Some Geometrical Aspects of Control Points for Toric Patches

Gheorghe Craciun[1], Luis David García-Puente[2], and Frank Sottile[3⋆]

[1] Department of Mathematics and Department of Biomolecular Chemistry, University of Wisconsin, Madison, WI 53706, USA, craciun@math.wisc.edu, www.math.wisc.edu/~craciun

[2] Department of Mathematics and Statistics, Sam Houston State University, Huntsville, TX 77341, USA, lgarcia@shsu.edu, www.shsu.edu/~ldg005

[3] Department of Mathematics, Texas A&M University, College Station, TX 77843, USA, sottile@math.tamu.edu www.math.tamu.edu/~sottile

Abstract. We use ideas from algebraic geometry and dynamical systems to explain some ways that control points influence the shape of a Bézier curve or patch. In particular, we establish a generalization of Birch's Theorem and use it to deduce sufficient conditions on the control points for a patch to be injective. We also explain a way that the control points influence the shape via degenerations to regular control polytopes. The natural objects of this investigation are irrational patches, which are a generalization of Krasauskas's toric patches, and include Bézier and tensor product patches as important special cases.

Introduction

The control points and weights of a Bézier curve, Bézier patch, or tensor-product patch govern many aspects of the curve or surface. For example, they provide an intuitive means to control its shape. Through de Castlejau's algorithm, they enable the computation of the curve or surface patch. Finer aspects of the patch, particularly continuity and smoothness at the boundary of two patches are determined by the control points and weights. Global properties, such as the location of a patch in space due to the convex hull property, also depend upon the control points. When the control points are in a particular convex position, then the patch is convex [1].

We apply methods from algebraic geometry, specifically toric geometry, to explain how some further global properties of a patch are governed by the control points. We first investigate the self-intersection, or injectivity of a patch. We give a simple and easy-to-verify condition on a set of control points which implies that the resulting patch has no self-intersection, for any choice of weights. For 3-dimensional patches as used for solid modeling, injectivity is equivalent to the

⋆ This material is based in part upon work supported by the Texas Advanced Research Program under Grant No. 010366-0054-2007 and NSF grant DMS-070105.

patch properly parameterizing the given solid. This uses Craciun and Feinberg's injectivity theorem [2] from the theory of chemical reaction networks, which may be seen as a generalization of Birch's Theorem from algebraic statistics.

A second global property that we investigate is how the shape of the patch is related to the shape of a *control polytope*. This is a piecewise linear triangulated surface whose vertices are the control points. It is *regular* if the underlying triangulation comes from a regular triangulation of the domain polytope of the patch. We show that regular control polytopes are the limits of patches as the weights undergo a toric deformation corresponding to the underlying regular triangulation, and that non-regular control polytopes can never be such a limit. This gives a precise meaning to the notion that the shape of the control net governs the shape of the patch.

This line of inquiry is pursued in terms of Krasauskas's toric patches [3], as it relies upon the structure of toric varieties from algebraic geometry. The correct level of generality is however that of *irrational (toric) patches*, which are analytic subvarieties of the simplex (realized as a compactified positive orthant) that are parameterized by monomials x^α, where x is a vector of positive numbers and the exponent vector α has real-number coordinates. (This is a usual toric variety when α has integer coordinates.) While irrational patches may seem exotic for modeling, they occur naturally in statistics as discrete exponential families [4] and their blending functions may be computed using iterative proportional fitting (IPF) [5], a popular numerical algorithm from statistics. Furthermore, these blending functions have linear precision. For toric patches, this was observed in [6] and developed in [7], and the analysis there carries over to irrational patches. While we work in this generality, our primary intent (and the main application) is to shed light on properties of Bézier curves, surfaces, and 3-dimensional patches.

We recall the standard definition of a mapping via control points and blending functions, and then the definitions of toric Bézier patches in Sect. 1. There, we also illustrate some of our results on examples of Bézier curves. In Sect. 2, we introduce irrational toric patches, recalling the geometric formulation of a toric patch and the use of iterative proportional fitting to compute these patches, explaining how these notions from [7] for toric patches extend to irrational patches. The next two sections contain our main results. We study injectivity of patches in Sect. 3, and discuss degenerations to control polytopes in Sect. 4. Appendices A and B contain technical proofs of some theorems.

1 Toric Bézier Patches

We interpret the standard definition of a mapping via control points and blending functions (see for example [8, §2]) in a general form convenient for our discussion. All functions here are smooth (C^∞) where defined and real-valued. Let $\mathbb{R}_>$ be the set of strictly positive real numbers and $\mathbb{R}_\geq$ the set of non-negative real numbers. We will use the following typographic conventions throughout. Vector constants (control points, indexing exponents, and standard basis vectors) will

be typeset in **boldface**, while vector variables will be typeset in standard math italics.

Let $\mathcal{A}$ be a finite set of points that affinely span $\mathbb{R}^d$, which we shall use as geometrically meaningful indices. A control point scheme for parametric patches, or *(parametric) patch*, is a collection $\beta = \{\beta_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}\}$ of non-negative functions, called *blending functions*. The common domain of the blending functions is the convex hull Δ of $\mathcal{A}$, which we call the *domain polytope*. We also assume that the blending functions do not vanish simultaneously at any point of Δ, so that there are no basepoints.

Lists $\mathcal{B} := \{\mathbf{b}_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}\} \subset \mathbb{R}^n$ of *control points* and positive *weights* $w := \{w_{\mathbf{a}} \in \mathbb{R}_> \mid \mathbf{a} \in \mathcal{A}\} \in \mathbb{R}_>^{\mathcal{A}}$ together give a map $F\colon \Delta \to \mathbb{R}^n$ defined by

$$F(x) \;:=\; \frac{\sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta_{\mathbf{a}}(x)\,\mathbf{b}_{\mathbf{a}}}{\sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta_{\mathbf{a}}(x)} \quad . \tag{1}$$

The denominator in (1) is positive on Δ and so the map F is well-defined.

Remark 1. We will refer to both $\beta_{\mathbf{a}}(x)$ and $w_{\mathbf{a}}\beta_{\mathbf{a}}(x)$ as the blending functions of a patch. This generality of separating the weights from the blending functions will be used in Sect. 4 when we investigate the effect of systematically varying the weights of a patch while keeping the control points and blending functions constant.

The control points and weights affect the shape of the patch which is the image of the map F (1). For example, the *convex hull property* asserts that the image $F(\Delta)$ of the patch lies in the convex hull of the control points. To see this, note that if we set

$$\overline{\beta}_{\mathbf{a}}(x) \;:=\; \frac{w_{\mathbf{a}}\beta_{\mathbf{a}}(x)}{\sum_{\mathbf{a}\in\mathcal{A}} w_{\mathbf{a}}\beta_{\mathbf{a}}(x)} \quad ,$$

then $\overline{\beta}_{\mathbf{a}}(x) \geq 0$ and $1 = \sum_{\mathbf{a}\in\mathcal{A}} \overline{\beta}_{\mathbf{a}}(x)$. Then formula (1) becomes

$$F(x) \;=\; \sum_{\mathbf{a}\in\mathcal{A}} \overline{\beta}_{\mathbf{a}}(x)\,\mathbf{b}_{\mathbf{a}} \quad ,$$

so that $F(x)$ is a convex combination of the control points and therefore lies in their convex hull. In fact, if there is a point $x \in \Delta$ at which no blending function vanishes, then *any* point in the interior of the convex hull of the control points is the image $F(x)$ of some patch for some choice of weights. In this way, the convex hull property is the strongest general statement that can be made about the location of a patch.

Another well-known manifestation of control points is the relation of a Bézier curve to its control polygon. Fix a positive integer m and let $\mathcal{A} := \{\frac{i}{m} \mid i = 0, \ldots, m\}$ so that Δ is the unit interval. The blending functions of a Bézier curve are the Bernstein polynomials,

$$\beta_i(x)\ (=\ \beta_{\frac{i}{m}}(x)) \;:=\; \tbinom{m}{i} x^i (1-x)^{m-i} \quad .$$

The *control polygon* of a Bézier curve with control points $\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_m$ is the union of the line segments $\overline{\mathbf{b}_0, \mathbf{b}_1}$, $\overline{\mathbf{b}_1, \mathbf{b}_2}$, $\ldots$, $\overline{\mathbf{b}_{m-1}, \mathbf{b}_m}$ between consecutive control points. Figure 1 displays two quintic plane Bézier curves with their control polygons (solid lines). The convex hulls of the control points are indicated

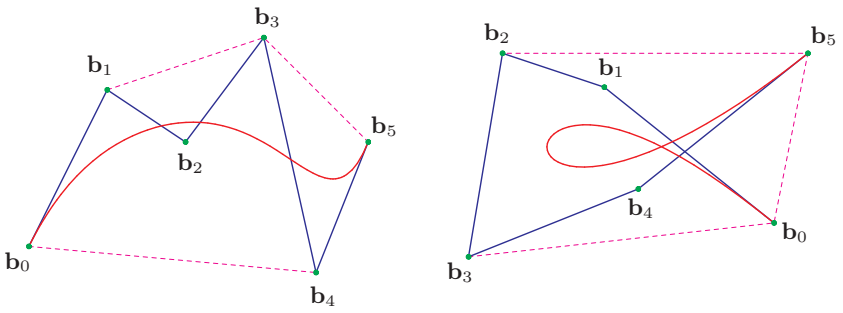

Fig. 1. Quintic Bézier curves.

by the dashed lines. The first curve has no points of self-intersection, while the second curve has one point of self-intersection. While this self-intersection may be removed by varying the weights attached to the control points, by Theorem 11 it is impossible to find weights so that a curve with the first set of control points has a point of self-intersection.

We will also show that the control polygon may be approximated by a Bézier curve. We state a simplified version of Theorem 13 from Sect. 4.

Theorem. *Given control points in $\mathbb{R}^n$ for a Bézier curve and some number $\epsilon > 0$, there is a choice of weights so that the image $F[0,1]$ of the Bézier curve lies within a distance ϵ of the control polygon.*

In Figure 2, we display one of the quintic curves from Figure 1, but with weights on $\mathbf{b}_0$—$\mathbf{b}_5$ of $(1, 20^2, 20^3, 20^3, 20^2, 1)$ and $(1, 300^2, 300^3, 300^3, 300^2, 1)$, respectively. The control polygon for the second curve is omitted, as it would obscure the curve. The first curve lies within a distance $\epsilon = 0.13$ of the control polygon and the second within a distance $\epsilon = 0.02$, if the control polygon has height 1.

1.1 Toric Patches

Krasauskas [3] introduced toric patches as a generalization of the classical Bézier and tensor product patches. These are based upon toric varieties from algebraic geometry and their shape may be any polytope with integer vertices. The articles [9, 6] provide an introduction to toric varieties for geometric modeling.

