A Coarse Stratification of the Monster Tower

ALEX CASTRO, SUSAN JANE COLLEY,
GARY KENNEDY, & COREY SHANBROM

ABSTRACT. The monster tower is a tower of spaces over a specified base; each space in the tower is a parameter space for curvilinear data up to a specified order. We describe and analyze a natural stratification of these spaces.

The monster tower, also known in the algebro-geometric literature as the Semple tower, is a tower of smooth spaces (varieties) over a specified smooth base M. Each space $M(k)$ in the tower, called a *monster space,* is a parameter space for curvilinear data up to order k on M. We will describe a coarse stratification of each monster space, with each stratum corresponding to a code word created out of a certain alphabet according to rules that we will specify. These strata parameterize curvilinear data "of the same type." The monster space can be regarded as an especially nice compactification of the parameter space for curvilinear data of nonsingular curves on M (as explained in Section 2), with the added points representing the data of singular curves; in our stratification the nonsingular data points form a single open dense stratum.

Versions of this coarse stratification have been observed by virtually everyone who has studied the monster construction. Here we develop the theory in full generality, beginning with a base space of arbitrary dimension and at all levels. A finer stratification would result from a thorough analysis of the orbits under the action of a suitable group acting on the base (whose action can be lifted to the tower) or, working locally, of the pseudogroup of local diffeomorphisms at a selected point. Results such as those in Section 5.7 of [12] show that we can expect there to be infinitely many strata, that is, that there are moduli. However, we seem to be very far from a full understanding of where and why moduli occur.

In Section 1, we recall the construction of the monster tower and selectively review prior literature. A brief Section 2 explains curvilinear data and baby monsters. Section 3 introduces code words for labeling the strata, which are explained in Section 4 via their closures, called *intersection loci.* Our main Theorem 4.2 gives an explicit description of these loci. To prove the theorem, we use coordinates on charts, as explained in Section 5; after this, the proof in Section 6 is nearly immediate. Finally, in Section 7, we count the strata.

1. The Monster Tower

Suppose that M is a smooth manifold, complex manifold, or nonsingular algebraic variety over a field of characteristic zero; denote its dimension by m. The *monster tower* is a sequence

$$\cdots \to M(k) \to M(k-1) \to \cdots \to M(2) \to M(1) \to M(0) = M \qquad (1.1)$$

in which each $M(k) \to M(k-1)$ is a fiber bundle with fiber $\mathbb{P}^{m-1}$. To define it, we begin with a more general construction.

The general construction begins with a pair $(X, \mathcal{B})$, where X is again a smooth manifold, complex manifold, or variety, and $\mathcal{B}$ is a rank b subbundle of its tangent bundle TX. Let $\widetilde{X} = \mathbb{P}\mathcal{B}$, and let $\pi : \widetilde{X} \to X$ be the projection. Then the tautological line bundle $\mathcal{O}_{\mathcal{B}}(-1)$ on $\widetilde{X}$ is a subbundle of $\pi^* TX$. Let

$$d\pi : T\widetilde{X} \to \pi^* TX$$

denote the derivative map of π. The pullback of $\mathcal{O}_{\mathcal{B}}(-1)$ by $d\pi$ is a subbundle of $T\widetilde{X}$, which we denote by $\widetilde{\mathcal{B}}$. In other words, a point P of $\widetilde{X}$ represents a tangent direction at a point of X, and $\widetilde{\mathcal{B}}$ is the subbundle of $T\widetilde{X}$ whose fiber at P consists of vectors mapping (via the derivative of projection) to vectors in that direction; we call them *focal vectors*. Note that the relative tangent bundle $T(\widetilde{X}/X)$ is a subbundle of $\widetilde{\mathcal{B}}$; its fiber consists of vectors mapping to zero, sometimes called *vertical vectors*. By construction, $\widetilde{\mathcal{B}}$ is a subbundle of $T\widetilde{X}$, and again its rank is b. Thus we can iterate this construction to obtain a tower of fibrations.

