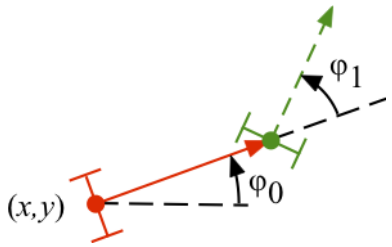


Figure 4: Coordinates for the truck-and-trailer configuration space.

The physical situation suggests, and the mathematical model confirms, that the path of the truck determines the path of the trailer. In other words, the nonholonomic constraint gives us a differential equation which, given the path of the truck, can be solved to determine the path of the trailer. Here are two examples:

- If the truck and trailer are aligned and we drive the truck forward, the trailer follows in a straight line.
- We can drive the truck in a unit circle (using the same unit as for the length of the trailer), so that the trailer stays fixed at a point.

Control theory is concerned with the opposite direction: if you want the trailer to follow a certain path, how should you drive the truck? The answer is, of course, a process involving differentiation. In fact, there's a nice geometric description: suppose we specify a differentiable path for the trailer. At each point, draw the unit tangent vector, and then mark the head of this vector. The heads of all these vectors trace out the required path of the truck. This process is illustrated in Figure 5.

As an example, suppose you want to move the trailer along a circle of radius r. Then the truck should be moving along the concentric circle of radius $\sqrt{r^2 + 1}$. An example is shown in Figure 6.

A related control theory problem is this: how do you get from one specified configuration to another? Consider this problem: you have your automobile lined up in the street parallel to and 3 meters from the sidewalk, next to an empty parking space. The spaces ahead and behind are occupied. You want to

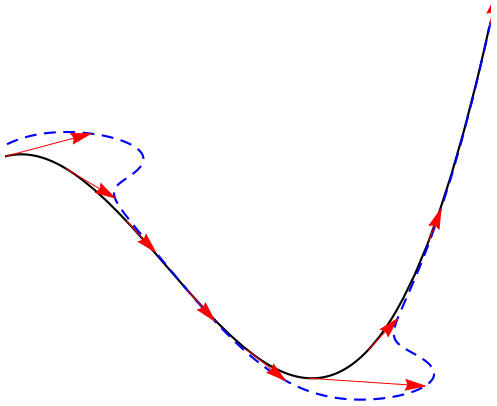


Figure 5: Determining the required truck path from the trailer path.

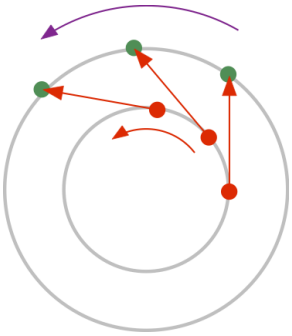


Figure 6: To move the trailer on a circle of radius r, drive the truck on a circle of radius $\sqrt{r^2+1}$.

move your vehicle to the side, but of course it can't jump to the side. So how do you get it into the space? When we learn to drive an automobile, the skill of *parallel parking* is one of the bigger challenges, as illustrated in Figure 7.

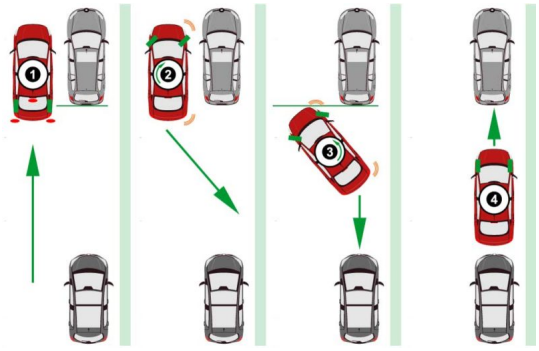


Figure 7: Parallel parking. (from officialdrivingschool.com)

To reduce the challenge somewhat, let's unhitch the trailer — let's suppose we just have the truck (or that we're riding a unicycle). To maneuver into the space, we should do these four things:

1. Pull forward.
2. Turn to face leftward (assuming the empty space is on the right).
3. Back up.
4. Turn rightward (to be facing forward again).

Mathematically, we've just described a commutator $FTF^{-1}T^{-1}$ of two basic motions. If we have a truck and a trailer, then a similar parsing into basic motions (of the truck) will reveal six individual steps.

Now let's consider what happens if we add additional trailers: after the first trailer we put a second trailer, following it in the same way that the first trailer follows the truck; and then a third trailer following the second, etc., forming a train. See Figure 8. In Lecture 3 we'll return to this figure; we'll look at the configuration space and dynamics for trains.

What we want to observe here is that the position of trailer k is determined from the path of trailer $k+1$ by a process of differentiation: the direction between the trailers is the slope of the tangent to this path. This probably sounds familiar! In fact one can show that the configuration space is one of the spaces in the Semple/monster tower over the plane, or nearly so. The "nearly so" comes from the fact that the trailers have a direction: they have a front and a back. Thus, to be precise, one needs to redo the basic construction: where previously we said "line" we need to say "ray." It's easy to see that if one applies this ray-monster construction, then one obtains an unbranched two-sheeted cover $\mathbb{R}^2_{\text{ray}}(1)$ of the usual monster space $\mathbb{R}^2(1)$. If you apply it

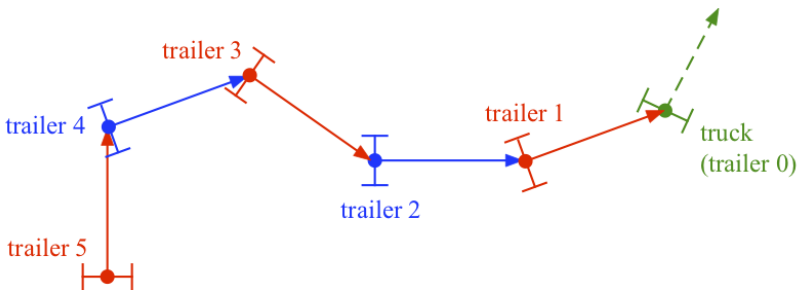


Figure 8: A train, made up of a truck pulling 5 trailers.

repeatedly, then you learn that the configuration space for the truck with n trailers is $\mathbb{R}^2_{\text{ray}}(n+1)$, the $(n+1)$st ray-monster space over the plane, and that's a 2^{n+1}-sheeted cover of the usual monster $\mathbb{R}^2(n+1)$.

Lecture 2: Combinatorial aspects

Here is an outline of this lecture:

- Charts and coordinates
- Lifting (prolonging)
- Coarse stratification
- Code words
- Moduli

Charts and coordinates

Begin with an open set U on M with coordinates $x_1, \ldots, x_m$ so that at each point the differentials $dx_1, \ldots, dx_m$ form a basis of the cotangent space. The monster tower over U is a tower of bundles with fiber $\mathbb{P}^{m-1}$. We explain here a systematic way of naming charts and coordinates for the spaces in the tower.

To avoid a cumbersome notation, we do this via an example: we suppose that M is 3-dimensional. Suppose first that we are just concerned with the data carried by a nonsingular point of a curve in U, and that at this point dx_1 doesn't vanish. Then the curvilinear data of the point is recorded by the values of the following quantities:

x_1	x_2	x_3
	$x_2' := \dfrac{dx_2}{dx_1}$	$x_3' := \dfrac{dx_3}{dx_1}$
	$x_2'' := \dfrac{dx_2'}{dx_1}$	$x_3'' := \dfrac{dx_3'}{dx_1}$
	$x_2^{(3)} := \dfrac{dx_2''}{dx_1}$	$x_3^{(3)} := \dfrac{dx_3''}{dx_1}$
	etc.	

To record the data of singular curves, however, we will need to turn some of these quantities upside-down, and we may need to work with derivatives with respect to quantities other than the three coordinates on the base. There will be 3^k charts (where k is the level), and each one is determined by a sequence of choices. Here we explain one possibility. To begin, we choose one of the three coordinates, and introduce derivatives of the other two coordinates with respect to it:

x_1	x_2	x_3
$x_1(3) := \dfrac{dx_1}{dx_3}$	$x_2(3) := \dfrac{dx_2}{dx_3}$	$\boxed{x_3(3) := x_3}$

The box indicates that we have introduced a (conveniently) redundant name. The three coordinates in the bottom row are called *active coordinates*; we choose one of them, and differentiate the two other active coordinates with respect to it:

x_1	x_2	x_3
$x_1(3) := \dfrac{dx_1}{dx_3}$	$x_2(3) := \dfrac{dx_2}{dx_3}$	$\boxed{x_3(3) := x_3}$
$x_1(32) := \dfrac{dx_1(3)}{dx_2(3)}$	$\boxed{x_2(32) := x_2(3)}$	$x_3(32) := \dfrac{dx_3(3)}{dx_2(3)}$

This shows three more steps:

x_1	x_2	x_3
$x_1(3) := \dfrac{dx_1}{dx_3}$	$x_2(3) := \dfrac{dx_2}{dx_3}$	$\boxed{x_3(3)}$
$x_1(32) := \dfrac{dx_1(3)}{dx_2(3)}$	$\boxed{x_2(32)}$	$x_3(32) := \dfrac{dx_3(3)}{dx_2(3)}$
$\boxed{x_1(321)}$	$x_2(321) := \dfrac{dx_2(32)}{dx_1(32)}$	$x_3(321) := \dfrac{dx_3(32)}{dx_1(32)}$
$x_1(3212) := \dfrac{dx_1(321)}{dx_2(321)}$	$\boxed{x_2(3212)}$	$x_3(3212) := \dfrac{dx_3(321)}{dx_2(321)}$
$x_1(32123) := \dfrac{dx_1(3212)}{dx_3(3212)}$	$x_2(32123) := \dfrac{dx_2(3212)}{dx_3(3212)}$	$\boxed{x_3(32123)}$

Thus on this chart, called $\mathcal{C}(32123)$, we have 13 coordinate functions. Note that each function extends to a rational function on the entire monster space, and there are no ambiguities in the names.

These coordinate systems are essentially due to Lejeune-Jalabert [25] (except for the redundant names).

In the case of a 2-dimensional base, we sometimes prefer a simpler notation, as follows: Begin with base coordinates x and y, and then use the standard prime notation for differentiation, e.g.,

$$y' = \frac{dy}{dx}$$
$$y'' = \frac{dy'}{dx}$$
$$x' = \frac{dx}{dy''}$$
$$x'' = \frac{dx'}{dy''}$$
$$y^{(3)} = \frac{dy''}{dx''}$$

This notation carries less information, and is ambiguous: x'' can mean different things in different charts. Note that at each step our choice can be framed as follows: either differentiate again with respect to the same coordinate as in the previous step (make the ***R**egular choice*), or differentiate with respect to the coordinate that has just been introduced (make the ***C**ritical choice*). Thus in the example, we could say we are working in a chart of type $RRCRC$.

Lifting (prolonging)

Robert Bryant: "Intuitively, prolongation is just differentiating the equations you have and then adjoining those equations as new equations in the system."

Consider a nonsingular point p of a curve C in M. We can associate to p the point of $\mathbb{P}TM$ representing the tangent direction of C at p. Thus, away from singularities, we have a curve $C(1)$ in $\mathbb{P}TM$, the *prolongation* or *lift*. In turn, we can lift this curve to the next level (arguing that in fact all of its tangent directions are focal), and continue upward through the tower, obtaining lifts $C(2)$, $C(3)$, etc. We also want to lift a singular curve C in M, and we do so by fiat: we lift at all nonsingular points of C, and then take the closure. This may give us several points over a singular point of C; even if it gives us a single point, the nature of the singularity will change, and we may even get a nonsingular point on the lift. In the mathematical literature, this process goes by at least three distinct names: prolongation, lifting, and Nash blow-up.