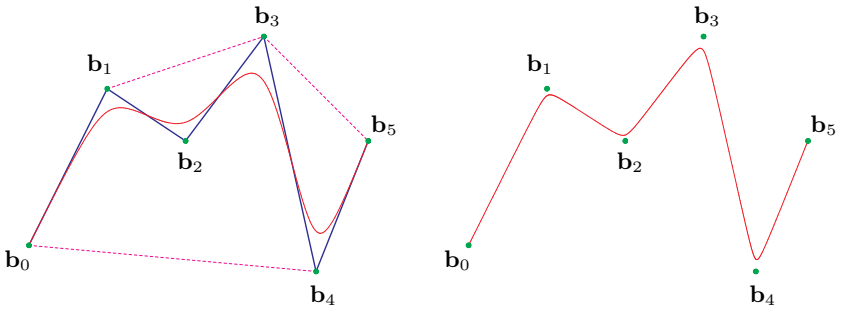

Fig. 2. Degenerating quintics.

A polytope Δ is defined by its *facet inequalities*

$$\Delta \;=\; \{x \in \mathbb{R}^d \mid 0 \le h_i(x)\,, i = 1, \ldots, \ell\} \;\;.$$

Here, Δ has ℓ facets (faces of maximal dimension) and for each $i = 1, \ldots, \ell$, $h_i(x) = \mathbf{v}_i \cdot x + c_i$ is the linear function defining the ith facet, where $\mathbf{v}_i \in \mathbb{Z}^d$ is the (inward oriented) primitive vector normal to the facet and $c_i \in \mathbb{Z}$.

For example, if our polytope is the triangle with vertices $(0,0)$, $(m,0)$, and $(0,m)$,

$$m\triangle \;:=\; \{(x,y) \;\in\; \mathbb{R}^2 \mid 0 \le x, y, \quad \text{and} \quad 0 \le m - (x+y)\} \;\;, \tag{2}$$

then we have $h_1 = x$, $h_2 = y$, and $h_3 = m - x - y$. Here, $m\triangle$ is the unit triangle $\triangle$ with vertices $(0,0)$, $(1,0)$, and $(0,1)$ scaled by a factor of m.

Let $\mathcal{A} \subset \Delta \cap \mathbb{Z}^d$ be any subset of the integer points of Δ which includes its vertices. For every $\mathbf{a} \in \mathcal{A}$, Krasauskas defined the *toric Bézier function*

$$\beta_{\mathbf{a}}(x) \;:=\; h_1(x)^{h_1(\mathbf{a})} h_2(x)^{h_2(\mathbf{a})} \cdots h_\ell(x)^{h_\ell(\mathbf{a})} \;\;, \tag{3}$$

which is non-negative on Δ, and the collection of all $\beta_{\mathbf{a}}$ has no common zeroes on Δ. These are blending functions for the *toric patch* of shape $\mathcal{A}$. If we choose weights $w \in \mathbb{R}^{\mathcal{A}}$ and multiply the formula (3) by $w_{\mathbf{a}}$, we obtain blending functions for the toric patch of shape $(\mathcal{A}, w)$.

Example 2 (Bézier triangles). When Δ is a scaled triangle or a product of such triangles, (3) gives the blending functions of the Bézier patch or Bézier simploid [10] with the corresponding shape. To see this for the scaled triangle $m\triangle$ (2), note that given an integer point $\mathbf{a} = (i,j) \in m\triangle$, and weight the multinomial coefficient $w_{(i,j)} := \frac{m!}{i!j!(m-i-j)!}$, then the corresponding blending function is

$$\beta_{(i,j)}(x,y) \;=\; \tfrac{m!}{i!j!(m-i-j)!} x^i y^j (m - x - y)^{m-i-j} \;\;.$$

This is almost the bivariate Bernstein polynomial, which is obtained by substituting mx and my for x and y, respectively, and dividing by m^m. (This has the effect of changing the domain from $m\triangle$ to the unit triangle $\triangle$.)

2 Irrational Patches

Krasauskas's definition (3) of toric Bézier functions still makes sense if we relax the requirement that the points $\mathcal{A} \subset \mathbb{R}^d$ have integer coordinates. This leads to the notion of an irrational patch (as its blending functions are no longer rational functions), which provides the level of generality appropriate for our investigation.

Let $\mathcal{A} \subset \mathbb{R}^d$ be a finite collection of points and set $\Delta \subset \mathbb{R}^d$ to be the convex hull of $\mathcal{A}$, which we assume is a full-dimensional polytope. We may also realize Δ as an intersection of half-spaces through its facet inequalities,

$$\Delta \;=\; \{x \in \mathbb{R}^d \mid h_i(x) \geq 0 \text{ for each } i = 1, \ldots, \ell\} \;. \tag{4}$$

Here, Δ has ℓ facets with the ith facet supported by the affine hyperplane $h_i(x) = 0$ where $h_i(x) = \mathbf{v}_i \cdot x + c_i$ with $c_i \in \mathbb{R}$ and $\mathbf{v}_i$ *an* inward pointing normal vector to the ith facet of Δ. There is no canonical choice for these data; multiplying a pair $(\mathbf{v}_i, c_i)$ by a positive scalar gives another pair defining the same half-space.

Following Krasauskas, we provisionally define *(irrational) toric Bézier functions* $\{\beta_{\mathbf{a}} \colon \Delta \to \mathbb{R}_{\geq} \mid \mathbf{a} \in \mathcal{A}\}$ by the same formula as (3),

$$\beta_{\mathbf{a}}(x) \;:=\; h_1(x)^{h_1(\mathbf{a})} h_2(x)^{h_2(\mathbf{a})} \cdots h_\ell(x)^{h_\ell(\mathbf{a})} \;.$$

These are blending functions for the *irrational toric patch* of shape $\mathcal{A}$. While these functions do depend upon the choice of data $(\mathbf{v}_i, c_i)$ for the facet inequalities defining Δ, we will see that the image $F(\Delta)$ of such a patch given by weights and control points is independent of these choices.

2.1 Geometric Formulation of a Patch

We follow Sect. 2.2 of [7], but drop the requirement that our objects are algebraic. Let $\mathcal{A} \subset \mathbb{R}^d$ be a finite subset indexing a collection of blending functions $\{\beta_{\mathbf{a}} \colon \Delta \to \mathbb{R}_{\geq} \mid \mathbf{a} \in \mathcal{A}\}$, where Δ is the convex hull of $\mathcal{A}$. Let $\mathbb{R}^{\mathcal{A}}$ be a real vector space with basis $\{e_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}\}$. Set $\mathbb{R}^{\mathcal{A}}_{\geq} \subset \mathbb{R}^{\mathcal{A}}$ to be the points with non-negative coordinates and let $\mathbb{R}^{\mathcal{A}}_{>}$ be those points with strictly positive coordinates.

For $z = (z_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}) \in \mathbb{R}^{\mathcal{A}}_{\geq}$, set $\sum z := \sum_{\mathbf{a} \in \mathcal{A}} z_{\mathbf{a}}$. The *$\mathcal{A}$-simplex*, $\Delta^{\mathcal{A}} \subset \mathbb{R}^{\mathcal{A}}_{\geq}$, is the set

$$\Delta^{\mathcal{A}} \;:=\; \{z \;\in\; \mathbb{R}^{\mathcal{A}}_{\geq} : \textstyle\sum z = 1\} \;.$$

We introduce homogeneous coordinates for $\Delta^{\mathcal{A}}$. If $z \in \mathbb{R}^{\mathcal{A}}_{\geq} \setminus \{0\}$, then we set

$$[z_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}] \;:=\; \frac{1}{\sum z}(z_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}) \;\in\; \Delta^{\mathcal{A}} \;.$$

The blending functions $\{\beta_{\mathbf{a}} \colon \Delta \to \mathbb{R}_{\geq} \mid \mathbf{a} \in \mathcal{A}\}$ give a C^∞ map,

$$\begin{aligned} \beta \;:\; \Delta &\longrightarrow \Delta^{\mathcal{A}} \\ x &\longmapsto [\beta_{\mathbf{a}}(x) \mid \mathbf{a} \in \mathcal{A}] \,. \end{aligned}$$

The reason for this definition is that a mapping $F\colon \Delta \to \mathbb{R}^n$ (1) given by the blending functions β, weights w, and control points $\mathcal{B}$ factors through the map $\beta\colon \Delta \to \Delta^{\mathcal{A}}$. To see this, first note that the weights $w \in \mathbb{R}^{\mathcal{A}}_{>}$ act on $\Delta^{\mathcal{A}}$: If $z = [z_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}] \in \Delta^{\mathcal{A}}$, then

$$w.z \;:=\; [w_{\mathbf{a}} z_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}] \;. \tag{5}$$

The control points $\mathcal{B}$ define the map $\pi_{\mathcal{B}}\colon \Delta^{\mathcal{A}} \to \mathbb{R}^n$ via

$$\pi_{\mathcal{B}} \;:\; z \;=\; (z_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}) \;\longmapsto\; \sum_{\mathbf{a}\in\mathcal{A}} z_{\mathbf{a}} \mathbf{b}_{\mathbf{a}} \;.$$

Then the mapping F (1) is simply the composition

$$\Delta \;\xrightarrow{\;\beta\;}\; \Delta^{\mathcal{A}} \;\xrightarrow{\;w.\;}\; \Delta^{\mathcal{A}} \;\xrightarrow{\;\pi_{\mathcal{B}}\;}\; \mathbb{R}^n \;. \tag{6}$$

In this way, we see that the image $\beta(\Delta) \subset \Delta^{\mathcal{A}}$ of Δ under the map β determines the shape of the patch $F(\Delta)$ (1). Internal structures of the patch, such as the mapping of texture, are determined by how β maps Δ to $\beta(\Delta)$. For example, precomposing β with any homeomorphism of Δ gives blending functions with the same image in $\Delta^{\mathcal{A}}$, but with a different internal structure.

For an irrational toric patch of shape $\mathcal{A}$ with blending functions (3), the image $\beta(\Delta) \subset \Delta^{\mathcal{A}}$ is independent of the choice of normal vectors. For this, we first define the map $\varphi_{\mathcal{A}}\colon \mathbb{R}^d_{>} \to \Delta^{\mathcal{A}}$ by

$$\varphi_{\mathcal{A}} \;:\; (x_1, \ldots, x_d) \;\longmapsto\; [x^{\mathbf{a}} \;:\; \mathbf{a} \in \mathcal{A}] \;. \tag{7}$$

Let $X_{\mathcal{A}}$ be the closure of the image of the map $\varphi_{\mathcal{A}}$. When $\mathcal{A} \subset \mathbb{Z}^d$, this is the positive part [11, §4] of the toric variety parameterized by the monomials of $\mathcal{A}$. When $\mathcal{A}$ is not integral, we call $X_{\mathcal{A}}$ the *(irrational) toric variety* parameterized by monomials in $\mathcal{A}$.

In Appendix B we prove the following theorem.

Theorem 3. *Suppose that $\mathcal{A} \subset \mathbb{R}^d$ is a finite set of points with convex hull Δ. Let $\beta = \{\beta_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}\}$ be a collection of irrational toric Bézier functions for $\mathcal{A}$. Then $\beta(\Delta) = X_{\mathcal{A}}$, the closure of the image of $\varphi_{\mathcal{A}}$.*

We prove this by showing that the restriction of the map β to the interior Δ° of Δ factors through the map $\varphi_{\mathcal{A}}$.