We will eventually apply this construction in several situations. To construct the monster tower, we simply apply it to the pair (M, TM) and then iterate. We denote the resulting spaces as in (1.1); $M(k)$ is called the *monster space at level k* or simply the *kth monster*. 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 $TM(k)$. For $k \geq 2$, the projectivization of the relative tangent bundle $\mathbb{P}T(M(k)/M(k-1))$ gives us a divisor on $M(k)$, which we call the *divisor at infinity* and denote by I_k. The pullback of I_k to any higher level is again a divisor, and for simplicity of notation, we again just denote it by I_k.

The earliest instance of the construction seems to be Gherardelli's paper [8]. The tower was explained by Semple [13; 14], and in the algebro-geometric literature, it bears his name. Two of the present authors used it to study problems of enumerative geometry in [4; 5; 6], and it was treated in greater generality in [10]. Demailly [7] used it to study positivity questions for hyperbolic varieties.

Working independently, a group of differential geometers studied the same construction using different techniques, language, and motivation. In this strand of literature the general construction is called *Cartan prolongation,* and the resulting tower is called the monster tower. The tower with base $\mathbb{R}^2$ was discovered by Montgomery and Zhitomirskii [11] in their study of singular Goursat distributions. They discovered connections with singular plane curves and the control-theoretic problem of a car pulling many trailers. Their detailed study of the monster tower with base $\mathbb{R}^2$ appears as the monograph [12]. Subsequent investigations of the monster tower with bases $\mathbb{R}^2, \mathbb{R}^3$, and $\mathbb{C}^3$ appear in [1; 2; 3; 15]. These

efforts were generally aimed at understanding the action of the diffeomorphism group on the tower: constructing invariants and counting and classifying orbits.

The equivalence of the two towers was first noticed by Castro in 2010. The present contribution represents the first collaboration between the two groups, and an effort to improve and standardize language and notation, most notably in the coding system presented in Section 3.

2. Curvilinear Data and Baby Monsters

Here we informally recall the concept of curvilinear data; for further information, see the works cited at the end of the previous section.

Again, let M be a smooth manifold, complex manifold, or nonsingular algebraic variety of dimension m. Suppose that 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. For each curve, assume that the restriction of the differential dx_1 does not 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. These are $(m-1)k$ conditions, and we can check that they are independent of local coordinates.

Thus a *nonsingular curvilinear datum* is a point in a manifold or smooth variety of dimension $m + (m-1)k$, and there is a tower of such manifolds over the base M, with fiber a projective space $\mathbb{P}^{m-1}$ at the first level, and then affine space fibers $\mathbb{A}^{m-1}$ thereafter. At the first level the manifold is exactly $M(1)$, the projectivization of the tangent bundle TM. With each point of the curve C in M, we can associate the point of $M(1)$ recording its tangent direction, and in this way, we obtain a copy of C inside $M(1)$, called the *lift* of C and denoted by $C(1)$; the process of passing from C to $C(1)$ is again called *prolongation*.

We can lift again to obtain a curve $C(2)$ inside the projectivization of the tangent bundle $TM(2)$. Observe that a tangent vector at a point of $C(1)$ is a focal vector; thus, in fact, $C(2)$ lies entirely inside the smaller space $M(2)$. Proceeding similarly, we obtain a copy $C(k)$ of C inside the monster space $M(k)$, called its *kth lift* or *kth prolongation*. At every stage, these focal vectors are nonvertical vectors; thus $C(k)$ avoids the divisor at infinity $I(k)$ and all pullbacks of prior divisors at infinity $C(2), \ldots, C(k-1)$. Conversely, given any point of $M(k)$ away from all divisors at infinity, there is a recipe for finding a curve $\widetilde{C}$ passing through this point and then, by a process of integration, producing a curve C on M for which $C(k) = \widetilde{C}$. Thus the monster spaces $M(k)$ are spaces that naturally compactify the spaces of nonsingular curvilinear data. As the reader undoubtedly suspects, a point on a divisor at infinity represents the data of some sort of singular curves; the process of lifting is essentially repeated Nash blowup, performed simultaneously at every point of the curve. The added points are also called *curvilinear data*, dropping the modifier "nonsingular." This is explained more carefully in the cited literature.

Inside the monster space $M(k)$, suppose that there is a submanifold X for which the intersection of its tangent bundle TX with the focal bundle Δ_k (inside

the tangent bundle of $M(k)$ restricted to X) has constant rank. Then we can apply the monster prolongation construction to the pair $(X, TX \cap \Delta_k)$; we call the resulting tower the *baby monster tower* associated with X.