Our examples are all done in the case where M is the plane. First, we work with a parametrized curve, taking an example from our paper [8]. Suppose that $x = t^2$ and $y = t^4 + t^5$. We calculate as follows:

$$\begin{aligned} y' &= \frac{dy}{dx} = \frac{4t^3 + 5t^4}{2t} = 2t^2 + \frac{5}{2}t^3 \\ y'' &= \frac{dy'}{dx} = \frac{4t + \frac{15}{2}t^2}{2t} = 2 + \frac{15}{4}t \\ x' &= \frac{dx}{dy''} = \frac{8}{15}t. \end{aligned}$$

These equations parametrize a curve in the third monster. (See Figure 9.) Setting $t = 0$, we find one point over the origin,

$$(x, y, y', y'', x') = (0, 0, 0, 2, 0),$$

which we can explicate as follows: the (generalized) tangent line is horizontal, and in fact the point carries the same data up to second order as the parabola $y = x^2$, but its third-order datum is infinite. Note this implicit equation for the curve: $(y - x^2)^2 = x^5$.

Our second example begins with cuspidal cubic curve $y^2 = x^3$. By implicit differentiation we find that

$$2yy' = 3x^2.$$

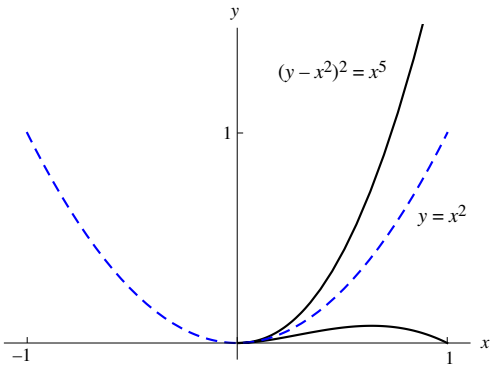


Figure 9: A ramphoid cusp.

Note that these two equations don't give an irreducible curve, since over the origin they allow any value for y'. See Figure 10, noting the spurious component. To remove it, we remark that, assuming we are away from the origin, we have

$$x = \frac{4}{9}(y')^2 \quad \text{and} \quad y = \frac{8}{27}(y')^3.$$

Thus these equations should also be used in defining the lift. To continue the calculation for one more step, we use $x' := dx/dy'$, and calculate that

$$x' = \frac{8}{9}y'.$$

Thus the second-order datum at the origin is infinite. If, loosely speaking, we think of this as curvature, we're claiming that as we approach the origin on this curve, its curvature goes to infinity. See Figure 11.

Going in the opposite direction: given a point on the kth monster, we want to generate a curve for which it is the curvilinear data. Montgomery and Zhitomirskii call this process *Legendrization*. Here's a calculation taken from our very first paper [7]. Using the chart with coordinates

$$y' = \frac{dy}{dx} \qquad x' = \frac{dx}{dy'} \qquad x'' = \frac{dx'}{dy'} \qquad y'' = \frac{dy'}{dx''}$$

we consider the point

$$(x, y, y', x', x'', y'') = (0, 0, 0, 0, 1, 1).$$

Figure 10: The lift of the cuspidal cubic.

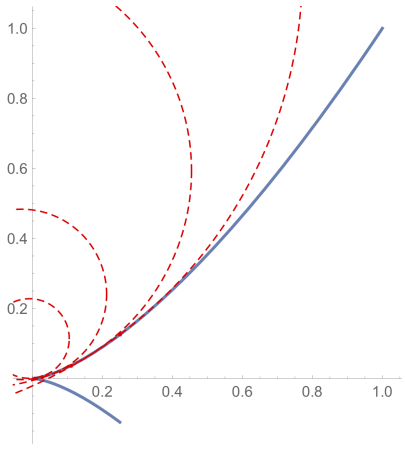


Figure 11: As one approaches the origin on the cuspidal cubic, the osculating circle shrinks to a point.

Beginning with two linear parametrizations, we calculate as follows:

$$\begin{aligned}
y'' &= 1+t \\
x'' &= 1+t \\
y' &= \int y''dx'' = \int (1+t)\,dt = \tfrac{1}{2}(1+t)^2 - \tfrac{1}{2} \\
x' &= \int x''dy' = \tfrac{1}{3}(1+t)^3 - \tfrac{1}{3} \\
x &= \int x'dy' = \tfrac{1}{15}(1+t)^5 - \tfrac{1}{6}(1+t)^2 + \tfrac{1}{10} \\
y &= \int y'dx = \tfrac{1}{42}(1+t)^7 - \tfrac{1}{30}(1+t)^5 - \tfrac{1}{24}(1+t)^4 + \tfrac{1}{12}(1+t)^2 - \tfrac{9}{280}
\end{aligned}$$

(Since the last coordinate of the point is nonzero, we could actually have started with the constant parametrization $y'' = 1$, obtaining a somewhat simpler result.)

Coarse stratification

We will now apply the basic construction in a different way. Refer to our one-page handout for a reminder of this construction. For an alternative explanation of the following, see [36].

Again suppose that $\mathcal{B}$ is a rank b subbundle of the tangent bundle TM; to avoid a silly trivial case, assume its rank isn't zero. Then the relative tangent bundle $\mathcal{V} = T(\widetilde{M}/M)$ is a subbundle of $\widetilde{\mathcal{B}}$; its fiber consists of vectors mapping to zero, sometimes called *vertical vectors*. Its rank is $b-1$. Applying the basic construction to the bundle $\mathcal{V}$ on $\widetilde{M}$, we obtain the space $\mathbb{P}\mathcal{V}$ carrying its focal bundle $\widetilde{\mathcal{V}}$, again of rank $b-1$. Since $\mathcal{V}$ is a subbundle of $\widetilde{\mathcal{B}}$, the space $\mathbb{P}\mathcal{V}$ is naturally a subset of $\mathbb{P}\widetilde{\mathcal{B}}$. In fact, each fiber of $\mathbb{P}\mathcal{V}$ is a hyperplane inside the $\mathbb{P}^{b-1}$-fiber of $\mathbb{P}\widetilde{\mathcal{B}}$. Thus $\mathbb{P}\mathcal{V}$ is a divisor (codimension one) subvariety of $\mathbb{P}\widetilde{\mathcal{B}}$. We call it a *baby monster*.

Here's where we want to apply the baby monster construction: to the monster spaces over some base manifold or variety M. We can do this to any projection map in the monster tower

$$M(j-1) \xrightarrow{\pi_{j-1}} M(j-2),$$

and what we obtain is a divisor on $M(j)$, which we call the jth *divisor at infinity* and denote by I_j. If the dimension of the base is m, then by construction I_j carries a bundle of rank $m-1$. We can iterate this construction, building a tower of $\mathbb{P}^{m-2}$-bundles over $M(j-1)$. Here is our notation:

$$\cdots \to I_j[3] \to I_j[2] \to I_j[1] \to I_j[0] = I_j \quad \to M(j-1)$$

Note that:

- I_j is a divisor on $M(j)$

- $I_j[1]$ has codimension 2 in $M(j+1)$
- $I_j[2]$ has codimension 3 in $M(j+2)$
- etc.

(If $m = 2$, then all these maps are isomorphisms.)

What's the geometric meaning of these loci? Using either our interpretation of the coordinate systems, or reflecting upon how the lift of a curve could possibly have a vertical tangent, we realize that the divisor at infinity consists of those curvilinear data points for which we consider the jth-order data to be infinite.

For example, we have seen that the lift of the cuspidal cubic curve $y^2 = x^3$ is tangent to the fiber of the first monster over the origin. Thus the second lift meets the divisor at infinity. See Figure 12.

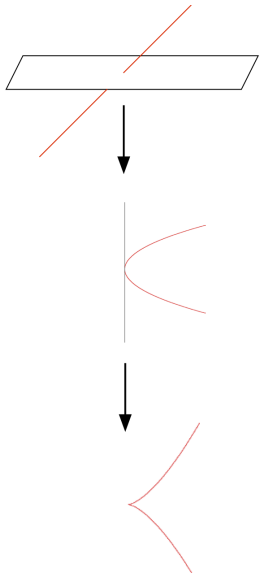


Figure 12: The second lift of the cuspidal cubic meets the divisor at infinity.

Let's also look again at the ramphoid cusp; see Figure 9. Here are our calculations:

$$\begin{aligned} x &= t^2 \\ y &= t^4 + t^5 \\ y' &= 2t^2 + \frac{5}{2}t^3 \\ y'' &= 2 + \frac{15}{4}t. \end{aligned}$$

These equations give us the second lift of the curve on the base, and it's tangent to the fiber of $M(2)$ over $M(1)$. We can compute

$$y''' = \frac{dy''}{dx} = \frac{15}{8t}$$

but of course we can't evaluate this at $t = 0$; that's because the third lift hits I_3, and that's why earlier we chose to compute

$$x' = \frac{dx}{dy''} = \frac{8}{15}t.$$

As this example suggests, it's straightforward to use coordinates to tell when you're on a divisor at infinity. We work this out in full detail in [8]. In this example, the divisor at infinity is the locus $x' = 0$.

What about $I_j[1]$? What's its geometric meaning? Well, suppose that when you lift your curve to $M(j-1)$, you are not just tangent to a fiber over $M(j-2)$, but tangent to second order. Then when you lift to $M(j)$, you won't just hit I_j; you'll be tangent to it. And when you lift yet again, then you'll be on $I_j[1]$.

To compare all these divisors at infinity and their associated baby monsters, let's pull everything back to our highest level k. The complete inverse image of I_j will be denoted in the same way, but now it's a divisor up on $M(k)$. With this convention, we now have a nest

$$I_j \supset I_j[1] \supset I_j[2] \supset \cdots \supset I_j[k-j].$$

We can also form intersections of spaces taken from different nests: we consider the *intersection locus*

$$I_W := \bigcap_{j=2}^{k} I_j[n_j - 1] \qquad (*)$$

for some specified nonnegative integers n_j (with the convention that $I_j[-1]$ is all of $M(k)$). The main theorem of [4] tells us this:

- This intersection is transverse.
- Its codimension (if it's nonempty) is $n_2 + \cdots + n_k$.
- If $k \leq \dim M$, then it's nonempty.

It also says more generally when the intersection is nonempty, and that naturally leads to our next topic.

Code words

Looking back at (*), note that we used a letter W. This indicates the *code word* of the intersection locus, which records, in a different way, which spaces we are intersecting. The alphabet for our code consists of all symbols V_A, where A is a finite subset of the integers strictly greater than 1. Although this is an

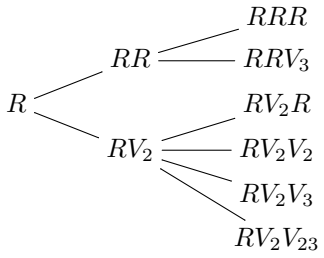


Figure 13: Code words up to length 3 (assuming $\dim M \geq 3$).

infinite alphabet, the rules for creating a valid word will imply that there are only finitely many words of each specified length. To be consistent with prior usage, we will use R in place of $V_\emptyset$. (The symbols R and V have been chosen to suggest "regular" and "vertical.") The rules for creating a code word are as follows:

1. The first symbol must be R.
2. Immediately following the symbol V_A, one may put any symbol V_B, where either B is a subset of A, or B is a subset of $A \cup \{j\}$, with j being the position of the symbol.
3. The cardinality of A is less than m.