By Theorem 3, the image of the irrational toric blending functions for $\mathcal{A}$ depends upon $\mathcal{A}$ and not upon the choice of toric blending functions for $\mathcal{A}$. Thus the shape of the corresponding patch $F(\Delta)$ (1) depends only upon $\mathcal{A}$, the weights w, and the control points $\mathcal{B}$. However, the actual parameterization of $X_{\mathcal{A}}$ by Δ, and hence of $F(\Delta)$ does depend upon the choice of toric blending functions for $\mathcal{A}$.

To ensure that the patch $F(\Delta)$ has shape reflecting that of Δ, we require that the map $\beta\colon \Delta \to X_{\mathcal{A}}$ be injective. This also guarantees that the patch $F(\Delta)$ is typically an immersion. In the context of irrational toric patches, this

injectivity is guaranteed by Birch's Theorem from algebraic statistics. For a standard reference, see [12, p. 168]. When $\mathcal{A} \subset \mathbb{Z}^d$, Birch's Theorem follows from general results on the moment map in symplectic geometry [11, §4.2 and Notes to Chapter 4, p. 140].

Theorem 4 (Birch's Theorem). *Suppose $\mathcal{A} \subset \mathbb{R}^d$ is finite and let β be a collection of toric Bézier functions. If we choose control points $\mathcal{B}$ to be the corresponding points of $\mathcal{A}$, $\{\mathbf{b_a} = \mathbf{a} \mid \mathbf{a} \in \mathcal{A}\}$, then the composition*

$$\Delta \xrightarrow{\beta} X_{\mathcal{A}} \xrightarrow{\pi_{\mathcal{B}}} \mathbb{R}^d$$

is a homeomorphism onto Δ.

By Birch's Theorem and Theorem 3, any two sets β, β' of toric Bézier functions of shape $\mathcal{A}$ differ only by a homeomorphism $h\colon \Delta \xrightarrow{\sim} \Delta$ of the polytope Δ, so that $\beta' = \beta \circ h$. In fact h restricts to a homeomorphism on all faces of Δ. As we are concerned with the shape of a patch and not its internal structure, we follow Krasauskas' lead and make the following definition.

Definition 5. *A(n irrational) toric patch of shape $\mathcal{A}$ is any set of blending functions $\beta := \{\beta_{\mathbf{a}}\colon \Delta \to \mathbb{R}_{\geq} \mid \mathbf{a} \in \mathcal{A}\}$ such that the map $\beta\colon \Delta \to X_{\mathcal{A}}$ is a homeomorphism.*

The projection map $\pi_{\mathcal{B}}\colon \Delta^{\mathcal{A}} \to \mathbb{R}^d$ appearing in Birch's Theorem induced by the choice $\mathcal{B} = \{\mathbf{b_a} = \mathbf{a} \mid \mathbf{a} \in \mathcal{A}\}$ of control points is called the *tautological projection* and written $\pi_{\mathcal{A}}$. Restricting the tautological projection to $X_{\mathcal{A}}$ gives the *algebraic moment map* $\mu\colon X_{\mathcal{A}} \to \Delta$. The components of its inverse $\mu^{-1}\colon \Delta \xrightarrow{\sim} X_{\mathcal{A}}$ provide a preferred set of blending functions for the patch. When $\mathcal{A} \subset \mathbb{Z}^d$, these were studied in [7, 6], where they were shown to have linear precision, and that they may be computed by *iterative proportional fitting (IPF)*, a numerical algorithm from statistics [5]. These same arguments apply to irrational toric patches—the preferred blending functions have linear precision and are computed by IPF.

Any patch has unique blending functions with linear precision [7, Theorem 1.11]. While the classification of toric patches for which these preferred blending functions are rational functions remains open in general, it has been settled for surface patches ($d = 2$) [13], and this places very strong restrictions on higher-dimensional patches.

2.2 Iterative Proportional Fitting for Toric Patches

In algebraic statistics, $\Delta^{\mathcal{A}}$ is identified with the probability simplex parameterizing probability distributions on data indexed by $\mathcal{A}$. The image $X_{\mathcal{A},w} := w.X_{\mathcal{A}}$ of $\mathbb{R}^d_{>}$ under the map $\varphi_{\mathcal{A}}$ (7) and translation by w (5) is known as a *toric model* [14, §1.2]. It is more common to call this a *log-linear model*, as the logarithms of the coordinates of $\varphi_{\mathcal{A}}$ are linear functions in the logarithms of the

coordinates of $\mathbb{R}^d_>$, or a *discrete exponential family* as the coordinates of $\varphi_{\mathcal{A}}$ are exponentials in the logarithms of the coordinates of $\mathbb{R}^d_>$.

The tautological map appears in statistics as follows. Given (observed) normalized data $q \in \Delta^{\mathcal{A}}$, the problem of *maximum likelihood estimation* asks for a probability distribution in the toric model, $p \in X_{\mathcal{A},w}$, with the same sufficient statistics as q, $\mu(p) = \pi_{\mathcal{A}}(p) = \pi_{\mathcal{A}}(q)$. By Birch's Theorem, the point $p \in X_{\mathcal{A},w}$ is unique and hence

$$p \;=\; \mu^{-1}(\pi_{\mathcal{A}}(q)) \;\; .$$

Thus inverting the tautological projection is necessary for maximum likelihood estimation.

Darroch and Ratcliff [5] introduced the numerical algorithm of iterative proportional fitting, also known as generalized iterative scaling, for computing the inverse μ^{-1} of the tautological projection. We now describe their algorithm.

Observe first that the toric patch $X_{\mathcal{A},w}$ does not change if we translate all elements of $\mathcal{A}$ by a fixed vector $\mathbf{b}$, $(\mathbf{a} \mapsto \mathbf{a}+\mathbf{b})$, so we may assume that $\mathcal{A}$ lies in the positive orthant $\mathbb{R}^d_>$. Scaling the exponent vectors in $\mathcal{A}$ by a fixed positive scalar $t \in \mathbb{R}_>$ also does not change $X_{\mathcal{A},w}$ as $x \mapsto x^t$ is a homeomorphism of $\mathbb{R}_>$ which extends to a homeomorphism of $\mathbb{R}^d_>$. Thus we may assume that $\mathcal{A}$ lies in the standard simplex Δ_d in $\mathbb{R}^d$,

$$\Delta_d \;=\; \{x \in \mathbb{R}^d_{\geq} \mid \textstyle\sum x \leq 1\} \;\; .$$

Lastly, we lift this to the probability simplex $\Delta^{d+1} \subset \mathbb{R}^{d+1}_{\geq}$,

$$\Delta^{d+1} \;:=\; \{y \in \mathbb{R}^{d+1}_{\geq} \mid \textstyle\sum y = 1\} \;\; ,$$

by

$$\mathcal{A} \ni \mathbf{a} \;\longmapsto\; \mathbf{a}^+ := (1-\textstyle\sum \mathbf{a}, \mathbf{a}) \in \Delta^{d+1} \;\; .$$

Since for $t \in \mathbb{R}_>$ and $x \in \mathbb{R}^d_>$,

$$(t, tx)^{\mathbf{a}^+} \;=\; t^{1-\sum \mathbf{a}}(tx)^{\mathbf{a}} \;=\; tx^{\mathbf{a}} \;\; ,$$

we see that replacing $\mathcal{A}$ by this homogeneous version also does not change $X_{\mathcal{A}}$.

We describe the algorithm of *iterative proportional fitting*, which is Theorem 1 in [5].

Proposition 6. *Suppose that $\mathcal{A} \subset \Delta^{d+1}$ has convex hull Δ and $q \in \Delta^{\mathcal{A}}$. Set $y := \pi_{\mathcal{A}}(q) \in \Delta$. Then the sequence of points*

$$\{p^{(m)} \;\mid\; m = 0, 1, 2 \ldots\} \;\subset\; \Delta^{\mathcal{A}}$$

whose $\mathbf{a}$-coordinates are defined by $p^{(0)}_{\mathbf{a}} := w_{\mathbf{a}}$ and, for $m \geq 0$,

$$p^{(m+1)}_{\mathbf{a}} \;:=\; p^{(m)}_{\mathbf{a}} \cdot \frac{y^{\mathbf{a}}}{(\pi_{\mathcal{A}}(p^{(m)}))^{\mathbf{a}}} \;\; ,$$

converges to the unique point $p \in X_{\mathcal{A},w}$ such that $\mu(p) = \pi_{\mathcal{A}}(p) = \pi_{\mathcal{A}}(q) = y$.

We remark that if $\mathcal{A}$ is not homogenized then to compute $\mu^{-1}(y)$ for $y \in \Delta$, we first put $\mathcal{A}$ into homogeneous form using an affine map ψ, and then use iterative proportional fitting to compute $\pi_{\mathcal{A}^+}^{-1}(\psi(y)) = \pi_{\mathcal{A}}^{-1}(y)$. We also call this modification of the algorithm of Proposition 6 iterative proportional fitting. Thus iterative proportional fitting computes the inverse image of the tautological projection.

3 Injectivity of Patches

Birch's Theorem (Theorem 4) states that for one particular choice of control points, namely $\{\mathbf{b}_\mathbf{a} := \mathbf{a} \mid \mathbf{a} \in \mathcal{A}\}$ and all weights $w \in \mathbb{R}_>^\mathcal{A}$, the mapping F (1) of a toric patch of shape $(\mathcal{A}, w)$ is a homeomorphism onto its image. From this, we can infer that for *most* choices of control points and weights, this mapping is at least an immersion. To study dynamical systems arising from chemical reaction networks, Craciun and Feinberg [2] prove an injectivity theorem for certain maps, which we adapt to generalize Birch's Theorem. This will give conditions on control points $\mathcal{B} \subset \mathbb{R}^d$ which guarantee that for *any* choice w of weights, the resulting mapping F (1) of a toric patch of shape $(\mathcal{A}, w)$ is a homeomorphism onto its image. This result has several consequences concerning the injectivity of toric patches.

Let us first give the Craciun-Feinberg Theorem. Let $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_m\} \subset \mathbb{R}^n$ be a finite set of points which affinely spans $\mathbb{R}^n$. For $k \in \mathbb{R}_>^m$ and $Z := \{\mathbf{z}_1, \ldots, \mathbf{z}_m\} \subset \mathbb{R}^n$, consider the map $G_k \colon \mathbb{R}_>^n \to \mathbb{R}^n$ defined by

$$G_k(x) \;:=\; \sum_{i=1}^{m} k_i \, x^{\mathbf{y}_i} \mathbf{z}_i \;\; . \tag{8}$$

Theorem 7 (Craciun-Feinberg). *The map G_k is injective for every $k \in \mathbb{R}_>^m$ if and only if the determinant of the Jacobian matrix,*

$$\mathrm{Jac}(G_k) \;=\; \left(\frac{\partial (G_k)_i}{\partial x_j} \right)_{i,j=1}^{n} \;\; ,$$

does not vanish for any $x \in \mathbb{R}_>^n$ and any $k \in \mathbb{R}_>^m$.