Here are three basic examples:

1. If C is a smooth curve in M and TC is its tangent bundle, then applying the baby monster construction to the pair (C, TC) produces a tower of copies of C, namely the lifts $C(k)$ in $M(k)$, as just explained.
2. If X is the fiber over a point of $M(k-1)$, then TX is already a subbundle of Δ_k. The resulting tower is what Castro and Montgomery [3] have also called a baby monster. (Thus our terminology generalizes theirs.)
3. If X is the divisor at infinity I_k on $M(k)$, then the intersection of TX and Δ_k is transverse; thus $TX \cap \Delta_k$ has rank $m-1$. This tower is described by Lejeune-Jalabert [10, p. 1287]. Denote the spaces in this tower by $I_k[n]$, beginning with $I_k[0] = I_k$; thus $I_k[n]$ is a subspace of $M(k+n)$ of codimension $n+1$. We use the same notation to denote the full inverse image of $I_k[n]$ in any higher monster space in the tower.

3. Code Words

We now introduce the code words to be used for labeling strata; the strata themselves are described later. The alphabet for our code consists of all symbols V_A, where A is a finite subset of the integers strictly greater than 1. To avoid a cumbersome notation, we write, for example, abc or a, b, c rather than $\{a, b, c\}$, and always arrange the elements of A in increasing order. Although this is an 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 use R instead of $V_\emptyset$. (The symbols R and V are 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, we 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. Also note that rule (3) is the only rule to use the specified value for m. The diagram in Figure 1 shows the code words of lengths 1, 2, and 3, assuming that $m \geq 3$. If $m > 3$, then there are twenty-four valid code words of length four. If, however, $m = 3$, then $RV_2V_{23}V_{234}$ is not a valid code word, since it violates rule (3).

Given a code word $V_{A_1}V_{A_2}V_{A_3} \cdots V_{A_k}$, for each $j = 2, 3, \ldots, k$, let n_j denote the number of times that j appears as a subscript. (If it does, then its first appearance is in position j, and its last appearance is in position $j + n_j - 1$.) For example, for the code word $RV_2V_{23}V_{23}V_{25}V_5V_5V_5$, we have $n_2 = 4$, $n_3 = 2$, $n_4 = 0$, and $n_5 = 4$. We note that $n_j \leq k+1-j$; the base dimension m adds

$$
R
\begin{cases}
RR
\begin{cases}
RRR \\
RRV_3
\end{cases} \\[2em]
RV_2
\begin{cases}
RV_2R \\
RV_2V_2 \\
RV_2V_3 \\
RV_2V_{23}
\end{cases}
\end{cases}
$$

Figure 1   Code words up to length 3

further restrictions. From a valid set of values for n_2 through n_k we can recover the code word.

In [3] and [2], an alternative coding system was developed for the case $m = 3$, called *RVT coding.* In this system, we work with the seven-symbol alphabet $\{R, V, T_1, T_2, L_1, L_2, L_3\}$, which corresponds to the seven cases considered by Semple [13, p. 151]. The rules for creating a code word are as follows:

- The first symbol must be R.
- The next symbol after R can be either R or V.
- The next symbol after V or T_1 can be R, V, T_1, or L_1.
- The next symbol after T_2 can be R, V, T_2, or L_3.
- The next symbol after L_1, L_2, or L_3 can be any symbol.

The symbol R plays the same role in both the RVT code and our new code. In the correspondence between the two codes, V, T_1, and T_2 always correspond to singly subscripted symbols, but the precise correspondence is conditioned by prior symbols of the code word; similarly, L_1, L_2, and L_3 correspond to doubly subscripted symbols. The precise correspondence is explained by these tables, in which the first row refers to the RVT code, and the second row to our new code.

kth symbol, immediately after R	R	V
kth symbol, immediately after R	R	V_k

kth symbol, after V or T_1	R	V	T_1	L_1
kth symbol, after V_j	R	V_k	V_j	V_{jk}

kth symbol, after T_2	R	V	T_2	L_3
kth symbol, after V_j	R	V_k	V_j	V_{jk}

kth symbol, after L_1, L_2, or L_3	R	V	T_1	T_2	L_1	L_2	L_3
kth symbol, after V_{ij}	R	V_k	V_j	V_i	V_{jk}	V_{ij}	V_{ik}