Note that j cannot appear in a subscript prior to position j.

To explain how the code word specifies an intersection locus, we present an example at level 8:

$$W = RV_2V_{23}V_{23}V_{25}V_5V_5V_5.$$

The number 2 first appears in position 2, and then appears in three subsequent positions. This means that one of the spaces used in the intersection is $I_2[3]$. Similarly, we use $I_3[1]$. The number 4 never appears in a subscript. Thus we can omit I_4, i.e., write $I_4[-1]$; similarly we can omit I_6, I_7, and I_8. Here is the result:

$$\begin{aligned} I_W &= I_2[3] \cap I_3[1] \cap I_4[-1] \cap I_5[3] \cap I_6[-1] \cap I_7[-1] \cap I_8[-1] \\ &= I_2[3] \cap I_3[1] \cap I_5[3]. \end{aligned}$$

Is this intersection locus nonempty? The answer depends on the dimension of the base. Note that the third rule for code words limits the length of the subscripts. Thus this intersection locus is nonempty if the base has dimension at least 3; over a surface it's empty.

Figure 13 shows the code words up to length 3, assuming the dimension of the base is at least 3. Over a surface, the last code word RV_2V_{23} should not be used; thus there are five intersection loci at level 3, rather than six.

When the base is a surface, note that at each position j of the code word there are at most three possibilities:

1. Use R.

2. Use V_j.

3. Repeat the previous symbol (assumed not to be R).

This leads to the alternative RVT code, originally developed by the differential geometers. (See [29, Chapter 3]; in [30], it's called the GST code.)

We illustrate it in Table 1, which gives representative curve germs for each of the five intersection loci at level 3.

Table 1: The five intersection loci at level 3.

RRR	RRV	RVR	RVV	RVT
RRR	RRV_3	RV_2R	RV_2V_3	RV_2V_2
$y=0$	$y^2=x^5$	$y^2=x^3$	$y^3=x^5$	$y^3=x^4$

How many intersection loci are there? To answer this, one can work out a simple recursion. In the case of a surface, the recursion immediately implies that the number of code words of specified length is a Fibonacci number: there are 2 intersection loci at level two, 5 at level three, 13 at level four, etc. Figure 14 shows the codewords up to level four.

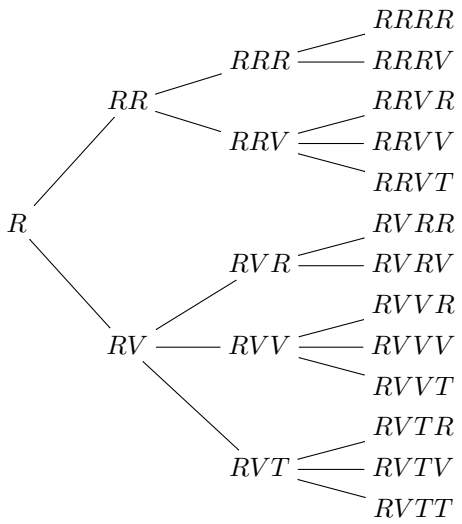


Figure 14: RVT code words up to length 4.

Moduli

Thus far we have only discussed discrete or coarse aspects of the monster spaces. A few researchers, most notably Mormul, have devoted considerable attention to understanding their moduli. (See, e.g., [31].) We find these results deep and mysterious in comparison to the prior considerations, and we're just going to give a few vague remarks. We're thinking about equivalence classes of plane curve germs or (for Mormul) equivalence classes of Goursat germs, where (in the case of curves) we allow arbitrary diffeomorphisms at a point of the base manifold M. If two curve germs give different code words, then certainly they are inequivalent, but suppose they have the same code word. We find it surprising that in fact there are moduli, i.e., there are continuous families of inequivalent curve germs (or Goursat germs) with the same code word. This is wide-open territory. Currently we have no general unifying insight as to where and why moduli should appear.

Lecture 3: Mechanical aspects

We are not as well acquainted with this aspect, so our reporting may be superficial or misguided. Nevertheless, we think that learning more about this connection may be helpful in both directions. You've just heard from an expert about the relevant dynamics, Alejandro Bravo-Doddoli. We may cover some of the same territory, while trying to tie it to the other aspects.

Here is an outline of this lecture:

- The configuration space for a truck with trailers
- Driving the train; rigid trains
- Vector fields (infinitesimal motions)
- Lie brackets
- Singular configurations
- The runaway train

The configuration space for a truck with trailers

We are examining the motion of a *truck* pulling n *trailers* of unit length; we call the entire configuration a *train*. When convenient, we may also say "trailer 0" instead of "truck." We idealize each trailer as a point together with a unit vector pointing in the direction of the previous trailer; the truck's unit vector may point in any direction. See Figure 15 for an example with $n = 5$.

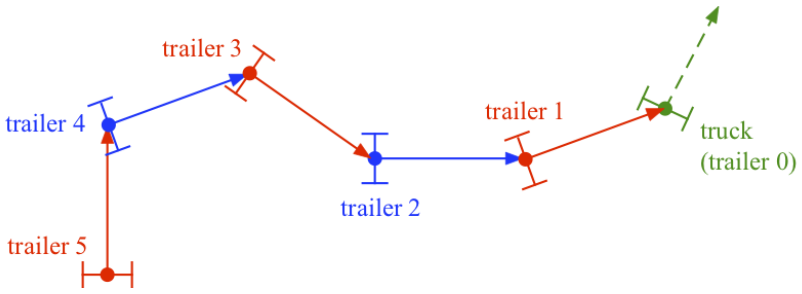


Figure 15: A truck pulling 5 trailers.

The configuration space was described by Jean [22], but we introduce coordinates in a somewhat different way. A somewhat confusing aspect, but one entrenched in the literature — and actually compatible with the way in which we think about higher-order data — is that the angles are named from back

(last trailer) to front (truck). The configuration space is $\mathbb{R}^2 \times (S^1)^{n+1}$, a manifold of dimension $n+3$. A point of the configuration space is specified by an $(n+3)$-tuple

$$(x, y, \varphi_0, \varphi_1, \ldots, \varphi_n).$$

The first two coordinates (x, y) record the position of the last trailer (trailer n). The coordinates $\varphi_1, \ldots, \varphi_n$ record the *bending angles* formed by successive trailers, ordered from the back of the train to the front: φ_k records the angle formed at trailer $n-k$ between the unit vectors associated to trailers $n-k$ and $n-k+1$. We define φ_0 to be the *heading angle* formed by the unit vector associated to the last trailer and the horizontal direction. (See Figure 16.) Note that the position of the truck, and indeed of all trailers except the last, is not directly recorded in the coordinates of the configuration space.

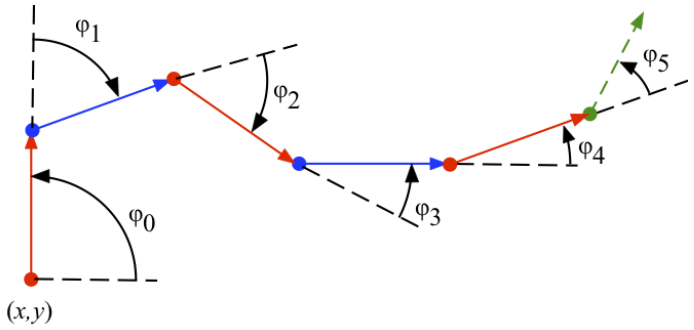


Figure 16: Coordinates for the truck-with-trailers configuration space.

Here's an unrealistic aspect of our model: the trailers can intersect. We also allow a bending angle to have value π, which leads to numerous "accordion configurations," with trailers totally overlapping. See Figure 17. For simplicity, in the following treatment we will assume that each of the angles is acute: $-\pi/2 \leq \varphi_k \leq \pi/2$. In particular, this assumes that the truck is pulling the train, not pushing it. This rules out the accordion configurations, but overlapping may still happen. Our convention is that bending to the left is positive; in Figure 16, angles φ_1 and φ_2 are negative, and the others are positive.

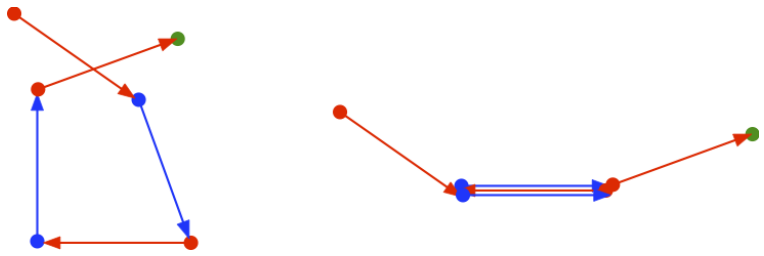


Figure 17: Overlapping trailers and "accordion configurations."

In Lecture 1, we noted that this configuration space is $\mathbb{R}^2_{\text{ray}}(n+1)$, the ray-monster space over the plane. The map

$$\mathbb{R}^2_{\text{ray}}(n+1) \to \mathbb{R}^2_{\text{ray}}(n)$$

just strips away the truck, and the first trailer then plays the role of the truck.

Driving the train; rigid trains

The truck can be driven following any differentiable path, subject to the following condition: the velocity vector must point in the direction that the truck is pointing (as indicated by its unit vector). If the velocity vector is zero, this is interpreted as a vacuous condition (automatically satisfied). It is thus legal to drive the truck simply by turning, i.e., by keeping its position fixed while changing φ_n.

The motion of each trailer is determined by two conditions:

1. The distance between successive trailers must always be the unit length. (This is a holonomic constraint.)

2. The velocity vector must point in the direction that the trailer is pointing (as specified by φ_i), i.e., the velocity vector must point in the direction of the previous trailer. (This is a nonholonomic constraint.)

These constraints imply a certain system of differential equations, which, given the path of the truck, can be solved to determine the paths of the trailers.

Going in the other direction, if we are given the path of a trailer, then the process of obtaining the path of the previous trailer is a geometric process involving derivatives. As we noted in Lecture 1, there's a nice geometric description: Suppose we specify a differentiable path for the last trailer. At each point, draw the unit tangent vector, and then mark the head of this vector. The heads of all these vectors trace out the path of the prior trailer. Now repeat the procedure, moving forward in the train until you reach the truck.

As an example, suppose you want to move the last trailer along a circle of radius r. Then the prior trailer should be moving along a circle of radius $\sqrt{r^2+1}$. Working forward, we obtain a sequence of concentric circles, with the truck driving along the outer circle. As a special case of this construction, take $r = 0$. Then the bending angles are

$$\varphi_k = \sin^{-1} \frac{1}{\sqrt{k}}$$

as shown on the left in Figure 18. Observe that the entire train will move rigidly, as if the angles between the successive trailers were welded at certain fixed angle. The last trailer isn't moving; it's just turning.