We give a proof in Appendix A.

The condition of Theorem 7 that the Jacobian $\mathrm{Jac}(G_k)$ does not vanish for any $x \in \mathbb{R}_>^n$ is reminiscent of the Jacobian conjecture [15], which is that the Jacobian of a polynomial map $G \colon \mathbb{C}^n \to \mathbb{C}^n$ does not vanish if and only if the map G is an isomorphism. Since we are restricted to $x \in \mathbb{R}_>^n$, it is closer to the real Jacobian conjecture, which is however false [16], and therefore not necessarily relevant.

The condition of Theorem 7 is conveniently restated in terms of Y and Z. For a list $I = \{i_1, \ldots, i_n\} \subset \{1, \ldots, m\}$, which we write as $I \in \binom{[m]}{n}$, let Y_I be the determinant of the matrix whose columns are the vectors $\mathbf{y}_{i_1}, \ldots, \mathbf{y}_{i_n}$, and define Z_I similarly. In Appendix A, we deduce the following corollary.

Corollary 8. *The map G_k (8) is injective for all $k \in \mathbb{R}^m_>$ if and only if $(Y_I \cdot Z_I) \cdot (Y_J \cdot Z_J) \geq 0$ for every $I, J \in \binom{[m]}{n}$ and at least one product $Y_I \cdot Z_I$ is non-zero.*

This leads to a generalization of Birch's Theorem. An ordered list $p_0, \ldots, p_d$ of affinely independent points in $\mathbb{R}^d$ determines an orientation of $\mathbb{R}^d$—simply consider the basis

$$p_1 - p_0\,,\ p_2 - p_0\,,\ \ldots\,,\ p_d - p_0\ .$$

Let $\mathcal{A}$ and $\mathcal{B} = \{\mathbf{b}_\mathbf{a} \mid \mathbf{a} \in \mathcal{A}\}$ be finite sets of points in $\mathbb{R}^d$. Suppose that $\{\mathbf{a}_0, \ldots, \mathbf{a}_d\}$ is an affinely independent subset of $\mathcal{A}$. If the corresponding subset $\{\mathbf{b}_{\mathbf{a}_0}, \ldots, \mathbf{b}_{\mathbf{a}_d}\}$ of $\mathcal{B}$ is also affinely independent, then each subset determines an orientation, and the two orientations are either the same or they are opposite. We say that $\mathcal{A}$ and $\mathcal{B}$ are *compatible* if either every such pair of orientations is the same, or if every such pair of orientations is opposite. We further need that there is at least one affinely independent subset of $\mathcal{A}$ such that the corresponding subset of $\mathcal{B}$ is also affinely independent. Observe that compatibility is preserved by invertible affine transformations acting separately on $\mathcal{A}$ and $\mathcal{B}$.

In Fig. 3 shows three sets of labeled points. The first and second sets are compatible, but neither is compatible with the third.

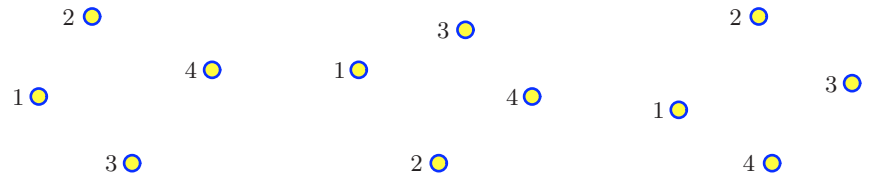

Fig. 3. Compatible and incompatible sets of points.

We give our generalization of Birch's Theorem. Suppose that $\Delta \subset \mathbb{R}^d$ is the convex hull of $\mathcal{A}$ and $\{\beta_\mathbf{a} \colon \Delta \to \mathbb{R}_\geq \mid \mathbf{a} \in \mathcal{A}\}$ are toric Bézier functions for $\mathcal{A}$. For any $w \in \mathbb{R}^\mathcal{A}_>$, let $F_w \colon \Delta \to \mathbb{R}^d$ be the toric patch of shape $(\mathcal{A}, w)$ given by the control points $\mathcal{B} \subset \mathbb{R}^d$:

$$F_w(x) \ := \ \frac{\sum_{\mathbf{a}\in\mathcal{A}} w_\mathbf{a}\beta_\mathbf{a}(x)\,\mathbf{b}_\mathbf{a}}{\sum_{\mathbf{a}\in\mathcal{A}} w_\mathbf{a}\beta_\mathbf{a}(x)}\ . \tag{9}$$

Theorem 9. *The map F_w is injective for all $w \in \mathbb{R}^\mathcal{A}_>$ if and only if $\mathcal{A}$ and $\mathcal{B}$ are compatible.*

As any set $\mathcal{A}$ is compatible with itself, this implies Birch's Theorem (Theorem 4).

Example 10. Let $\triangle$ be the convex hull of $\{(0,0),(1,0),(0,1)\}$. Set $\mathcal{A} := 3\triangle \cap \mathbb{Z}^2$ and let $w \in \mathbb{R}^\mathcal{A}_>$ be the weights of a cubic Bézier patch (Example 2 with $m = 3$).

We consider choices $\mathcal{B} \subset \mathbb{R}^2$ of control points that are compatible with $\mathcal{A}$. For convenience, we will require that $\mathbf{b_a} = \mathbf{a}$ when $\mathbf{a}$ is a vertex and that if $\mathbf{a}$ lies on an edge of 3Δ, then so does $\mathbf{b_a}$. For these edge control points, compatibility imposes the restriction that they appear along the edge in the same order as the corresponding exponents from $\mathcal{A}$. The placement of the center control point is however constrained. We show two compatible choices of $\mathcal{B}$ in Fig. 4. On the left is the situation of Birch's Theorem, in which $\mathbf{b_a} = \mathbf{a}$, and on the right we have moved the edge control points. The region in which we are free to move the center point is shaded in each picture.

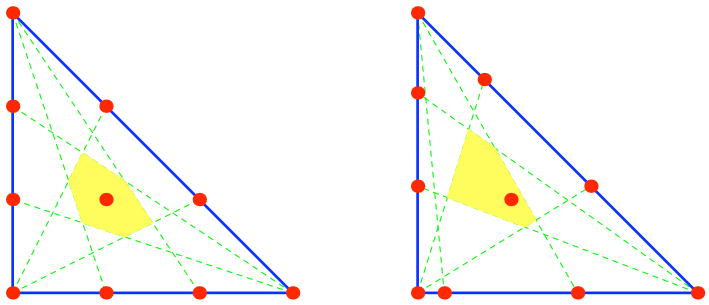

Fig. 4. Compatible control points for the Bézier triangle.

Proof (Theorem 9). Let (t, x) be coordinates for $\mathbb{R}^{d+1}$ and consider the map $G_w \colon \mathbb{R}^{d+1}_{>} \to \mathbb{R}^{d+1}$ defined by

$$G_w(t,x) \;=\; \sum_{\mathbf{a}\in\mathcal{A}} t x^{\mathbf{a}} w_{\mathbf{a}} (1, \mathbf{b_a}) \;\; .$$

We claim that F_w is injective if and only if G_w is injective.

Since F_w is the composition (6) $\Delta \xrightarrow{\beta} X_{\mathcal{A}} \xrightarrow{w.} w.X_{\mathcal{A}} \xrightarrow{\pi_{\mathcal{B}}} \mathbb{R}^d$, with the first map an isomorphism, F_w is injective if and only if the composition of the last two maps is injective. Since $X_{\mathcal{A}}$ is compact, this will be injective if and only if its restriction to the interior $X^\circ_{\mathcal{A}}$ of $X_{\mathcal{A}}$ is injective. Precomposing with the monomial parametrization (7) of $X^\circ_{\mathcal{A}}$, we see that F_w is injective if and only if the map $H_w \colon \mathbb{R}^d_{>} \to \mathbb{R}^d$ defined by

$$H_w(x) \;=\; \frac{\sum_{\mathbf{a}\in\mathcal{A}} x^{\mathbf{a}} w_{\mathbf{a}} \mathbf{b_a}}{\sum_{\mathbf{a}\in\mathcal{A}} x^{\mathbf{a}} w_{\mathbf{a}}} \;:\; \mathbb{R}^d_{>} \xrightarrow{\varphi_{\mathcal{A}}} X_{\mathcal{A}} \xrightarrow{w.} w.X_{\mathcal{A}} \xrightarrow{\pi_{\mathcal{B}}} \mathbb{R}^d$$

is injective.

Since $G_w(t,x) = t \cdot G_w(1,x)$, these values lie on a ray through the origin, and we invite the reader to check that this ray meets the hyperplane with first coordinate 1 at the point $(1, H_w(x))$. Thus G_w is injective if and only if H_w is injective, which is equivalent to the map F_w being injective.

We deduce the theorem by showing that G_w is injective. This follows from Corollary 8 as the condition that $\mathcal{A}$ and $\mathcal{B}$ are compatible is equivalent to $Y_I \cdot Z_I \geq 0$ for all $I \in \binom{[m]}{d+1}$, where $Y = \{(1, \mathbf{a}) \mid \mathbf{a} \in \mathcal{A}\}$, $Z = \{(1, \mathbf{b_a}) \mid \mathbf{a} \in \mathcal{A}\}$ and we have $m = \#\mathcal{A}$.

We now describe two applications of Theorem 9 to modeling.

3.1 Solid Modeling with (Toric) Bézier Patches

In solid modeling, we represent a 3-dimensional solid by covering it with 3-dimensional patches, for example using Bézier toric patches as finite elements. Besides the obvious C^0 or higher continuity along the boundary as required, such Bézier finite elements should at least provide a one-to-one parametrization of their image, i.e. they should be injective. By Theorem 9, we may guarantee injectivity by requiring that the control points $\mathcal{B}$ be compatible with the exponents $\mathcal{A}$. Moreover, if these sets are incompatible, then there is some choice of weights for which the patch is not injective.

3.2 Injectivity of Bézier Curves and Surfaces

Typically, the exponents $\mathcal{A}$ and the control points do not lie in the same space; surfaces ($\mathcal{A} \subset \mathbb{R}^2$) are modeled in 3-space ($\mathcal{B} \subset \mathbb{R}^3$), or curves ($\mathcal{A} \subset \mathbb{R}$) in 2– or 3–space ($\mathcal{B} \subset \mathbb{R}^2$ or $\mathbb{R}^3$). Nevertheless, Theorem 9 gives conditions that imply injectivity of patches.

Let $p \in \mathbb{R}^{n+1}$ be a point disjoint from a hyperplane, H. The projection $\mathbb{R}^{n+1} \dashrightarrow H$ with center p is the map which associates a point $x \in \mathbb{R}^{n+1}$ to the intersection of the line $\overline{px}$ with H. We use a broken arrow as the projection is not defined on the plane through p parallel to H. Identifying H with $\mathbb{R}^n$ gives a projection map $\mathbb{R}^{n+1} \dashrightarrow \mathbb{R}^n$. A coordinate projection $(x_1, \ldots, x_n, x_{n+1}) \mapsto (x_1, \ldots, x_n)$ is a projection with center at infinity. More generally, a *projection* $\mathbb{R}^n \dashrightarrow \mathbb{R}^d$ is a sequence of such projections from points.