It is straightforward to use these tables to translate an RVT code word to a new code word. Here is one example:

$$RVT_1VVT_1T_1T_1L_1 \longleftrightarrow RV_2V_2V_4V_5V_5V_5V_5V_{59}.$$

In the other direction, when translating the symbol of a new code word immediately following some V_j, we must be careful to choose the correct table. For example, although $RV_2V_{23}V_2V_2$ and $RV_2V_2V_2V_2$ end with the same pair of symbols, the final symbols of their corresponding RVT code words differ:

$$RV_2V_{23}V_2V_2 \longleftrightarrow RVL_1T_2T_2,$$
$$RV_2V_2V_2V_2 \longleftrightarrow RVT_1T_1T_1.$$

4. Intersection Loci and Strata

We now give a recipe for converting a code word $V_{A_1}V_{A_2}V_{A_3}\cdots V_{A_k}$ to a description of an intersection locus. Recall that, for each number $j = 2, 3, \ldots, k$, we let n_j denote the number of times that j appears as a subscript. If j never appears, then set $n_j = -1$. To state Definition 4.1 cleanly, we adopt the following convention: $I_j[-1]$ is interpreted as the entire monster space $M(k)$.

DEFINITION 4.1. On $M(k)$, the intersection locus corresponding to the code word $W = RV_{A_2}V_{A_3}\cdots V_{A_k}$ is

$$I_W := \bigcap_{j=2}^{k} I_j[n_j - 1]. \tag{4.1}$$

THEOREM 4.2. *For each code word, intersection* (4.1) *is transverse and nonempty. The codimension of the intersection locus is the sum of cardinalities* $|A_2| + \cdots + |A_k|$ *or, equivalently,* $n_2 + \cdots + n_k$. *It contains the intersection locus corresponding to the code word* W' *if and only if each* $n'_j \geq n_j$.

In Section 6, we prove Theorem 4.2 by an explicit calculation using coordinates to be introduced in Section 5.

EXAMPLE 4.3. For the code word consisting entirely of Rs, the intersection locus is $M(k)$ itself.

EXAMPLE 4.4. Assume that $m \geq 3$. On the monster $M(8)$, the codimension-7 intersection locus corresponding to the code word $RV_2V_3V_{34}V_{35}V_3RR$ is the transverse intersection

$$I_2[0] \cap I_3[3] \cap I_4[0] \cap I_5[0],$$

where $I_j[0]$ means the full inverse image of I_j, and $I_3[3]$ denotes the inverse image of the third space in the baby monster tower over I_3.

Using Theorem 4.2, we obtain a natural stratification of the monster space $M(k)$ as follows: from each intersection locus I_W excise all those intersection loci $I_{W'}$ that it contains. In fact, it suffices just to excise each locus whose code word W' is obtained from W by increasing a single n_j by one (ignoring those W' that are not valid code words).

5. Coordinate Charts

Here we describe a natural system of coordinate charts on the monster spaces. The notation is basically that of Lejeune-Jalabert [10], except for the conveniently redundant names. Note that a similar system was developed from the differential geometry side in [9] (see [12] for more details).

We begin with an open set U on M with coordinates $x_1, \ldots, x_m$ for which, at each point, the differentials $dx_1, \ldots, dx_m$ form a basis of the cotangent space. Then over U the monster space $M(1)$ is isomorphic to $U \times \mathbb{P}^{m-1}$, and it is covered by m charts $\mathcal{C}(1), \ldots, \mathcal{C}(m)$, each isomorphic to $U \times \mathbb{A}^{m-1}$. The chart $\mathcal{C}(p)$ represents one-dimensional quotients of the cotangent bundle on which dx_p does not vanish, and the coordinates for the second factor are defined by

$$x_q(p) := \frac{dx_q}{dx_p}$$

for $q \neq p$. For convenience, we also define

$$x_p(p) := x_p$$

(i.e., we give a new name to the pullback of this coordinate function). Note that at each point in this chart the differentials $dx_1(p), \ldots, dx_m(p)$ form a basis for the dual focal bundle $\Delta_1^\vee$. For this reason, the coordinates $x_q(p)$ are called *active coordinates;* the active coordinate $x_p(p)$ is also called the *retained coordinate.*