Now note what we could do to embellish this example: we could attach additional trailers to the end, as shown on the right in Figure 18. Since the end isn't moving, we can satisfy the motion constraints by not moving these

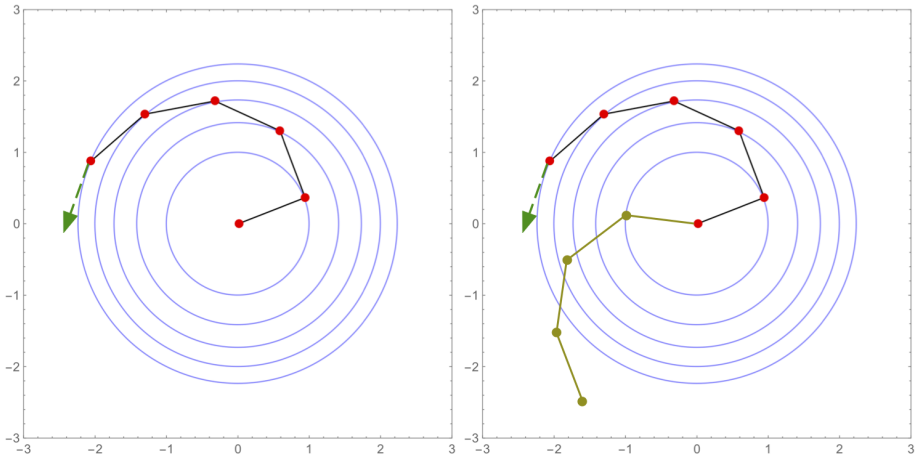


Figure 18: A rigid train and an embellishment.

additional trailers at all. The overall motion then looks like this: the rear of the train isn't moving at all, while the front of the train is moving rigidly. Just one angle is changing.

Vector fields (infinitesimal motions)

To probe the dynamics of the train, we will look at infinitesimal motions, i.e., vector fields on the configuration space. We want to distinguish two sorts of infinitesimal motions:

- motions that simply preserve the distance between the trailers,
- motions that correspond to legal ways of moving the train.

The latter sort of motion is quite restrictive, since the only available infinitesimal motions of the truck are linear combinations of these two:

1. $v =$ infinitesimally turn the truck leftward;
2. $f =$ infinitesimally drive the truck forward.

The first motion v extends to a legal motion v_n of the train in a trivial way: the trailers don't move at all. In coordinates, we have

$$v_n = \frac{\partial}{\partial \varphi_n}.$$

The second motion f likewise extends to a legal motion f_n of the entire train. If we drive the truck forward, however, then we certainly expect the trailers to move. Thus the coordinate expression for f_n is more complicated, and we omit it.

Our choice of notation comes from the connection with the monster spaces: When we regard the configuration space as the ray-monster $\mathbb{R}^2_{\text{ray}}(n+1)$, the vector field v_n is a field of vertical vectors. Physically, this means that when we forget about the truck, the remainder of the train is stationary. Taking both v_n and f_n, they everywhere span a rank 2 subbundle of the tangent bundle of the configuration space; in fact it's the focal bundle. (In the present context, the words "veer" and "forward" may also be memorable.)

Jean introduces two sorts of vector fields on the configuration space, for a total of $2(n+1)$ vector fields. We have already met v_n and f_n. For $0 \le k \le n$, we define two vector fields as follows:

- The vector field f_k treats trailer $n-k$ as a truck, driving it forward infinitesimally and pulling the following trailers behind it; in doing so, it does not change the angles φ_i at the earlier trailers ($i > k$). This is of course not a legitimate way of driving the train; we are treating the previous trailers as if they were a hood ornament, with the wheels removed and all angles welded. See Figure 19.
- The vector field v_k also treats trailer $n-k$ as a truck, turning it infinitesimally to the left; thus the following trailers don't move. To compensate for this turning, turn the prior trailer infinitesimally to the right, and again treat all trailers forward of that as a hood ornament. See Figure 20.

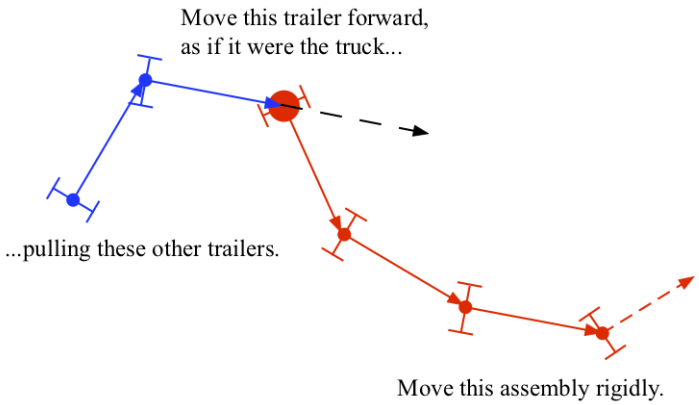


Figure 19: The infinitesimal motion (vector field) f_k. (Here $n = 5$ and $k = 2$.)

One has this basic formula:

$$f_k = v_{k-1} \sin\varphi_k + f_{k-1} \cos\varphi_k.$$

By repeated use of this formula, one can express vector f_n, as a linear combination of f_0 and $v_0, v_1, \ldots, v_n$. (If you noticed that the total number of vector fields doesn't quite reach the dimension $n+3$ of the configuration space, that's because none of these vector fields moves the last trailer sideways. You could introduce one more vector field v_{-1} to handle this possibility.)

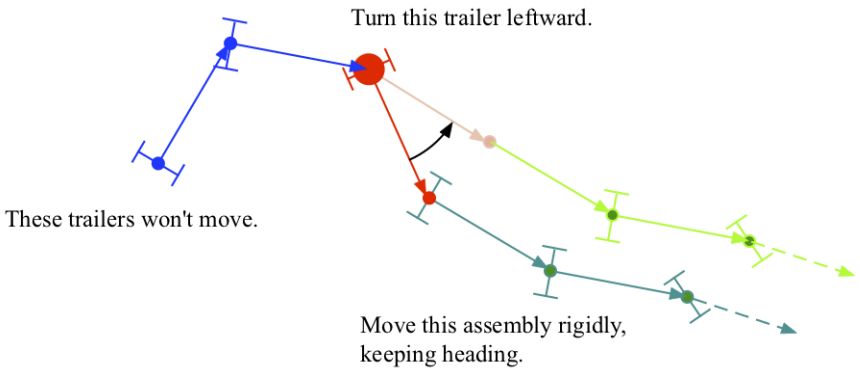


Figure 20: The infinitesimal motion (vector field) v_k. (Here $n = 5$ and $k = 2$.)

Lie brackets

In Lecture 1, we looked at the problem of parallel parking: maneuvering a vehicle sideways into a location. This led us to the notion of a commutator of two motions. We also encountered the notion of Lie bracket of vector fields. The following standard formula relates the two notions:

$$[X, Y] = \frac{1}{2}\frac{\partial^2}{\partial t^2}\Big|_0 \left(\mathrm{Fl}^Y_{-t} \circ \mathrm{Fl}^X_{-t} \circ \mathrm{Fl}^Y_t \circ \mathrm{Fl}^X_t\right).$$

(We took this formulation from [26].) Here X and Y are two vector fields on a smooth manifold. The motion Fl^X_t is the *flow* along X: starting at each point of X, integrate the vector field from time 0 to time t; this determines where the point should be moved. Observe that Fl^X_{-t} is its inverse. Thus the expression in parentheses is a commutator. The formula says that Lie bracket is akin to an "infinitesimal commutator." See Figure 21.

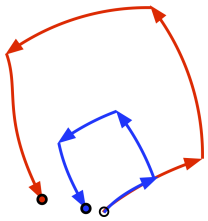


Figure 21: The Lie bracket is an "infinitesimal commutator."

Thus Lie brackets are an important idea in control theory. One natural question to ask is: if we are given a distribution $\mathcal{D}$ on a manifold (a subbundle of the tangent bundle) what other vector fields can we generate from the ones in the distribution? We think of the vector fields within the distribution as the ones we can realize directly by single motions, and the ones we obtain via Lie bracketing are those which we can obtain by combining motions.

Recall that in Lecture 1 we considered the *Lie squares sequence*

$$\mathcal{D} = \mathcal{D}_1 \subset \mathcal{D}_2 \subset \mathcal{D}_3 \subset \cdots$$

in which each bundle is the Lie square of the previous bundle:

$$\mathcal{D}_i = (\mathcal{D}_{i-1})^2 = \mathcal{D}_{i-1} + [\mathcal{D}_{i-1}, \mathcal{D}_{i-1}].$$

Also recall that a *Goursat distribution* is one for which all of these are subbundles (i.e., the rank is constant on the manifold), with the rank increasing by one at each step until we reach the tangent bundle.

This is exactly the situation on the configuration space for a truck with n trailers. At each point of the configuration space we have two vectors v_n and f_n, the infinitesimal motions of the train caused by turning and driving the truck, respectively. They fit together into a rank 2 distribution. Here is an intuition for why this is a Goursat distribution: By bracketing v_n and f_n, we obtain a third independent vector field; in the distribution spanned by the three vector fields we find both v_{n-1} and f_{n-1}. Thus we now have control of the first trailer, and as if it were the truck. Repeat this argument for n steps.

Note that when we're working in $\mathcal{D}_3$, we're allowed to input four vector fields into calculations, e.g.,

$$\Big[[v_n, f_n], [v_n, f_n]\Big]$$

is a vector field in $\mathcal{D}_3$ (when $\mathcal{D}$ is the focal bundle on the configuration space). We can represent this calculation by a tree, as shown on the left in Figure 22. A more restrictive sort of calculation is represented by the tree on the right, and this leads us to the idea of the *slow growth sequence*

$$\mathcal{D} = \mathcal{D}^0 \subset \mathcal{D}^1 \subset \mathcal{D}^2 \subset \mathcal{D}^3 \subset \cdots$$

in which

$$\mathcal{D}^i = \mathcal{D}_{i-1} + [\mathcal{D}_{i-1}, \mathcal{D}].$$

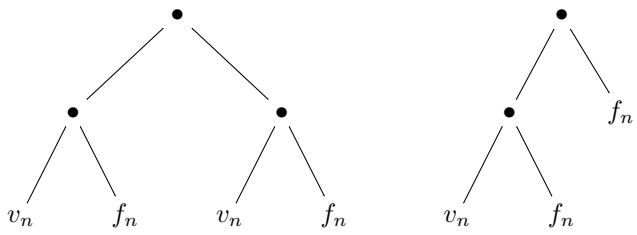


Figure 22: Lie bracketing trees.

We don't want to assume that these are subbundles; that would be too restrictive. They are *subsheaves* of the sheaf of the tangent bundle. In fact the

idea is to study the distribution by studying the associated sequence of ranks, which is a function on the manifold.

Let's do this for the focal bundle on the configuration space. One can show that there is a generic behavior, for all points representing configurations without any right angles; the generic behavior is that $\mathcal{D}^i = \mathcal{D}_i$ at such points, and thus the *slow growth vector* is simply $(2, 3, 4, \ldots, n+3)$. The other extreme was pointed out by Luca and Risler in [32]: for configurations in which all the bending angles are right angles, the slow growth vector is

$$(2, 3, 4, 4, 5, 5, 5, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 7, 7, 8, \ldots),$$

where the number of repetitions is a Fibonacci number. The same information is recorded by Jean's *beta vector*, which here is

$$(1, 2, 3, 5, 8, \ldots).$$

Its entries tell us when we first reach rank 2, then rank 3, etc. This worst-case growth is telling us, in some sense, that this particular train configuration is the hardest to control.

Singular configurations

Jean [22] calls a configuration *singular* if the slow growth vector is something other than the generic $(2, 3, 4, \ldots, n+3)$, and he provides an elegant recursion for computing the slow growth vector at all points.