Theorem 11. *Let $\mathcal{A} \subset \mathbb{R}^d$, $w \in \mathbb{R}^{\mathcal{A}}_{>}$, and $\mathcal{B} \subset \mathbb{R}^n$ be the exponents, weights, and control points of a toric patch and let Δ be the convex hull of $\mathcal{A}$. If there is a projection $\pi \colon \mathbb{R}^n \dashrightarrow \mathbb{R}^d$ such that $\mathcal{A}$ is compatible with the image $\pi(\mathcal{B})$ of $\mathcal{B}$, then the mapping $F \colon \Delta \to \mathbb{R}^n$ given by the toric blending functions associated to $\mathcal{A}$, the weights w, and control points $\mathcal{B}$ is injective.*

Proof. By Theorem 9, the composition $\pi \circ F$ is injective, from which it follows that F must have been injective.

Example 12. For the curve on the left in Fig. 1 (which is reproduced below), the vertical projection $\mathbb{R}^2 \to \mathbb{R}^1$ maps the control points $\{\mathbf{b}_0, \ldots, \mathbf{b}_5\}$ to points on the line in the same order as the exponents $\mathcal{A} = \{0, 1, 2, 3, 4, 5\}$, which implies that the curve has no self-intersections.

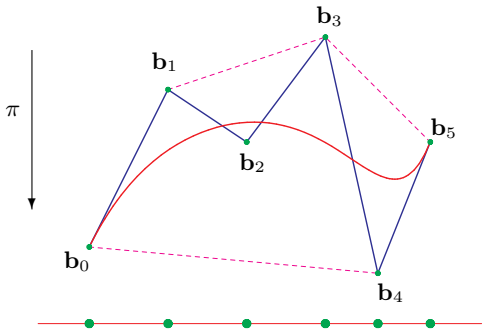

Fig. 5. A compatible projection.

4 Control Polytopes and Toric Degenerations

The convex hull property asserts that the image, $F(\Delta)$, of a toric Bézier patch of shape $(\mathcal{A}, w)$ given by control points $\mathcal{B} = \{\mathbf{b}_\mathbf{a} \mid \mathbf{a} \in \mathcal{A}\} \subset \mathbb{R}^n$ and weights $w \in \mathbb{R}^{\mathcal{A}}_{>}$ lies in the convex hull of the control points. When $F(\Delta)$ is a curve, the control points may be joined sequentially to form the *control polygon*, which is a piecewise linear representation of the curve. When $F(\Delta)$ is however a surface patch, there are many ways to interpolate the control points by triangles or other polygons to obtain a piecewise linear surface, called a *control polytope*, that represents the patch. The shape of this control polytope affects the shape of the patch. For example, when the control points have the form $(\mathbf{a}, \lambda(\mathbf{a}))$ for λ a convex function, then the patch is convex [17, 1]. Also, Leroy [18] uses a particular control polytope for the graph of a function to obtain certificates of positivity for polynomials.

Among all control polytopes for a given set of control points, we identify the class of regular control polytopes, which come from regular triangulations of the exponents $\mathcal{A}$. These regular control polytopes are related to the shape of the patch in the following precise manner: There is a choice of weights so that a toric Bézier patch is arbitrarily close to a given control polytope if and only if that polytope is regular.

4.1 Bézier Curves

It is instructive to begin with Bézier curves. A Bézier curve of degree m in $\mathbb{R}^n$ with weights w is the composition (6),

$$[0,1] \xrightarrow{\beta} \Delta^{m+1} \xrightarrow{w.} \Delta^{m+1} \xrightarrow{\pi_\mathcal{B}} \mathbb{R}^n \ ,$$

where $\beta = (\beta_0, \ldots, \beta_m)$ with $\beta_i(x) = \binom{m}{i} x^i (1-x)^{m-i}$ for $x \in [0,1]$. Then the map β is given by $z_i = \binom{m}{i} x^i (1-x)^{m-i}$, for $i = 0, \ldots, m$. Here, $(z_0, \ldots, z_m) \in \mathbb{R}^{d+1}_{\geq}$ with $z_0 + \cdots + z_m = 1$ are the coordinates for Δ^{m+1}. The image $\beta[0,1] \subset$

$\triangle^{m+1}$ is defined by the binomials

$$\tbinom{m}{i}\tbinom{m}{j} z_a z_b \;-\; \tbinom{m}{a}\tbinom{m}{b} z_i z_j \;=\; 0\,, \qquad \text{for} \quad a+b = i+j\ . \tag{10}$$

To see this, suppose that $(z_0, \ldots, z_m) \in \mathbb{R}^{m+1}_{\geq}$ satisfies (10). Setting $x := z_1/(mz_0 + z_1)$, then we may solve these equations to obtain $z_i = \binom{m}{i} x^i (1-x)^{m-i} = \beta_i(x)$.

If $w = (w_0, \ldots, w_m) \in \mathbb{R}^{m+1}_{>}$ are weights, then $w.\beta[0,1]$ is defined in $\triangle^{m+1}$ by

$$w_i w_j \tbinom{m}{i}\tbinom{m}{j} z_a z_b \;-\; w_a w_b \tbinom{m}{a}\tbinom{m}{b} z_i z_j \;=\; 0\,, \qquad \text{for} \quad a+b = i+j\ . \tag{11}$$

Suppose that we choose weights $w_i := t^{i(m-i)}$. Dividing by $t^{i(m-i)+j(m-j)}$, (11) becomes

$$\tbinom{m}{i}\tbinom{m}{j} z_a z_b \;-\; t^{i^2+j^2-a^2-b^2} \tbinom{m}{a}\tbinom{m}{b} z_i z_j \;=\; 0\,, \qquad \text{for} \quad a+b = i+j\ . \tag{12}$$

Since $i + j = a + b$, we may assume that $a < i \leq j < b$. Setting $c := i - a = b - j \geq 1$, we see that

$$i^2 + j^2 - a^2 - b^2 \;=\; c(i+a) - c(j+b) \;=\; c(i-j+a-b) \leq -2c \;<\; 0\ .$$

If we consider the limit of these binomials (12) as $t \to \infty$, we obtain

$$z_a \cdot z_b \;=\; 0 \qquad \text{if} \quad |a-b| > 1\ .$$

These define the polygonal path in $\triangle^{m+1}$ whose ith segment is the edge

$$(\underbrace{0, \ldots, 0}_{i-1}, x\,, 1-x\,, \underbrace{0, \ldots, 0}_{m-i}) \qquad \text{for} \quad x \in [0,1]\ ,$$

and whose projection to $\mathbb{R}^n$ is the *control polygon* of the Bézier curve, which is the collection of line segments $\overline{\mathbf{b}_0, \mathbf{b}_1}$, $\overline{\mathbf{b}_1, \mathbf{b}_2}$, $\ldots$, $\overline{\mathbf{b}_{m-1}, \mathbf{b}_m}$.

We illustrate this when $m = 3$ in Fig. 6, which shows three different Bézier curves having the same control points, but different weights $w_i = t^{i(3-i)}$ for $t = 1, 3, 9$. In algebraic geometry, altering the weights in this manner is called a *toric degeneration*. The Bézier cubics are displayed together with the cubics $w.\beta[0,1]$ lying in the 3-simplex, $\triangle^4$, which is drawn in $\mathbb{R}^3$. In these pictures, the projection $\pi_{\mathcal{B}}$ is simply the vertical projection forgetting the third coordinate. The progression indicated in Fig. 6, where the Bézier curve approaches the control polygon as the parameter t increases, is a general phenomenon. Let $\|\cdot\|$ be the usual Euclidean distance in $\mathbb{R}^n$.

Theorem 13. *Suppose that $F_t\colon [0,1] \to \mathbb{R}^n$ is a Bézier curve of degree m with control points $\mathcal{B}$ and weights $w_i = t^{i(m-i)}$. Set $\kappa := \max\{\|\mathbf{b_a}\| : \mathbf{b_a} \in \mathcal{B}\}$. For any $\epsilon > 0$, if we have $t > \kappa m/\epsilon$, then the distance between the control polygon and any point of the Bézier curve $F_t[0,1]$ is less than ϵ.*

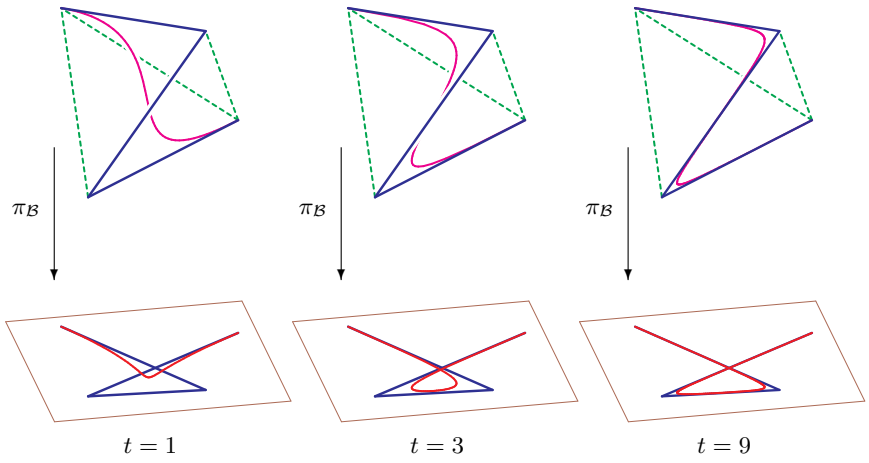

Fig. 6. Toric degenerations of a Bézier cubic.