This is the beginning of a recursive construction of a system of m^k charts $\mathcal{C}(p_1 p_2 \cdots p_k)$ on $M(k)$. Each chart is isomorphic to

$$\mathcal{C}(p_1 p_2 \cdots p_{k-1}) \times \mathbb{A}^{m-1},$$

and there are m *active coordinates* $x_q(p_1 p_2 \cdots p_k)$, where $1 \leq q \leq m$. The *retained (active) coordinate* is

$$x_{p_k}(p_1 p_2 \cdots p_k) := x_{p_k}(p_1 p_2 \cdots p_{k-1})$$

(the pullback of a coordinate from below, given a new name); the others are defined by

$$x_q(p_1 p_2 \cdots p_k) := \frac{dx_q(p_1 p_2 \cdots p_{k-1})}{dx_{p_k}(p_1 p_2 \cdots p_{k-1})}$$

and serve as affine coordinates for the second factor $\mathbb{A}^{m-1}$. The chart represents one-dimensional quotients of $\Delta_{k-1}^\vee$ on which the differential of the retained coordinate does not vanish. We easily verify that, at each point, the differentials of the active coordinates form a basis for $\Delta_k^\vee$.

Thus, for each chart, we have shown how to systematically construct $(k+1)m$ coordinate names

$$x_1, \ldots, x_m,$$
$$x_1(p_1), \ldots, x_m(p_1),$$
$$x_1(p_1 p_2), \ldots, x_m(p_1 p_2),$$
$$\vdots$$
$$x_1(p_1 p_2 \cdots p_k), \ldots, x_m(p_1 p_2 \cdots p_k).$$

Because of redundant names, there are in fact $m + k(m-1)$ distinct coordinates. For a coordinate with redundant names, there is a *shortest name,* characterized by the fact that its subscript differs from the final symbol appearing in parentheses. (The names $x_1, \ldots, x_m$ are always shortest names.)

EXAMPLE 5.1. Assuming that $m = 3$, in chart $\mathcal{C}(32123)$ on $M(5)$, we have the following coordinate names:

x_1	x_2	x_3
$x_1(3)$	$x_2(3)$	$\boxed{x_3(3)}$
$x_1(32)$	$\boxed{x_2(32)}$	$x_3(32)$
$\boxed{x_1(321)}$	$x_2(321)$	$x_3(321)$
$x_1(3212)$	$\boxed{x_2(3212)}$	$x_3(3212)$
$x_1(32123)$	$x_2(32123)$	$\boxed{x_3(32123)}$

All names are shortest names except those in the boxes. Here $x_3(3212)$ is the retained active coordinate.

We now identify, in each chart $\mathcal{C}(p_1 p_2 \cdots p_k)$, the equations for the loci appearing in Theorem 4.2. The divisor at infinity I_j first appears on $M(j)$. It represents one-dimensional quotients of $\Delta^\vee_{j-1}$ in which the differential of every coordinate pulled back from $M(j-2)$ vanishes. There are two possibilities. If $p_{j-1} = p_j$, then the retained coordinate is pulled back from $M(j-2)$; thus I_j does not meet the chart $\mathcal{C}(p_1 p_2 \cdots p_j)$. If $p_{j-1} \neq p_j$, then we claim that the differential of every coordinate pulled back from $M(j-2)$ vanishes if and only if

$$x_{p_{j-1}}(p_1 p_2 \cdots p_j) = 0. \tag{5.1}$$

Indeed, this equation is satisfied if and only if the differential of the previous retained coordinate vanishes, that is,

$$dx_{p_{j-1}}(p_1 p_2 \cdots p_{j-1}) = 0,$$

and in $\Delta^\vee_{j-1}$ the differentials of all inactive coordinates are multiples of this differential.