To explain his analysis, we need to delve into how the trigonometric coordinates of the configuration space relate to the algebro-geometric coordinates on the monster space over $\mathbb{R}^2$. On the second monster we have a chart with coordinates (x, y, y', y''), whereas for the configuration space of the truck with one trailer we use coordinates $(x, y, \varphi_0, \varphi_1)$. For simplicity let's assume $\varphi_1 \geq 0$ and $\varphi_2 \geq 0$. The relations between these coordinate systems will look familiar:

$$\tan \varphi_0 = y'$$
$$\tan \varphi_1 = \frac{y''}{(1 + (y')^2)^{3/2}}$$

Adding another trailer, here is a formula relating φ_2 to third-order data:

$$\tan \varphi_2 = \frac{y^{(3)} + y'' + (y'')^3}{(1 + (y'')^2)^{3/2}}$$

Here we assume that $y' = 0$. If we use the chart in which

$$x' = \frac{dx}{dy'} \quad \text{and} \quad x'' = \frac{dx'}{dy'}$$

then here are the relations:

$$\cot \varphi_1 = x'(1 + (y')^2)^{3/2}$$
$$\tan \varphi_2 = \frac{-x'' + 1 + (x')^2}{((x')^2 + 1)^{3/2}}$$

(again assuming for the latter that $y' = 0$).

One could continue in this fashion, but the formulas are not particularly illuminating. More important is to understand how the coarse stratification developed in Lecture 2 is interpreted on the configuration space. The divisor at infinity I_2 is defined by the vanishing of x', and thus by $\varphi_1 = \pi/2$. Its baby-monster prolongation $I_2[1]$ is defined by the vanishing of both x' and x'', and thus by the additional condition $\varphi_2 = \pi/4$. Similarly, at all levels, we can see what configurations are *singular*, meaning that they aren't on the open dense locus with code word $RRR\ldots$, and we can see what code words they give.

Given a configuration of the train, here's how to read its RVT code word from its bending angles. If $\varphi_i = \pm\pi/2$, then write V in position i. If you see the sequence of special bending angles

$$\pm\left(\pi/2, \quad \pi/4, \quad \sin^{-1}(1/\sqrt{3}), \quad \sin^{-1}(1/\sqrt{4}), \quad \cdots\right)$$

starting at position i, then write V in position i followed by a sequence of T's for the other special angles. In all other cases write R. For example, the trains shown in Figure 18 have code words $RVTTTT$ and $RRRRRVTTTT$; the train shown in Figure 23 has code word $RVRVTTVR$.

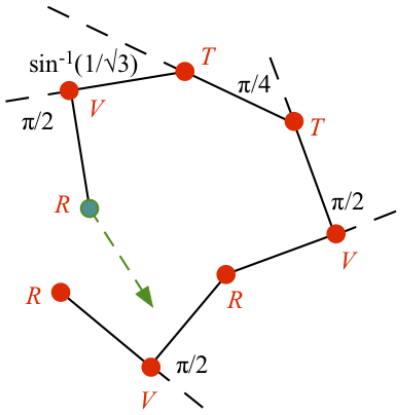


Figure 23: A train with code word $RVRVTTVR$.

The code word W of a configuration determines its slow growth vector. Jean expresses his recursion using his beta vector

$$\beta(W) = (\beta_1(W), \beta_2(W), \ldots, \beta_k(W)).$$

The first two entries are always 1 and 2. He proves this recursion (valid for $j \geq 3$, and with X indicating a single symbol):

$$\begin{aligned}
\beta_j(WR) &= \beta_{j-1}(W) \\
\beta_j(WXV) &= \beta_{j-2}(W) + \beta_{j-1}(WX) \\
\beta_j(WXT) &= 2\beta_{j-1}(WX) - \beta_{j-2}(W)
\end{aligned}$$

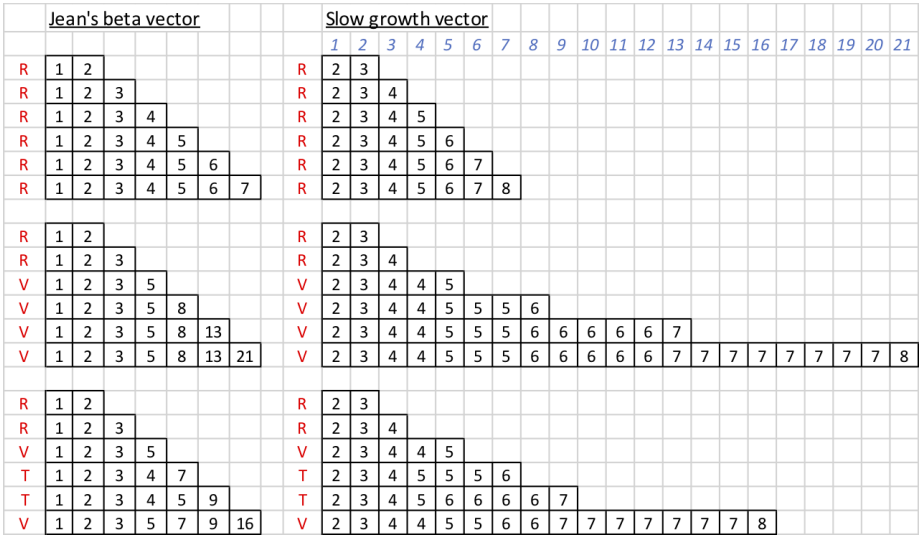


Figure 24: Recursive calculation of Jean's beta and the slow growth vector.

Figure 24 presents three examples.

Now remember that the context for these calculations is on the configuration space. We believe that the same rules should apply in the wider context of the monster over any surface, and that probably is an easy consequence of Jean's work. We would prefer, however, to give a fresh proof of this fact, not invoking bending angles. Together with Corey Shanbrom, we're in the midst of a project which we think will explain Jean's recursion by developing a theory of calculational trees like those we saw in Figure 22. For example, we speculate that the tree shown in Figure 25 should be associated to the code word $RRVT$, and we are developing a recursion for trees which should yield Jean's numerical recursion. Note in this example that the number of leaves is 7, which is the last entry in Jean's beta vector.

The runaway train

Alejandro Bravo-Doddoli has reported to you about his paper with García-Naranjo [2], and we won't attempt to repeat the details. We'll just point out how the special strata on the monster come up in their construction.

We consider a physically-motivated problem. Suppose that each trailer including the truck) has an equal point mass. We assume that the train is in a certain configuration, and that the truck is set into motion with a certain initial velocity and initial angular velocity. Thereafter the truck and trailers move according to Newtonian mechanics: they form a runaway train. We want to know how it moves. This is an initial value problem, and not really a control theory problem. In fact it's an example of out-of-control theory!

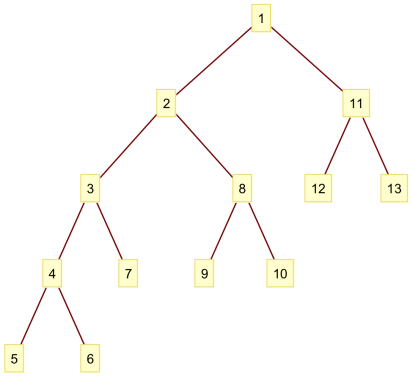


Figure 25: Speculative tree associated to the code word $RRVT$.

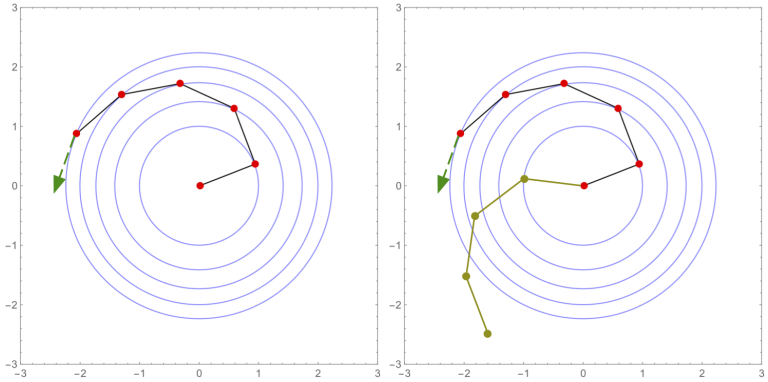


Figure 26: These configurations lead to periodic solutions.

In Newtonian mechanics, the energy is conserved, so in our analysis it makes sense to specify the level of energy and then to examine the possible trajectories at that level. Bravo-Doddoli and García-Naranjo show that, on a trajectory, the angular velocity of the truck is constant, so let's also fix that constant and assume it's not zero.

In this sort of analysis, one would like to look for periodic solutions, and we actually know some already. The relevant configurations were depicted back in Figure 18, which we repeat in Figure 26. Recall that such a train moves just by rotating around one of its trailers, while the back of the train remains motionless. The rigid train leads to a single periodic trajectory, while the embellishments lead to entire tori of periodic trajectories. Thus there are $n+1$ *critical values* of the energy for which we find this abundance of periodic solutions. Bravo-Doddoli and García-Naranjo do a bifurcation analysis, showing that at each of these critical values, there is a change in the dynamics of the system.

Lecture 4: Enumerative aspects

Thanks to Ferran Dachs Cadefau for his preliminary talk.

Here is an outline of this lecture:

- Enumerative geometry
- Contact formulas
- A strategy for proving a contact formula
- Intersection theory
- A higher-order contact formula
- Orbits and strata
- Prolongation in families

Enumerative geometry

Enumerative geometry is a branch of algebraic geometry. It is concerned with problems of counting objects of a specified type, in situations where one expects the answer to be finite.

Example. Given a nonsingular cubic surface in $\mathbb{P}^3$ (the locus of solutions to a degree 3 homogeneous equation in w, x, y, z), how many lines does it contain?

The answer is: 27.

Actually, since this is algebraic geometry, one ought to be asking "Over what field is the surface defined?" The default answer in enumerative geometry is $\mathbb{C}$, the field of complex numbers. But in this problem, the answer is correct even for a real cubic surface: it contains 27 real lines. Figure 27 shows an example.

Example. Let C be an algebraic curve of degree d in the complex projective plane. Its *class* is

$$d^\vee := \text{the number of tangents to } C \text{ through a general point of } \mathbb{P}^2.$$

For a nonsingular curve $d^\vee = d(d-1)$, but for a singular curve we need correction terms. Suppose that C has (ordinary) nodes and cusps, but no more complicated singularities. Then

$$d^\vee = d(d-1) - 2\delta - 3\kappa,$$

where δ is the number of nodes and κ is the number of cusps. If we define two additional quantities

$$\beta := \text{the number of bitangent lines to } C,$$
$$\varphi := \text{the number of inflection points (flexes) on } C,$$

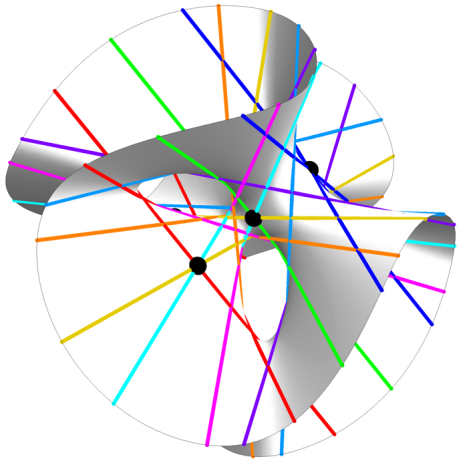


Figure 27: A real cubic surface and its 27 lines (figure by Greg Egan).

then there are three additional equations:

$$\begin{aligned}\varphi &= 3d(d-2) - 6\delta - 8\kappa;\\ d &= d^\vee(d^\vee - 1) - 2\beta - 3\varphi;\\ \kappa &= 3d^\vee(d^\vee - 2) - 6\beta - 8\varphi.\end{aligned}$$

Collectively, the four equations are known as the *Plücker formulas*. According to Kleiman [24, pp. 307–308], the nonsingular case dates back to a 1756 anonymous work on plane curves, and the idea of having "correction terms" for singularities was apparently due to Poncelet. The formulas above — and the original theory of projective duality that facilitates obtaining them — are due to Plücker.