Proof. Let $z \in w.\beta[0,1] \subset \Delta^{m+1}$, where the weights are $w_i = t^{i(m-i)}$ with $t > \kappa m/\epsilon$. Suppose that $b-a > 1$ are integers in $[0,m]$. Then there exist integers $i \le j$ with $a < i \le j < b$ and $a+b = i+j$. By (12), we have

$$z_a z_b \;=\; t^{i^2+j^2-a^2-b^2} \frac{\binom{m}{a}\binom{m}{b}}{\binom{m}{i}\binom{m}{j}} z_i z_j \;\; .$$

Since the binomial coefficients are log-concave, we have

$$\binom{m}{a}\binom{m}{b} \;<\; \binom{m}{i}\binom{m}{j} \;\; .$$

Using $i^2+j^2-a^2-b^2 < -2$ and $z_i + z_j \le 1$, we see that

$$z_a z_b \;<\; \frac{1}{4t^2} \;\; .$$

In particular, if $|b-a| > 1$, then at most one of z_a or z_b exceeds $1/2t$. We conclude that at most two, necessarily consecutive, coordinates of z may exceed $1/2t$. Suppose that i is an index such that $z_j < 1/2t$ if $j \ne i-1, i$ and let $x := z_{i-1}\mathbf{b}_{i-1} + (1-z_{i-1})\mathbf{b}_i$, a point along the ith segment of the control polygon. Since

$$1 \;\ge\; z_{i-1} + z_i \;=\; 1 - \sum_{j \ne i-1,i} z_j \;>\; 1 - \frac{m-1}{2t} \;\; ,$$

we have $0 \leq 1 - z_{i-1} - z_i < \frac{m-1}{2t} < \frac{m}{2t}$, and we see that

$$\begin{aligned}\|\pi_{\mathcal{B}}(z) - x\| &= \Big\|\sum z_j \mathbf{b}_j \;-\; (z_{i-1}\mathbf{b}_{i-1} + (1 - z_{i-1})\mathbf{b}_i)\Big\| \\ &\leq \sum_{j \neq i-1,i} z_j \|\mathbf{b}_j\| \;+\; |z_i - (1 - z_{i-1})| \|\mathbf{b}_i\| \\ &< \kappa\,\frac{m-1}{2t} \;+\; \kappa\,\frac{m}{2t} \;<\; \frac{\kappa m}{t} \;=\; \epsilon\,.\end{aligned}$$

This proves the theorem as $z \in w.\beta[0,1]$ is an arbitrary point of the curve $F[0,1]$.

4.2 Regular Triangulations and Control Polytopes

In dimensions 2 and higher, the analog of Theorem 13 requires the notion of a regular triangulation from geometric combinatorics. Let $\mathcal{A} \subset \mathbb{R}^d$ be a finite set of points and consider a lifting function $\lambda\colon \mathcal{A} \to \mathbb{R}$. Let P_λ be the convex hull of the lifted points

$$\lambda(\mathcal{A}) \;:=\; \{(\mathbf{a},\, \lambda(\mathbf{a})) \mid \mathbf{a} \in \mathcal{A}\} \;\subset\; \mathbb{R}^{d+1}\;.$$

We assume that P_λ is full-dimensional in that $\mathbb{R}^{d+1}$ is its affine span.

The *upper facets* of P_λ are those facets whose outward pointing normal vector has positive last coordinate. Any face of an upper facet is an *upper face*. We illustrate this below when $d = 1$, where the displayed arrows are outward pointing normal vectors to upper facets.

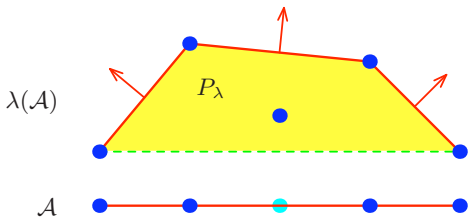

Fig. 7. Upper faces and a regular subdivision.

Projecting these upper facets to $\mathbb{R}^d$ yields a *regular polyhedral subdivision* of the convex hull of $\mathcal{A}$, which is the image of P_λ. For our purposes, we will need to assume that the lifting function is generic in that all upper facets of P_λ are simplices. In this case, we obtain a *regular triangulation* of $\mathcal{A}$. This consists of a collection

$$\{\mathcal{A}_i \mid i = 1, \ldots, m\}$$

of subsets of $\mathcal{A}$, where each subset $\mathcal{A}_i$ consists of $d{+}1$ elements and spans a d-dimensional simplex. We regard all subsets of the facets $\mathcal{A}_i$ as faces of the

triangulation. These simplices form a subdivision in that they cover the convex hull of $\mathcal{A}$ and any two with a non-empty intersection meet along a common face.

The subdivision induced by the lifting function in Fig. (7) consists of three intervals resulting from removal of the middle point of $\mathcal{A}$, which does not participate in the subdivision as it is not lifted high enough.

A set may have many regular triangulations, and not every point needs to participate in a given triangulation. Figure 8 shows the edges in four regular triangulations of $3\triangle \cap \mathbb{Z}^2$.

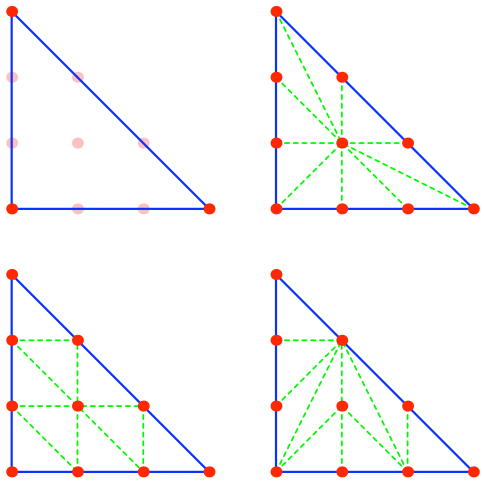

Fig. 8. Some regular triangulations.

Not every triangulation is regular. We may assume that a lifting function λ for the triangulation of $4\triangle \cap \mathbb{Z}^2$ in Fig. 9 takes a constant value at the three interior points. The clockwise neighbor of any vertex of the big triangle must be lifted lower than that vertex. (Consider the figure they form with the parallel edge of the interior triangle.) Since the edge of the big triangle is lifted to a convex path, this is impossible, except in some M.C. Escher woodcuts.

Definition 14. *Let $\mathcal{B} = \{\mathbf{b_a} \mid \mathbf{a} \in \mathcal{A}\} \subset \mathbb{R}^n$ be a collection of control points indexed by a finite set of exponents $\mathcal{A} \subset \mathbb{R}^d$ with $d \leq n$. Given a regular triangulation $\mathcal{T} = \{\mathcal{A}_i \mid i = 1, \ldots, m\}$ of $\mathcal{A}$ we define the control polytope as follows. For each d-simplex $\mathcal{A}_i$ in $\mathcal{T}$, the corresponding points of $\mathcal{B}$ span a (possibly degenerate) simplex*

$$conv\{\mathbf{b_a} \mid \mathbf{a} \in \mathcal{A}_i\} \ .$$

The union of these simplices in $\mathbb{R}^n$ forms the regular control polytope $\mathcal{B}(\mathcal{T})$ that is induced by the regular triangulation $\mathcal{T}$ of $\mathcal{A}$. This is a simplicial complex in $\mathbb{R}^n$ with vertices in $\mathcal{B}$ that has the same combinatorial type as the triangulation $\mathcal{T}$ of $\mathcal{A}$.

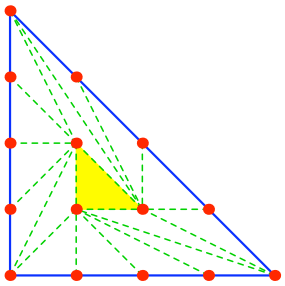

Fig. 9. An irregular triangulation.

If the coordinate points $(\mathbf{e_a} \mid \mathbf{a} \in \mathcal{A})$ of $\mathbb{R}^{\mathcal{A}}$ are our control points (these are the vertices of $\triangle^{\mathcal{A}}$), then the regular control polytope is just the geometric realization $|\mathcal{T}|$ of the simplicial complex $\mathcal{T}$, which is a subcomplex of the simplex $\triangle^{\mathcal{A}}$. In general, $\mathcal{B}(\mathcal{T})$ is the image of this geometric realization $|\mathcal{T}| \subset \triangle^{\mathcal{A}}$ under the projection $\pi_{\mathcal{B}}$.

Example 15. Let $\mathcal{A} := 3\triangle \cap \mathbb{Z}^2$, the exponents for a cubic Bézier triangle. Figure 10 shows the three control polytopes corresponding to the last three regular triangulations of (8), all with the same control points.

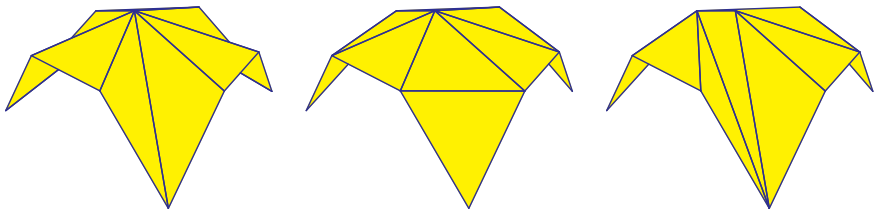

Fig. 10. Three control polytopes.

The reason that we introduce regular control polytopes is that they may be approximated by toric Bézier patches.

Theorem 16. *Let $\mathcal{A} \subset \mathbb{R}^d$, $w \in \mathbb{R}^{\mathcal{A}}_{>}$, and $\mathcal{B} \subset \mathbb{R}^n$ be exponents, weights, and control points for a toric Bézier patch. Suppose that $\mathcal{T}$ is a regular triangulation of $\mathcal{A}$ induced by a lifting function $\lambda\colon \mathcal{A} \to \mathbb{R}$. For each $t > 0$, let $F_t\colon \Delta \to \mathbb{R}^n$ be the toric Bézier patch of shape $\mathcal{A}$ with control points $\mathcal{B}$ and weights $t^{\lambda(\mathbf{a})}w_{\mathbf{a}}$. Then, for any $\epsilon > 0$ there exists a t_0 such that if $t > t_0$, the image $F_t(\Delta)$ lies within ϵ of the control polytope $\mathcal{B}(\mathcal{T})$.*

We prove Theorem 16 in Appendix B. Figure 11 illustrates Theorem 16 for a cubic Bézier triangle with the control points of Example 15. The patch on the left is the cubic Bézier triangle with the weights of Example 2. The second and third patches are its deformations corresponding to the lifting function inducing the leftmost control polytope of Fig. 10. The values of t are 1, 5, and 100, as we move from left to right.

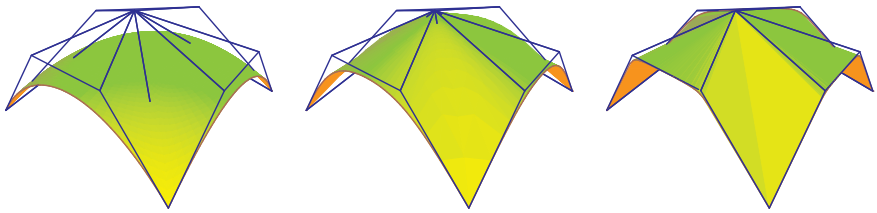

Fig. 11. Degeneration to the control polytope.

An absolutely unpractical consequence of Theorem 16 is a universality result: Any surface which admits a triangulation that forms a regular control polytope may be approximated by a single Bézier patch.

As with Theorem 13, the main idea behind the proof of Theorem 16 (which is given in Appendix B) is that for t large enough, the translated patch $t.X_{\mathcal{A},w} \subset \triangle^{\mathcal{A}}$ can be made arbitrarily close to the geometric realization $|\mathcal{T}| \subset \triangle^{\mathcal{A}}$ of the regular triangulation $\mathcal{T}$. The result follows by projecting this into $\mathbb{R}^n$ using $\pi_{\mathcal{B}}$.

In Appendix B we also prove a weak converse to Theorem 16. Namely if $w.X_{\mathcal{A}}$ is sufficiently close to the geometric realization $|\mathcal{T}|$ of a triangulation $\mathcal{T}$, then $\mathcal{T}$ is in fact the regular triangulation of $\mathcal{A}$ induced by the lifting function $\lambda(\mathbf{a}) = \log(w_{\mathbf{a}})$.