Thus on the divisor at infinity I_j the differential $dx_{p_{j-1}}(p_1 p_2 \cdots p_j)$ vanishes. When we prolong, there are again two cases. If $p_{j-1} = p_{j+1}$, then again $I_j[1]$

does not meet the chart $\mathcal{C}(p_1 p_2 \cdots p_{j+1})$. If $p_{j-1} \neq p_{j+1}$, then when we prolong, we have

$$x_{p_{j-1}}(p_1 p_2 \cdots p_j p_{j+1}) = 0, \tag{5.2}$$

so that $I_j[1]$ is defined by both (5.1) and (5.2). Continuing in this manner, we see that $I_j[n_j - 1]$ is defined by the vanishing of n_j coordinates

$$x_{p_{j-1}}(p_1 p_2 \cdots p_j), \quad x_{p_{j-1}}(p_1 p_2 \cdots p_{j+1}), \quad \ldots, \quad x_{p_{j-1}}(p_1 p_2 \cdots p_{j+n_j-1}),$$

assuming that the symbol p_{j-1} does not appear in the string $p_j \cdots p_{j+n_j-1}$; if it does, then $I_j[n_j - 1]$ does not meet the chart.

EXAMPLE 5.2. Suppose $m \geq 3$ and consider the code word $R V_2 V_2 V_{24} R$, for which $n_2 = 3$, $n_3 = 0$, $n_4 = 1$, and $n_5 = 0$. It represents the stratum on $M(5)$ whose closure is the intersection locus $I_2[2] \cap I_4$; to obtain the stratum, we need to excise these smaller intersection loci:

$$I_2[3] \cap I_4, \quad I_2[2] \cap I_3 \cap I_4, \quad I_2[2] \cap I_4[1], \quad I_2[2] \cap I_4 \cap I_5.$$

This stratum has codimension 4.

Now assume that $m = 3$. Here are the equations of $I_2[2] \cap I_4$ in the chart $\mathcal{C}(32123)$ using the coordinate names shown in Example 5.1:

$$x_3(32) = 0,$$
$$x_3(321) = 0,$$
$$x_3(3212) = 0,$$
$$x_1(3212) = 0.$$

The first three equations define $I_2[2]$, whereas the fourth defines I_4. The following table displays the additional equation needed to define each of the excised intersection loci:

$I_2[3] \cap I_4$	does not meet the chart,
$I_2[2] \cap I_3 \cap I_4$	$x_2(321) = 0,$
$I_2[2] \cap I_4[1]$	$x_1(32123) = 0,$
$I_2[2] \cap I_4 \cap I_5$	$x_2(32123) = 0.$

6. Proof of the Main Theorem

We now prove Theorem 4.2. If the intersection locus

$$\bigcap_j I_j[n_j - 1] \tag{6.1}$$

meets a chart $\mathcal{C}(p_1 p_2 \cdots p_k)$, then (as we have explained in Section 5) it is defined by the vanishing of $\sum_{j=2}^{k} n_j$ coordinates, noting that the names we have given these coordinates are all shortest names and so there are no repetitions. Thus the intersection is transverse, and the codimension of the intersection locus is as

claimed. These explicit equations also make clear the claim about when one inter-section locus contains another. The stratum for the code word $V_{A_1} V_{A_2} V_{A_3} \cdots V_{A_k}$ is obtained from (6.1) by removing all the smaller intersection loci.

Finally, we argue that each intersection locus, and hence each stratum, is nonempty. Indeed, suppose that $V_{A_1} V_{A_2} V_{A_3} \cdots V_{A_k}$ is a valid code word. Then the chart $\mathcal{C}(p_1 p_2 \cdots p_k)$ will meet the intersection locus (6.1) if and only if each subscript p_j avoids a certain subset of the previous subscripts, and this subset has the cardinality of A_j. Since this cardinality is less than m, we can always choose such a p_j.

7. Counting Code Words

In this section, we explain how to count the number of code words. As before, fix the dimension of the base space to be m. Let $N(k, r)$ be the number of code words of length k in which the last subscript has length r. Assume that $k \geq 2$ and that $1 \leq r \leq k - 1$. Then

$$N(k, r) = \sum_{i=r-1}^{m-1} \binom{i+1}{r} N(k-1, i).$$

Indeed, to obtain such a code word, take any code word of length $k - 1$ in which the last subscript A has length $i \geq r - 1$; then create the new subscript B by choosing any r symbols from the set $A \cup \{k\}$. Clearly,

$$N(k, 0) = \sum_{i=0}^{m-1} N(k-1, i) = \text{number of code words of length } k - 1.$$

To use these recursive formulas, begin with $N(1, 0) = 1$ and $N(1, r) = 0$ for $r > 0$.