To discover and prove such formulas, one works with appropriate *parameter spaces*, i.e., spaces whose points represent objects of a certain type, in such a way that "moving continuously through the space has the effect of continuously varying the object." If you're wondering about lines on a cubic surface, then it's natural to work with the *Grassmannian of lines in* $\mathbb{P}^3$, which is a beautiful 4-dimensional algebraic variety. If you want to understand the Plücker formulas, you want an object already mentioned in these lectures; it's the first monster space over the projective plane, which is also the total space of the projectivized tangent bundle:

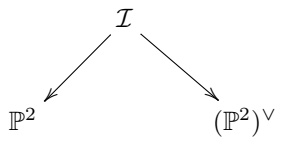
$$\mathbb{P}^2(1) = \mathbb{P}T\mathbb{P}^2.$$

Classically this space was understood as the *incidence correspondence of points and lines*:

$$\mathcal{I} = \{(p, \ell) : p \in \ell\}.$$

This point of view reflects an understanding of *projective duality*: the lines in $\mathbb{P}^2$

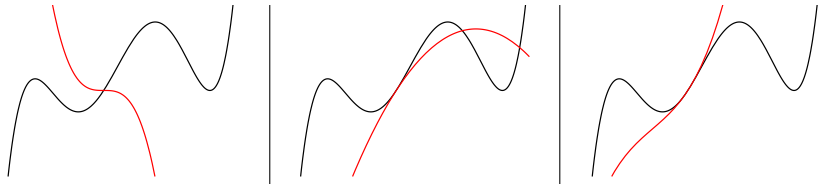
are naturally parametrized by another projective plane $(\mathbb{P}^2)^\vee$, and the incidence correspondence naturally maps to both planes.



Given our algebraic curve C in $\mathbb{P}^2$, we can lift it up to $\mathcal{I}$ and then project it down to $(\mathbb{P}^2)^\vee$, obtaining its *dual curve* $C^\vee$. Under this beautiful duality, the nodes on C correspond to bitangent lines on $C^\vee$, and vice versa. The cusps on C correspond to flexes on $C^\vee$, and vice versa. Furthermore, the class of C is the degree of $C^\vee$, and vice versa. Thus the Plücker formulas record an elaborate interplay between the geometry of the two curves, via the intermediary $\mathcal{I}$.

Contact formulas

To establish our conventions:



Transverse intersection	Tangency	Osculation
1st-order contact	2nd-order contact (Ordinary contact)	3rd-order contact
Data agree to order 0.	Data agree to order 1.	Data agree to order 2.

To justify these conventions, we remark that a tangency can be thought of as what happens when two intersection points coalesce. Alternatively, one can do a calculation in local commutative algebra, such as this one:

$$\langle y - x^2 \rangle \cap \langle y \rangle = \langle y, x^2 \rangle.$$

Example. Bézout's Theorem says that two plane curves of degrees d and e meet in de points. See Figure 28 for an example.

To have a completely correct statement, we need to invoke several provisos:

- We're working over $\mathbb{C}$.
- To capture possible "intersections at infinity," we're working in the projective plane.

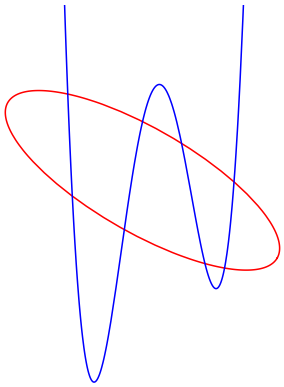


Figure 28: A quartic curve and a conic meeting in 8 points.

- We must assume that the two curves don't have any components in common, e.g., that they don't both contain the same line. (We could utter the magic word "general.")
- If necessary, we count intersections with multiplicity, e.g., a point where the curves are tangent should be counted twice. (We could avoid this by uttering the magic word.)

There is a natural generalization to hypersurfaces in projective space $\mathbb{P}^n$: the intersection of hypersurfaces of degree $d_1, d_2, \ldots, d_n$ consists of $d_1 d_2 \cdots d_n$ points; there are similar (somewhat more elaborate) provisos. This is also called Bézout's Theorem.

Example. Given five general conics in the plane, how many conics are tangent to each of them? The first answer to this was given in 1848 by J. Steiner [37]. The equation of a conic

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

uses six coefficients, and multiplying them all by the same constant preserves the curve. Thus the parameter space for conics is $\mathbb{P}^5$. (More generally, the parameter space for degree d plane curves is $\mathbb{P}^N$, where $N = \binom{d+2}{2} - 1$.) It's pretty easy to see that the condition of being tangent to some specified conic is an equation of degree 6 in A, B, C, D, E, F, i.e., it defines a hypersurface of this degree in $\mathbb{P}^5$. Since there are five specified conics, this gives us five hypersurfaces. According to (the generalized) Bézout's Theorem, they intersect in $6^5 = 7776$ points. Thus the answer to the enumerative problem is 7776.

However, this answer is wrong! The difficulty involves one of the provisos of Bézout's Theorem. The parameter space for conics includes all curves of degree 2. Most conics are nonsingular, but some of them are pairs of lines, and, even worse, some of them are just a single line counted twice, e.g.,

$$x^2 + 2xy + y^2 = 0.$$

If we examine the equation of degree 6 specifying conics tangent to a specified conic, we'll see that all of these double lines satisfy it. Thus those five hypersurfaces don't meet in a finite set of points at all! In fact they meet along the 2-dimensional locus of double lines, and at a finite number of additional points. This problem was pointed out *circa* 1859 by de Jonquières (who did not publish) and in 1864 by Chasles. They also gave the correct count of 3264. We don't want to go into how they did this, except to say—to those who know this notion—that they taught us to use a different parameter space, obtained by blowing up $\mathbb{P}^5$ along the locus of double lines. This clever construction removes all the intersections except those we really want to count.

Example. Let $\mathcal{X} = \{X_s\}$ denote a 2-parameter family of curves in $\mathbb{P}^2$, and let Y be a fixed plane curve of degree d, of class $d^\vee$, and having e flexes. Define the following *characteristic numbers* for the family $\mathcal{X}$:

$$\begin{aligned} M &:= \#\text{ of } X_s \text{ tangent to a specified line at a specified point on it,} \\ K &:= \#\text{ of } X_s \text{ with a specified general point as cusp,} \\ K^\vee &:= \#\text{ of } X_s \text{ with a specified general line as inflectional tangent.} \end{aligned}$$

In 1880, Schubert [34] gave the following formula for the number of triple contacts between Y and members of $\mathcal{X}$:

$$dK^\vee + d^\vee K + (3d + e)M.$$

The formula was verified by Roberts and Speiser in [33], and in [8] we gave an alternative proof that uses the second monster.

When is this formula valid? Hilbert asked the same question, in a much broader way, in the fifteenth of his famous 23 problems of 1900. Hilbert [21] wrote:

> The problem consists in this: To establish rigorously and with an exact determination of the limits of their validity those geometrical numbers which Schubert especially has determined on the basis of the so-called principle of special position, or conservation of number, by means of the enumerative calculus developed by him.

For this particular formula, the required hypotheses are quite mild; one should assume:

- Y doesn't contain a line.
- The general member of $\mathcal{X}$ doesn't contain a line. (Special members may, however.)
- $\mathcal{X}$ and Y are in general position with respect to the action of the group $\mathbb{P}GL(3)$ of projective motions of the plane.

A strategy for proving a contact formula

We're considering a situation involving specified plane curves or families of curves, and asking an enumerative question involving contacts among the curves or members of the family. We're looking for a formula whose output is the answer to the question.

Here's a strategy for finding such a formula:

- The inputs should be numbers measuring some aspect of the curves or families. Following Chasles, we call them *characteristic numbers*. As Fulton, Kleiman, and MacPherson say [17, p. 161], "The word 'characteristic' reflects the fact that these numbers suffice to characterize the family in enumeration of contacts." In practice, the exact sort of characteristic numbers that are needed, and what they really measure, emerge from the next steps in the strategy.
- Construct an appropriate parameter space.
- Develop its *intersection theory*. We'll say more about this in a moment.
- Apply the intersection theory to the curves or families, obtaining a *proto-contact formula*.
- Establish the *enumerative significance* of the inputs and outputs. In other words, explain under what conditions they have their intended meanings. Thus the proto-contact becomes an actual contact formula.

Intersection theory

What is intersection theory? We give a breezy explanation. It's a type of cohomology theory appropriate to the study of an algebraic variety X. The basic objects are certain equivalence classes of subvarieties of X; linear combinations of these objects are called *algebraic cycle classes*. If X is nonsingular, then there is an *intersection ring* A^*X whose product reflects the way in which subvarieties intersect. If V and W are nonsingular subvarieties intersecting transversally, i.e., if wherever they meet their tangent spaces span the entire tangent space of X, then

$$[V] \cdot [W] = [V \cap W].$$

For both intersection theory and enumerative geometry, we recommend the textbook [13]. The standard advanced treatise is [16]. The development of the subject stretches back to the 19th century, and is entwined with the development of homology and cohomology in algebraic topology.

Example. The intersection ring of n-dimensional projective space is

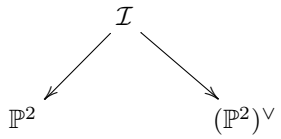
$$A^*(\mathbb{P}^n) = \frac{\mathbb{Z}[h]}{\langle h^{n+1} \rangle},$$

where h is the class of a hyperplane. In this ring, the class of a hypersurface of degree d is dh and the class of a point is h^n. The ring is graded by codimension,

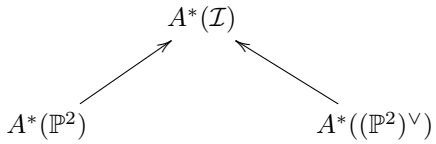
and there is a *degree homomorphism* from the nth graded piece to the integers; it counts, with multiplicity, the number of points, and it's traditionally denoted by $\int$. Here's Bézout's theorem:

$$\int (d_1 h) \cdot (d_2 h) \cdots (d_n h) = d_1 d_2 \cdots d_n.$$

Example. Recall the incidence correspondence of points and lines in the plane:



For the intersection rings, the arrows are reversed: we can "pull back" classes.



Let $h \in A^*(\mathbb{P}^2)$ be the class of a line. Let $h^\vee \in A^*((\mathbb{P}^2)^\vee)$ be the class of all lines through a specified point of $\mathbb{P}^2$ (any point). Then

$$A^*(\mathcal{I}) = \frac{\mathbb{Z}[h, h^\vee]}{\langle h^3, (h^\vee)^3, h^2 - hh^\vee + (h^\vee)^2 \rangle}.$$

In fact, we could omit either of the first two relations, but we prefer to retain them all in order to display the symmetry, which is a consequence of projective duality.