Theorem 17. *Let $\mathcal{A} \subset \mathbb{R}^d$ be a finite set of exponents. Suppose that $|\mathcal{T}| \subset \triangle^{\mathcal{A}}$ is the geometric realization of a triangulation $\mathcal{T}$ of $\mathcal{A}$ and there is a weight w such that the distance between $X_{\mathcal{A},w}$ and $|\mathcal{T}|$ is less than $1/2(d+1)$. Then $\mathcal{T}$ is the regular triangulation induced by the lifting function $\lambda(\mathbf{a}) = \log(w_{\mathbf{a}})$.*

A Proofs of Injectivity Theorems

Theorem 7. (Craciun-Feinberg) *The map G_k is injective for every $k \in \mathbb{R}^m_>$ if and only if the determinant of the Jacobian matrix,*

$$\mathrm{Jac}(G_k) \;=\; \left(\frac{\partial (G_k)_i}{\partial x_j}\right)^n_{i,j=1} \quad ,$$

does not vanish for any $x \in \mathbb{R}^n_>$ and any $k \in \mathbb{R}^m_>$.

Proof. We show the equivalence of the two statements, transforming one into the other. First, suppose there is a $k \in \mathbb{R}^m_>$ so that G_k is not injective. Then there exist $c, d \in \mathbb{R}^n_>$ so that $G_k(c) = G_k(d)$. Then we have

$$0 \;=\; \sum_{i=1}^{m} k_i(c^{\mathbf{y}_i} - d^{\mathbf{y}_i})\mathbf{z}_i \;=\; \sum_{i=1}^{m} k'_i((\tfrac{c}{d})^{\mathbf{y}_i} - 1)\mathbf{z}_i \;\;,$$

where $k' \in \mathbb{R}^m_>$ is defined by $k'_i = k_i d^{\mathbf{y}_i}$, and $\frac{c}{d} \in \mathbb{R}^n_>$ has ith coordinate $\frac{c_i}{d_i}$. In particular, $G_{k'}(\frac{c}{d}) = G_{k'}(\iota)$, where $\iota := (1, \ldots, 1)$. Define $\mathbf{v} \in \mathbb{R}^n$ by

$$v_i \;:=\; \log(c_i) - \log(d_i) \;\;,$$

so that $e^{\mathbf{y}_i \cdot \mathbf{v}} = (\frac{c}{d})^{\mathbf{y}_i}$, where $\mathbf{y}_i \cdot \mathbf{v}$ is the Euclidean dot product, and we now have

$$0 \;=\; \sum_{i=1}^{m} k'_i(e^{\mathbf{y}_i \cdot \mathbf{v}} - 1)\mathbf{z}_i \;\;. \tag{13}$$

Define the univariate function by $f(t) = (e^t - 1)/t$ for $t \neq 0$ and set $f(0) = e$. Then f is an increasing continuous bijection between $\mathbb{R}$ and $\mathbb{R}_>$. Define $k'' \in \mathbb{R}^m_>$ by $k''_i := k'_i f(\mathbf{y}_i \cdot \mathbf{v})$. Then $k'_i(e^{\mathbf{y}_i \cdot \mathbf{v}} - 1) = k''_i(\mathbf{y}_i \cdot \mathbf{v})$, and (13) becomes

$$0 \;=\; \sum_{i=1}^{m} k''_i(\mathbf{y}_i \cdot \mathbf{v})\mathbf{z}_i \;\;. \tag{14}$$

We claim that $\mathbf{v}$ lies in the kernel of the Jacobian matrix of $G_{k''}$ evaluated at the point ι. Indeed, let $\mathbf{e}_j := (0, \ldots, 1, \ldots, 0) \in \mathbb{R}^n$ be the unit vector in the jth direction. Then

$$\frac{\partial G_{k''}}{\partial x_j}(x) \;=\; \sum_{i=1}^{m} k''_i x^{\mathbf{y}_i - \mathbf{e}_j} y_{i,j}\mathbf{z}_i \;\;,$$

where $\mathbf{y}_i = (y_{i,1}, \ldots, y_{i,n})$. Since $\iota^{\mathbf{y}_i - \mathbf{e}_j} = 1$ and $\sum_j y_i v_j = \mathbf{y}_i \cdot \mathbf{v}$, we see that

$$\sum_{j=1}^{n} \frac{\partial G_{k''}}{\partial x_j}(\iota)\, v_j \;=\; \sum_{j=1}^{n}\sum_{i=1}^{m} k''_i \iota^{\mathbf{y}_i - \mathbf{e}_j} y_{i,j} v_j \mathbf{z}_i \;=\; \sum_{i=1}^{m} k''_i(\mathbf{y}_i \cdot \mathbf{v})\mathbf{z}_i \;=\; 0 \;\;,$$

so that $\mathbf{v}$ lies in the kernel of the Jacobian matrix of $G_{k''}$ evaluated at ι, which implies that the Jacobian determinant of $G_{k''}$ vanishes at ι.

The theorem follows as these arguments are reversible. 

Corollary 8. *The map G_k (8) is injective for all $k \in \mathbb{R}^m_>$ if and only if $(Y_I \cdot Z_I) \cdot (Y_J \cdot Z_J) \geq 0$ for every $I, J \in \binom{[m]}{n}$ and at least one product $Y_I \cdot Z_I$ is non-zero.*

Proof. Observe first that the Jacobian matrix $\mathrm{Jac}(G_k)$ factors as the product of matrices $\delta^{-1} Y D Z^T$, where δ is the diagonal matrix with entries $(x_1, \ldots, x_n)$, D is the diagonal matrix with entries $(k_1 x^{\mathbf{y}_1}, \ldots, k_m x^{\mathbf{y}_m})$ and Y and Z are the

matrices whose columns are the vectors $\mathbf{y}_i$ and $\mathbf{z}_i$, respectively. If we apply the Binet-Cauchy Theorem to this factorization, we see that

$$\det \operatorname{Jac}(G_k) \;=\; x^{-\iota} \cdot \sum_{I \in \binom{[m]}{n}} \prod_{i \in I} k_i x^{\mathbf{y}_i} \cdot Y_I \cdot Z_I \;, \tag{15}$$

where $\iota = (1, \ldots, 1)$.

Suppose that $(Y_I \cdot Z_I) \cdot (Y_J \cdot Z_J) \geq 0$ for every $I, J \in \binom{[m]}{n}$, and at least one product $Y_I \cdot Z_I$ is non-zero. Then all terms in the sum (15) have the same sign and not all are zero, and so the Jacobian does not vanish for any $x \in \mathbb{R}^n_>$ and $k \in \mathbb{R}^m_>$. Thus G_k is injective for all $k \in \mathbb{R}^m_>$, by Theorem 7.

Suppose that there are two subsets $I, J \in \binom{[m]}{n}$ such that $Y_I Z_I > 0$ and $Y_J Z_J < 0$. For $t \in \mathbb{R}_>$ and $K \in \binom{[m]}{n}$, define $k(K,t) \in \mathbb{R}^m_>$ by

$$k(K,t)_j \;:=\; \begin{cases} t & \text{if } j \in K \\ 1 & \text{otherwise} \end{cases}$$

If we fix $x \in \mathbb{R}^d_>$, then the expansion (15), implies that $\det \operatorname{Jac}(G_{k(K,t)})(x)$ has the same sign as $Y_K Z_K$ when $t \gg 0$, at least when $Y_K Z_K \neq 0$.

We conclude that there is some k, x such that $\det \operatorname{Jac}(G_k)(x) > 0$ and some k, x such that $\det \operatorname{Jac}(G_k)(x) < 0$, and therefore some k, x such that $\det \operatorname{Jac}(G_k)(x) = 0$. This implies the corollary. 

B Three Toric Theorems

Theorem 3. *Suppose that $\mathcal{A} \subset \mathbb{R}^d$ is a finite set of points with convex hull Δ. Let $\beta = \{\beta_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}\}$ be a collection of irrational toric Bézier functions for $\mathcal{A}$. Then $\beta(\Delta) = X_{\mathcal{A}}$, the closure of the image of $\varphi_{\mathcal{A}}$.*

Proof. Let Δ° be the interior of Δ, which we assume has ℓ facets and is given by the facet inequalities $0 \leq h_i(x), i = 1, \ldots, \ell$. Define two maps $H \colon \Delta^\circ \to \mathbb{R}^\ell_>$ and $\psi \colon \mathbb{R}^\ell_> \to \triangle^{\mathcal{A}}$ by

$$\begin{aligned} H \;&:\; x \longmapsto (h_1(x), \ldots, h_\ell(x)) \,, \\ \psi \;&:\; u \longmapsto [u_1^{h_1(\mathbf{a})} \cdots u_\ell^{h_\ell(\mathbf{a})} \;:\; \mathbf{a} \in \mathcal{A}] \,. \end{aligned}$$

Then the map $\beta \colon \Delta^\circ \to \triangle^{\mathcal{A}}$ (whose image is dense in $X_{\mathcal{A}}$) is the composition of the maps H and ψ.

$$\beta \;:\; \Delta^\circ \xrightarrow{\;H\;} \mathbb{R}^\ell \xrightarrow{\;\psi\;} \triangle^{\mathcal{A}} \;.$$

Let us recall the definiton of the map $\varphi_{\mathcal{A}} \colon \mathbb{R}^d_> \to \triangle^{\mathcal{A}}$,

$$\varphi_{\mathcal{A}} \;:\; (x_1, \ldots, x_d) \;\longmapsto\; [x^{\mathbf{a}} \;:\; \mathbf{a} \in \mathcal{A}] \;.$$

The theorem follows once we show that the map ψ factors through the map $\varphi_{\mathcal{A}}$. For this, define a new map $f_\Delta \colon \mathbb{R}^\ell_> \to \mathbb{R}^d_>$ by

$$f_\Delta \;:\; u \;\longmapsto\; t = (t_1, \ldots, t_d) \qquad \text{where} \quad t_j \;:=\; \prod_{i=1}^{\ell} u_i^{\mathbf{v}_i \cdot \mathbf{e}_j} \;\; .$$

Then we claim that

$$\psi \;:\; \mathbb{R}^\ell \;\xrightarrow{\;f_\Delta\;}\; \mathbb{R}^d_> \;\xrightarrow{\;\varphi_{\mathcal{A}}\;}\; \Delta^{\mathcal{A}} \;\; .$$

To see this, we compute the component of $\psi(u)$ corresponding to $\mathbf{a} \in \mathcal{A}$,

$$\prod_{i=1}^{\ell} u_i^{h_i(\mathbf{a})} \;=\; \prod_{i=1}^{\ell} u_i^{\mathbf{v}_i \cdot \mathbf{a} + c_i} \;=\; \prod_{i=1}^{\ell} u_i^{c_i} \;\cdot\; \prod_{i=1}^{\ell} u_i^{\mathbf{v}_i \cdot \mathbf{a}} \;=\; u^c \cdot t^{\mathbf{a}} \;\; .$$

Thus $\psi(u) = u^c \varphi_{\mathcal{A}}(f_\Delta(u))$, as maps to $\mathbb{R}^{\mathcal{A}}_>$. The common factor u^c does not affect the image in $\Delta^{\mathcal{A}}$, which shows that $\psi = \varphi_{\mathcal{A}} \circ f_\Delta$ and proves the theorem.