We claim that if $k \leq m$, then $N(k, r)$ equals the unsigned Stirling number of the first kind $c(k, r + 1)$. One way to establish this is to remark that these Stirling numbers satisfy the same recurrence just established for the number of code words.

Alternatively, we can establish a bijection between the set of code words of length k and the set of all rooted trees on the vertex set $\{0, 1, \ldots, k\}$ in which 0 is the root and the labels increase as we move away from the root (i.e., *increasing trees*). Given a code word, recall that n_j denotes the number of times that j appears as a subscript. If the code word has length k, then create a tree with vertices labeled by the integers $1, 2, \ldots, k + 1$, with root at $k + 1$. For each integer $j = 2, \ldots, k$, draw an edge connecting $j - 1$ and $j + n_j$; also draw an edge connecting k and $k + 1$. Note that the degree of the root is $r + 1$ and that labels increase as we move toward the root. Replacing each label j by $k + 1 - j$, we obtain an increasing tree. The process clearly can be reversed; thus we have the desired bijection. By Proposition 1.5.5(b) of [16], $N(k, r) = c(k, r + 1)$.

References

[1] A. Castro and W. Howard, *A Monster tower approach to Goursat multi-flags,* Differential Geom. Appl. 30 (2012), 405–427.

[2] A. Castro, W. Howard, and C. Shanbrom, *Stratification of the Monster/Semple tower under diffeomorphisms of three-space,* 2015, arXiv:1512.00107.

[3] A. Castro, R. Montgomery, and appendix by W. Howard, *Spatial curve singularities and the Monster/Semple tower,* Israel J. Math. 192 (2012), 381–427.

[4] S. J. Colley and G. Kennedy, *A higher-order contact formula for plane curves,* Comm. Algebra 12 (1991), no. 2, 479–508.

[5] ⸻ , *Triple and quadruple contact of plane curves,* Enumerative algebraic geometry, Copenhagen, 1989, Contemp. Math., 123, pp. 31–59, 1991.

[6] ⸻ , *The enumeration of simultaneous higher-order contacts between plane curves,* Compos. Math. 93 (1994), 171–209.

[7] J. P. Demailly, *Algebraic criteria for Kobayashi hyperbolic projective varieties and jet differentials,* Algebraic geometry—Santa, Cruz, 1995, Proc. Sympos. Pure Math., 62, pp. 285–360, 1997.

[8] G. 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 (1941), 821–828.

[9] A. Giaro, A. Kumpera, and C. Ruiz, *Sur la lecture correcte d'un résultat d'Élie Cartan,* C. R. Math. Acad. Sci. Paris 287 (1978), 241–244.

[10] M. Lejeune-Jalabert, *Chains of points in the Semple tower,* Amer. J. Math. 128 (2006), no. 5, 1283–1311.

[11] R. Montgomery and M. Zhitomirskii, *Geometric approach to Goursat flags,* Ann. Inst. H. Poincaré Anal. Non Linéaire 18 (2001), 459–493.

[12] ⸻ , *Points and curves in the Monster tower,* Mem. Am. Math. Soc., 956, 2010.

[13] J. G. Semple, *Singularities of space algebraic curves,* Proc. Lond. Math. Soc. 44 (1938), no. 2, 149–174.

[14] ⸻ , *Some investigations in the geometry of curve and surface elements,* Proc. Lond. Math. Soc. 4 (1954), no. 3, 24–49.

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

[16] R. Stanley, *Enumerative combinatorics, volume I,* second edition, Cambridge University Press, New York, 2012.

A. Castro
OVO Energy
33 Notting Hill Gate
London W11 3JQ
UK

alex.decastro@ovoenergy.com

S. J. Colley
Department of Mathematics
Oberlin College
Oberlin, Ohio 44074
USA

sjcolley@math.oberlin.edu

G. Kennedy
Ohio State University at Mansfield
1760 University Drive
Mansfield, Ohio 44906
USA

kennedy@math.ohio-state.edu

C. Shanbrom
California State University
6000 J St.
Sacramento, CA 95819
USA

corey.shanbrom@csus.edu