Example. Suppose we have a variety X carrying a rank two bundle $\mathcal{B}$. Let $\mathbb{P}\mathcal{B}$ denote the total space of the projectivization of $\mathcal{B}$. Then

$$A^*(\mathbb{P}\mathcal{B}) \cong \frac{A^*(X)[\varphi]}{\langle \varphi^2 - c_1(\mathcal{B})\varphi + c_2(\mathcal{B}) \rangle}$$

where $c_1(\mathcal{B})$ and $c_2(\mathcal{B})$ are certain cycle classes on X called the *Chern classes* of $\mathcal{B}$. This is part of the theory of characteristic classes, explained in [27] or [13].

Example. Of course this applies to the construction of the monster tower. In all our subsequent examples, we will work with the tower over the projective plane $\mathbb{P}^2$, and we confine our remarks to that case. We denote by i_k the cycle class of the divisor at infinity:

$$i_k = [I_k].$$

From the previous formula, a simple change of variables tells us this:

$$A^*(\mathbb{P}^2(k)) = \frac{A^*(\mathbb{P}^2(k-1))[i_k]}{\langle \text{explicit quadratic in } i_k \rangle}.$$

Thus as one works out the intersection rings of the monster spaces, at each level one needs to adjoin a single new indeterminate satisfying a single new quadratic relation. Table 2 shows our presentation for $A^*(\mathbb{P}^2(4))$. To obtain a presentation for $A^*(\mathbb{P}^2(3))$, simply omit the last row.

Table 2: Presentation for $A^*(\mathbb{P}^2(4))$.

$\mathbb{P}^2$	h	h^3
$\mathbb{P}^2(1)$	$h^\vee$	$(h^\vee)^2 - hh^\vee + h^2$
$\mathbb{P}^2(2)$	i_2	$i_2(i_2 + 3h^\vee - 3h)$
$\mathbb{P}^2(3)$	i_3	$i_3(i_3 + 3i_2 + 4h^\vee - 5h)$
$\mathbb{P}^2(4)$	i_4	$i_4(i_4 + 3i_3 + 4i_2 + 5h^\vee - 7h)$

From these presentations, one can immediately infer the ranks of the graded pieces, for example:

$$\begin{aligned}
\operatorname{rank} A^0(\mathbb{P}^2(3)) &= 1 \\
\operatorname{rank} A^1(\mathbb{P}^2(3)) &= 4 \\
\operatorname{rank} A^2(\mathbb{P}^2(3)) &= 7 \\
\operatorname{rank} A^3(\mathbb{P}^2(3)) &= 7 \\
\operatorname{rank} A^4(\mathbb{P}^2(3)) &= 4 \\
\operatorname{rank} A^5(\mathbb{P}^2(3)) &= 1
\end{aligned}$$

The elements of $A^1(\mathbb{P}^2(3))$ are called *divisor classes*.

A higher-order contact formula

Here we present one formula of our own, taken from our paper [8]. Suppose that Y is a curve in the projective plane. Suppose we have a 3-parameter family of plane curves

$$\mathcal{X} \subset \mathbb{P}^2 \times S \to S;$$

here S is a 3-dimensional variety, and each fiber of $\mathcal{X} \to S$ is a plane curve. We want to know the number of members of $\mathcal{X}$ having quadruple contact with Y.

Now where does one look for quadruple contacts? Our answer is: on $\mathbb{P}^2(3)$, the third monster over the projective plane. This is a nonsingular variety of dimension 5. The curve Y has a lift $Y(3)$. There is a also a lift $\mathcal{X}(3) \to S$; for a general member of the family we just use its lift, but for certain members the

lifts may have additional components. (We will discuss this more carefully at the end of the lecture.) We have

$$\mathcal{X}(3) \subset \mathbb{P}^2(3) \times S \xrightarrow{\sigma} \mathbb{P}^2(3).$$

According to our strategy, we need to work with the intersection ring of $\mathbb{P}^2(3)$. Our aim is to find

$$Q = \int \sigma_*[\mathcal{X}(3)] \cdot [Y(3)],$$

which we feel ought to calculate the number we want.

In addition to the divisor classes h, $h^\vee$, i_2, and i_3, we find it convenient to work with a class that we call z_3: it is the class of the points of $\mathbb{P}^2(3)$ representing the data of lines in the plane. This is a cycle class of codimension 2. The following matrix shows bases for the pieces of the intersection ring in codimensions 1 and 4, together with all their intersection numbers:

	$h^2h^\vee i_2$	$h^2h^\vee i_3$	$h^2 i_2 i_3$	$(h^\vee)^2 z_3$
h	0	0	0	1
$h^\vee$	0	0	1	0
i_2	0	1	-3	0
i_3	1	-3	5	0.

We will need the inverse transpose of this intersection matrix:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 4 & 3 & 1 & 0 \\ 3 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Using the displayed bases, we define characteristic numbers for the curve Y and the family $\mathcal{X}$ in a formal way, as follows. We also indicate their intended geometric meanings:

	intended meaning
$d := \int h[Y(3)]$	degree of Y
$d^\vee := \int h^\vee[Y(3)]$	class of Y
$k_2 := \int i_2[Y(3)]$	# of cusps on Y
$k_3 := \int i_3[Y(3)]$	# of 3rd-order cusps ($y^2 = x^5$) on Y
	# of members of $\mathcal{X}$ having ...
$\gamma_3 := \int h^2 h^\vee i_2 \cdot \sigma_*[\mathcal{X}(3)]$	... cusp at specified point with specified tangent
$\gamma_2 := \int h^2 h^\vee i_3 \cdot \sigma_*[\mathcal{X}(3)]$	... 3rd-order cusp at specified point with specified tangent
$\gamma_1 := \int h^2 i_2 i_3 \cdot \sigma_*[\mathcal{X}(3)]$	... profound cusp ($y^3 = x^5$) at specified point
$\lambda := \int (h^\vee)^2 z_3 \cdot \sigma_*[\mathcal{X}(3)]$	... 3rd-order flex ($y = x^4$) with specified tangent

From these definitions and the inverse transpose matrix, we obtain the proto-contact formula

$$Q = d\lambda + d^\vee\gamma_1 + (3d^\vee + k_2)\gamma_2 + (4d^\vee + 3k_2 + k_3)\gamma_3.$$

To finish, we need to say when the quantities appearing in it have their intended meanings. We have done this quite thoroughly in our cited paper, but we want to avoid the details here. Very roughly, the required assumptions are of these types:

- Y has no line components, nor does the general member of $\mathcal{X}$.
- Restrictions on the types of more complicated singularities, for both Y and the general member of $\mathcal{X}$
- General position with respect to the action of $\mathbb{P}GL(3)$.

Orbits and strata

In the previous section, we defined a *proto-contact number*

$$Q = \int_{\mathbb{P}^2(3)} \sigma_*[\mathcal{X}(3)] \cdot [Y(3)],$$

and claimed that we could elucidate conditions under which it counted the intended contact number, namely the number of members of $\mathcal{X}$ having quadruple contact with Y. Here we want to explain how one analyzes this situation and comes up with appropriate conditions. The basic idea is that $\mathbb{P}GL(3)$ acts on

$\mathbb{P}^2(3)$, and that this action has an open dense orbit. We say that $\mathcal{X}$ and Y are *suitably transverse* if $\sigma(\mathcal{X}(3))$ and $Y(3)$ meet only in the dense orbit, and if this is a transverse intersection. The basic reference for this sort of analysis is [23]; the theory developed there is often called *Kleiman transversality theory.* With this theory in hand, we need to simply carry out dimension counts of the intersections of $\sigma(\mathcal{X}(3))$ and $Y(3)$ with the various orbits.

Table 3, taken mostly from our earlier paper [8, p. 45], describes the eight orbits of $\mathbb{P}^2(3)$ under the action of $\mathbb{P}GL(3)$. Figure 29, which is a repeat of Figure 9, show the representative germ of the orbit $\mathcal{O}(-,\infty)$.

Table 3: The eight $\mathbb{P}GL(3)$-orbits of $\mathbb{P}^2(3)$.

Orbit	Stratum	Dimension	Represented by	Parametrization
$\mathcal{O}(0,0)$	RRR	3	$y=0$	$x=t, y=0$
$\mathcal{O}(0,\infty)$	RRV	3	$y^2=x^5$	$x=t^2, y=t^5$
$\mathcal{O}(0,-)$	RRR	4	$y=x^3$	$x=t, y=t^3$
$\mathcal{O}(\infty,0)$	RVT	3	$y^3=x^4$	$x=t^3, y=t^4$
$\mathcal{O}(\infty,\infty)$	RVV	3	$y^3=x^5$	$x=t^3, y=t^5$
$\mathcal{O}(\infty,-)$	RVR	4	$y^2=x^3$	$x=t^2, y=t^3$
$\mathcal{O}(-,\infty)$	RRV	4	$(y-x^2)^2=x^5$	$x=t^2, y=t^4+t^5$
$\mathcal{O}(-,*)$	RRR	5	$y=x^2$	$x=t, y=t^2$

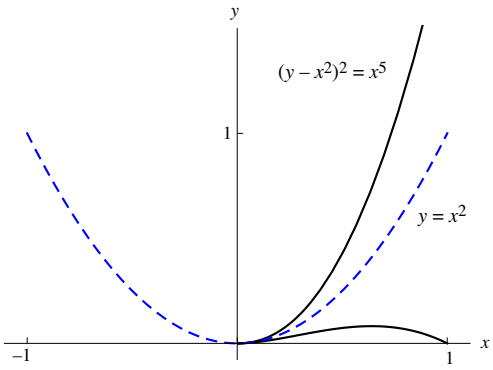
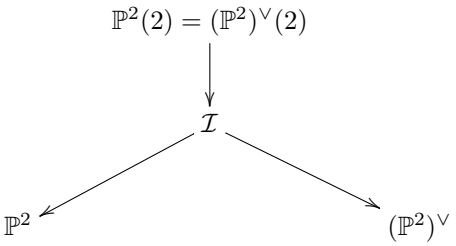


Figure 29: The ramphoid cusp is a representative germ of the orbit $\mathcal{O}(-,\infty)$.

The column on the left uses a code we developed in our enumerative work; it describes a finer stratification than the RVT stratification that we described in Lecture 2. This stratification takes into account not only the divisors at infinity but also the special locus of data of lines in the plane; since $\mathbb{P}GL(3)$ preserves lines, this locus is likewise preserved. Alternatively, this locus can be obtained by observing that the incidence correspondence $\mathcal{I}$ can be interpreted as the first

monster of both $\mathbb{P}^2$ and $(\mathbb{P}^2)^\vee$. The spaces $\mathbb{P}^2(2)$ and $(\mathbb{P}^2)^\vee(2)$ are the same, but the two divisors at infinity are disjoint. In fact the locus of tangency to lines is just the divisor at infinity for the monster over $(\mathbb{P}^2)^\vee$.



If we compare the orbit structure with the strata associated to the two sorts of code words, here's what we find:

Level	# of RVT strata	# of $0-*\infty$ strata	# of $\mathbb{P}GL(3)$ orbits
0	1	1	1
1	1	1	1
2	2	3	3
3	5	8	8
4	13	21	21
5	34	55	≥ 56
6	89	144	∞
7	233	377	∞

Some of the values in the right column were first obtained by Oberlin College students in unpublished work *circa* 1991–1995. Dan Frankowski found the value 21 and we later confirmed it. Since $\mathbb{P}GL(3)$ has dimension 8, there are infinitely many orbits on $\mathbb{P}^2(n)$, none of them dense, when $n \geq 7$. Ian Robertson and Sue Sierra, however, showed that there are infinitely many orbits even on $\mathbb{P}^2(6)$. We believe that the value 56 is correct; the extra orbit comes from the points of $\mathbb{P}^2(5)$ representing the data of nonsingular points of conics. These form the *sextactic locus*, this classical terminology stems from [6].

Prolongation in families

Understanding how to lift (prolong) a family of curves is a subtle part of the analysis, and we'd like to finish by saying a bit more about this topic. If you consider a singular curve all by itself, the recipe for lifting says: lift at all nonsingular points, and then take the closure. Trying this out on the pair of coordinate axes $xy = 0$, we find that the lift is disconnected: it's the disjoint union of two lines:

$$\begin{cases} y = 0 \\ y' = 0 \end{cases} \quad \text{(in one chart)} \quad \bigcup \quad \begin{cases} x = 0 \\ x' = 0 \end{cases} \quad \text{(in another chart)}$$

But if this curve is the central member of the family $\mathcal{X}$ consisting of the curves $xy = t$, one should lift it first by lifting all nonsingular members and

then taking the closure: this gives the two components we've already seen, together with a third component

$$\begin{cases} x = 0 \\ y = 0 \end{cases}$$

representing the data of all directions over the origin. This component appears because, on any nearby curve of the family, we do see all possible directions (except the strictly horizontal and vertical). Note that the lift of the central member fits into a nice family $\mathcal{X}(1)$ of lifts, whose other members are just the usual lifts of the other curves in the family.

For precise work in algebraic geometry, one would like to demand even more: one wants the family to be *flat*. This is a technical condition which gives the best analogue of a continuous family in topology, but it carries more information: each member of the family is a *scheme*, meaning that locally it is cut out by equations. The middle component of the lift of $xy = 0$ within the family $xy = t$ is cut out by these equations:

$$\begin{cases} x^2 = 0 \\ xy = 0 \\ y^2 = 0 \\ xy' + y = 0. \end{cases}$$

As a point set it's just the line $x = y = 0$, but the elements x and y aren't in the ideal we've just described, so as a scheme it's a "thickened line," and in fact one can measure its thickness; it's 2. The fact that the family is flat means that when we do intersection calculations with members of the family, the result doesn't suddenly change when $t = 0$. For example, suppose one intersects the members of the family with the class $h^\vee$ (represented by the set of all lifts of lines through one specified point). For a nonsingular member $xy = t \neq 0$ one obtains 2, which we have called the *class* of the curve; it's the number of tangent lines passing through the specified point. For the central member $xy = 0$ the intersection occurs on the middle component, and it counts twice.

This kind of thinking is what led us to our most recent published project [10], an analysis of the lifts of this family to a an arbitrary level k of the monster tower. We give an explicit recursive recipe for deriving all the equations of $\mathcal{X}(k)$ in all relevant charts of the monster, and show how they cut out a chain of $2^k + 1$ component curves, each of which is a projective line; we call them *twigs*. Figure 30 depicts $\mathcal{X}(3)$, showing the multiplicities (thicknesses) of each twig. The end twigs are the lifts of the two individual axes; the twig right in the middle is the lift of the component we saw at level 1.

The thesis of Swaminathan [38] carried out a similar project, as applied to a single-branched curve with local parametrization

$$\begin{cases} x = f(t) \\ y = g(t). \end{cases}$$

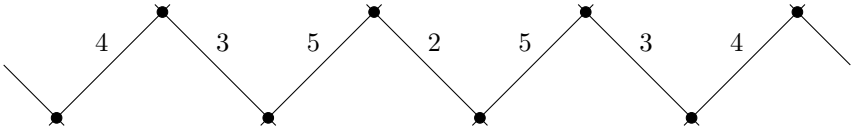


Figure 30: The third prolongation of the nodal curve.

This was simply a topological analysis, and took no account of scheme structures or flat families. The main result is this: one gets the same result as one does by blowing up at an appropriate sequence of points. Let us note that one is comparing objects that live in different ambient spaces: the prolongation (iterated Nash blow-up) naturally lives in the monster space, whereas the curve obtained by blowing up still lives in a surface, namely the plane blown up at the same series of points. We intend to repeat this refined analysis for more complicated curves in the context of algebraic geometry.

References

[1] A. M. Bloch. *Nonholonomic mechanics and control*, volume 24 of *Interdisciplinary Applied Mathematics*. Springer, New York, second edition, 2015. With the collaboration of J. Bailieul, P. E. Crouch, J. E. Marsden and D. Zenkov, With scientific input from P. S. Krishnaprasad and R. M. Murray.

[2] Alejandro Bravo-Doddoli and Luis C. García-Naranjo. The dynamics of an articulated n-trailer vehicle. *Regul. Chaotic Dyn.*, 20(5):497–517, 2015.

[3] E. Cartan. Sur l'équivalence absolue de certains systèmes d'équations différentielles et sur certaines familles de courbes. *Bull. Soc. Math. France*, 42:12–48, 1914.

[4] Alex Castro, Susan Jane Colley, Gary Kennedy, and Corey Shanbrom. A coarse stratification of the monster tower. *Michigan Math. J.*, 66(4):855–866, 2017.

[5] Alex L. Castro. *Chains and monsters: From Cauchy-Riemann geometry to Semple towers and singular space curves*. ProQuest LLC, Ann Arbor, MI, 2010. Thesis (Ph.D.)–University of California, Santa Cruz.

[6] Arthur Cayley. On the sextactic points of a plane curve. *Philosophical Transactions of the Royal Society of London*, 155(3):545–578, 1865.

[7] Susan Jane Colley and Gary Kennedy. A higher-order contact formula for plane curves. *Comm. Algebra*, 19(2):479–508, 1991.

[8] Susan Jane Colley and Gary Kennedy. Triple and quadruple contact of plane curves. In *Enumerative algebraic geometry (Copenhagen, 1989)*, volume 123 of *Contemp. Math.*, pages 31–59. Amer. Math. Soc., Providence, RI, 1991.

[9] Susan Jane Colley and Gary Kennedy. The enumeration of simultaneous higher-order contacts between plane curves. *Compositio Math.*, 93(2):171–209, 1994.

[10] Susan Jane Colley and Gary Kennedy. Cartan prolongation of a family of curves acquiring a node. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 14:Paper No. 031, 16, 2018.

[11] Alberto Collino. Evidence for a conjecture of Ellingsrud and Strømme on the Chow ring of $\mathrm{Hilb}_d(\mathbf{P}^2)$. *Illinois J. Math.*, 32(2):171–210, 1988.

[12] Jean-Pierre Demailly. Algebraic criteria for Kobayashi hyperbolic projective varieties and jet differentials. In *Algebraic geometry—Santa Cruz 1995*, volume 62 of *Proc. Sympos. Pure Math.*, pages 285–360. Amer. Math. Soc., Providence, RI, 1997.

[13] David Eisenbud and Joe Harris. *3264 and all that—a second course in algebraic geometry*. Cambridge University Press, Cambridge, 2016.

[14] F. Engel. Zur invariantentheorie der systeme von pfaffschen gleichungen. *Berichte Verhandlungen der Koniglich Sachsischen Gesellschaft der Wissenschaften Mathematisch- Physikalische Klasse*, 41, 1889.

[15] Robert Foote, Mark Levi, and Serge Tabachnikov. Tractrices, bicycle tire tracks, hatchet planimeters, and a 100-year-old conjecture. *Amer. Math. Monthly*, 120(3):199–216, 2013.

[16] William Fulton. *Intersection theory*, volume 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, second edition, 1998.

[17] William Fulton, Steven Kleiman, and Robert MacPherson. About the enumeration of contacts. In *Algebraic geometry—open problems (Ravello, 1982)*, volume 997 of *Lecture Notes in Math.*, pages 156–196. Springer, Berlin, 1983.

[18] Giuseppe Gherardelli. Sul modello minimo della varietà degli elementi differenziali del 2° ordine del piano projettivo. *Atti Accad. Italia. Rend. Cl. Sci. Fis. Mat. Nat. (7)*, 2:821–828, 1941.

[19] G.-H. Halphen. Sur les invariants différentiels des courbes gauches. In *Oeuvres de G.-H. Halphen, Tome II*, pages 353–446. Gauthier-Villars, Paris, 1918.

[20] G.-H. Halphen. Sur les invariants différentiels (thèse présentée à la faculté des sciences de paris; 1878). In *Oeuvres de G.-H. Halphen, Tome II*, pages 197–253. Gauthier-Villars, Paris, 1918.

[21] David Hilbert. Mathematical problems. *Bull. Amer. Math. Soc. (N.S.)*, 37(4):407–436, 2000. Reprinted from Bull. Amer. Math. Soc. **8** (1902), 437–479.

[22] Frédéric Jean. The car with n trailers: characterisation of the singular configurations. *ESAIM Contrôle Optim. Calc. Var.*, 1:241–266, 1995/96.

[23] Steven L. Kleiman. The transversality of a general translate. *Compositio Math.*, 28:287–297, 1974.

[24] Steven L. Kleiman. The enumerative theory of singularities. In *Real and complex singularities (Proc. Ninth Nordic Summer School/NAVF Sympos. Math., Oslo, 1976)*, pages 297–396. Sijthoff and Noordhoff, Alphen aan den Rijn, 1977.

[25] Monique Lejeune-Jalabert. Chains of points in the Semple tower. *Amer. J. Math.*, 128(5):1283–1311, 2006.

[26] Markus Mauhart and Peter W. Michor. Commutators of flows and fields. *Arch. Math. (Brno)*, 28(3-4):229–236, 1992.

[27] John W. Milnor and James D. Stasheff. *Characteristic classes*. Princeton University Press, Princeton, N. J.; University of Tokyo Press, Tokyo, 1974. Annals of Mathematics Studies, No. 76.

[28] Richard Montgomery and Michail Zhitomirskii. Geometric approach to Goursat flags. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 18(4):459–493, 2001.

[29] Richard Montgomery and Michail Zhitomirskii. Points and curves in the Monster tower. *Mem. Amer. Math. Soc.*, 203(956):x+137, 2010.

[30] Piotr Mormul. Geometric classes of Goursat flags and the arithmetics of their encoding by small growth vectors. *Cent. Eur. J. Math.*, 2(5):859–883, 2004.

[31] Piotr Mormul. Exotic moduli of Goursat distributions exist already in codimension three. In *Real and complex singularities*, volume 459 of *Contemp. Math.*, pages 131–145. Amer. Math. Soc., Providence, RI, 2008.

[32] Jean-jacques Risler and Florian Luca. The maximum of the degree of nonholonomy for the car with trailers. *IFAC Proceedings Volumes*, 27, 12 1993.

[33] Joel Roberts and Robert Speiser. Enumerative geometry of triangles. II. *Comm. Algebra*, 14(1):155–191, 1986.

[34] H. Schubert. Anzahlgeometrische Behandlung des Dreiecks. *Math. Ann.*, 17(2):153–212, 1880.

[35] J. G. Semple. Some investigations in the geometry of curve and surface elements. *Proc. London Math. Soc. (3)*, 4:24–49, 1954.

[36] Corey Shanbrom. The Puiseux characteristic of a Goursat germ. *J. Dyn. Control Syst.*, 20(1):33–46, 2014.

[37] J. Steiner. Elementare lösung einer geometrischen aufgabe, und über einige damit in beziehung stehende eigenschaften der kegelschnitte. *J. Reine Angew. Math.*, 37:161–192, 1848.

[38] Vidya Swaminathan. *A comparison of two methods of resolution: Blow up and prolongation*. ProQuest LLC, Ann Arbor, MI, 2009. Thesis (Ph.D.)–University of California, Santa Cruz.