A consequence of this proof of Theorem 3 is the derivation of equations which define the points of $X_{\mathcal{A}}$. This derivation is similar to, but easier than, the development of toric ideals in [19, Ch. 4], as we have monomials with arbitrary real-number exponents.

Suppose that we have a linear relation among the points of $\mathcal{A}$,

$$\sum_{\mathbf{a} \in \mathcal{A}} \mu_{\mathbf{a}} \cdot \mathbf{a} \;=\; \sum_{\mathbf{a} \in \mathcal{A}} \nu_{\mathbf{a}} \cdot \mathbf{a} \;\; , \tag{16}$$

for some $\mu, \nu \in \mathbb{R}^{\mathcal{A}}$. Then the analytic binomial

$$\prod_{\mathbf{a} \in \mathcal{A}} z_{\mathbf{a}}^{\mu_{\mathbf{a}}} \;-\; \prod_{\mathbf{a} \in \mathcal{A}} z_{\mathbf{a}}^{\nu_{\mathbf{a}}} \;=:\; z^\mu \;-\; z^\nu \tag{17}$$

vanishes on $\varphi_{\mathcal{A}}(\mathbb{R}^d_>)$, considered as a point in $\mathbb{R}^{\mathcal{A}}_>$. This follows from the easy calculation

$$\varphi^*_{\mathcal{A}}(z^\mu) \;=\; \prod_{\mathbf{a} \in \mathcal{A}} (x^{\mathbf{a}})^{\mu_{\mathbf{a}}} \;=\; x^{\sum_{\mathbf{a}} \mu_{\mathbf{a}} \cdot \mathbf{a}} \;\; .$$

Even after clearing denominators, the common zero set of the binomials (17) is not exactly the image $\varphi_{\mathcal{A}}(\mathbb{R}^d_>)$ in the simplex $\Delta^{\mathcal{A}}$, as the point $\varphi_{\mathcal{A}}(x) \in \Delta^{\mathcal{A}}$ is where the ray

$$\mathbb{R}_> \cdot (x^{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}) \;\subset\; \mathbb{R}^{\mathcal{A}}_>$$

meets the simplex $\Delta^{\mathcal{A}}$. If we require that the binomial (17) is homogeneous in that $\sum_{\mathbf{a} \in \mathcal{A}} \mu_{\mathbf{a}} = \sum_{\mathbf{a} \in \mathcal{A}} \nu_{\mathbf{a}}$, then it vanishes at every point of this ray and therefore on the image of $\varphi_{\mathcal{A}}$ in $\Delta^{\mathcal{A}}$. Since the coordinates are positive numbers, we may further assume that (16) is an affine relation in that

$$\sum_{\mathbf{a} \in \mathcal{A}} \mu_{\mathbf{a}} \;=\; \sum_{\mathbf{a} \in \mathcal{A}} \nu_{\mathbf{a}} \;=\; 1 \;\; .$$

These necessary conditions are also sufficient.

Proposition 18. *A point z in $\triangle^{\mathcal{A}}$ lies in $X_{\mathcal{A}}$ if and only if we have*

$$z^{\mu} \;-\; z^{\nu} \;=\; 0$$

for all $\mu, \nu \in \mathbb{R}^{\mathcal{A}}$ with $\sum_{\mathbf{a}\in\mathcal{A}} \mu_{\mathbf{a}} = \sum_{\mathbf{a}\in\mathcal{A}} \nu_{\mathbf{a}} = 1$ and

$$\sum_{\mathbf{a}\in\mathcal{A}} \mu_{\mathbf{a}} \cdot \mathbf{a} \;=\; \sum_{\mathbf{a}\in\mathcal{A}} \nu_{\mathbf{a}} \cdot \mathbf{a} \;.$$

One way to see the sufficiency is to pick an affinely independent subset $\mathcal{C}$ of $\mathcal{A}$ that affinely spans $\mathbb{R}^d$ and use the formula $x^{\mathbf{a}} = z_{\mathbf{a}}$ for $\mathbf{a} \in \mathcal{C}$ to solve for $x \in \mathbb{R}^d$. Then the point $z \in \triangle^{\mathcal{A}}$ satisfies this collection of binomials if and only if $z = \varphi_{\mathcal{A}}(x)$. This is also evident if we take logarithms of the coordinates.

These arguments only work for points z in the interior of $\triangle^{\mathcal{A}}$. For points of $X_{\mathcal{A}}$ with some coordinates zero, we use the recursive nature of polytopes and the toric Bézier functions. Namely, if we restrict the collection $\{\beta_{\mathbf{a}} \mid \mathbf{a} \in \mathcal{A}\}$ of toric Bézier functions to a face $\mathcal{F}$ of the convex hull of $\mathcal{A}$, then those whose index $\mathbf{a}$ does not lie in $\mathcal{F}$ vanish, while those indexed by points of $\mathcal{A}$ lying in $\mathcal{F}$ specialize to toric Bézier functions for $\mathcal{F}$.

Theorem 16. *Let $\mathcal{A} \subset \mathbb{R}^d$, $w \in \mathbb{R}^{\mathcal{A}}_{>}$, and $\mathcal{B} \subset \mathbb{R}^n$ be exponents, weights, and control points for a toric Bézier patch. Suppose that $\mathcal{T}$ is a regular triangulation of $\mathcal{A}$ induced by a lifting function $\lambda\colon \mathcal{A} \to \mathbb{R}$. For each $t > 0$, let $F_t\colon \Delta \to \mathbb{R}^n$ be the toric Bézier patch of shape $\mathcal{A}$ with control points $\mathcal{B}$ and weights $t^{\lambda(\mathbf{a})} w_{\mathbf{a}}$. Then, for any $\epsilon > 0$ there exists a t_0 such that if $t > t_0$, the image $F_t(\Delta)$ lies within ϵ of the control polytope $\mathcal{B}(\mathcal{T})$.*

Proof. The lifting function $\lambda\colon \mathcal{A} \to \mathbb{R}$ inducing the triangulation $\mathcal{T}$ also induces an action of $\mathbb{R}_{>}$ on $\triangle^{\mathcal{A}}$ where $t \in \mathbb{R}_{>}$ acts on a point $z \in \triangle^{\mathcal{A}}$ by scaling its coordinates, $(t.z)_{\mathbf{a}} = t^{\lambda(\mathbf{a})} z_{\mathbf{a}}$. Then $F_t(\Delta)$ is the image of $t.X_{\mathcal{A},w}$ under the projection $\pi_{\mathcal{B}}\colon \triangle^{\mathcal{A}} \to \mathbb{R}^n$. It suffices to show that $t.X_{\mathcal{A},w}$ can be made arbtrarily close to $|\mathcal{T}| \subset \triangle^{\mathcal{A}}$, if we choose t large enough.

We single out some equations from Proposition 18. Suppose that $\mathbf{a}, \mathbf{b} \in \mathcal{A}$ are points that do not lie in a common simplex of $\mathcal{T}$. That is, the segment $\overline{\mathbf{a}, \mathbf{b}}$ is not an edge in the triangulation $\mathcal{T}$, and therefore it meets the interior of some face $\mathcal{F}$ of $\mathcal{T}$ so that there is a point common to the interiors of $\mathcal{F}$ and of $\overline{\mathbf{a}, \mathbf{b}}$. (If $\mathbf{a} = \mathbf{b}$, so that $\mathbf{a}$ does not participate in the triangulation $\mathcal{T}$, then this point is just $\mathbf{a}$.) This gives the equality of convex combinations

$$\mu\mathbf{a} + (1-\mu)\mathbf{b} \;=\; \sum_{\mathbf{c}\in\mathcal{F}} \nu_{\mathbf{c}} \cdot \mathbf{c} \;, \tag{18}$$

where all coefficients are positive and $\sum_{\mathbf{c}} \nu_{\mathbf{c}} = 1$. Thus

$$z_{\mathbf{a}}^{\mu} z_{\mathbf{b}}^{1-\mu} \;=\; \prod_{\mathbf{c}\in\mathcal{F}} z_{\mathbf{c}}^{\nu_{\mathbf{c}}}$$

holds on $X_{\mathcal{A}}$. The corresponding equation on $t.X_{\mathcal{A},w}$ is

$$z_{\mathbf{a}}^{\mu} z_{\mathbf{b}}^{1-\mu} \;=\; t^{\mu\mathbf{a}+(1-\mu)\mathbf{b}-\sum_{\mathbf{c}\in\mathcal{F}}\nu_{\mathbf{c}}\cdot\mathbf{c}} \cdot \frac{w_{\mathbf{a}}^{\mu} w_{\mathbf{b}}^{1-\mu}}{\prod_{\mathbf{c}\in\mathcal{F}} w_{\mathbf{c}}^{\nu_{\mathbf{c}}}} \cdot \prod_{\mathbf{c}\in\mathcal{F}} z_{\mathbf{c}}^{\nu_{\mathbf{c}}} \;\; . \tag{19}$$

Since $\overline{\mathbf{a},\mathbf{b}}$ is not in the triangulation, points in the interior of the lifted segment

$$\overline{(\mathbf{a},\lambda(\mathbf{a})),\,(\mathbf{b},\lambda(\mathbf{b}))}$$

lie below points of upper faces of the polytope P_λ. We apply this observation to the point (18). Its height in the lifted segment is $\mu\mathbf{a}+(1-\mu)\mathbf{b}$, while its height in the lift of the face $\mathcal{F}$ is $\sum_{\mathbf{c}\in\mathcal{F}}\nu_{\mathbf{c}}\cdot\mathbf{c}$, and so

$$\mu\mathbf{a}+(1-\mu)\mathbf{b} \;<\; \sum_{\mathbf{c}\in\mathcal{F}} \nu_{\mathbf{c}}\cdot\mathbf{c} \;\; .$$

This implies that the exponent of t in (19) is negative. Since the other terms on the right hand side are bounded, we see that the left hand side, and in fact the simple product $z_{\mathbf{a}}z_{\mathbf{b}}$, may be made as small as we please by requiring that t be sufficiently large.

Suppose that $\mathcal{A}$ consists of $\ell+d+1$ elements. Repeating the previous argument for the (finitely many) pairs of points $\mathbf{a},\mathbf{b}$ which are not both in any simplex of $\mathcal{T}$, we see that for any $\epsilon>0$, there is a t_0 such that if $t>t_0$ and $z\in X_{\mathcal{A},w}$, then

$$z_{\mathbf{a}}z_{\mathbf{b}} \;<\; \epsilon^2/4\ell^2 \;\; ,$$

whenever $\mathbf{a},\mathbf{b}$ do not lie in a common simplex of $\mathcal{T}$. In particular, at most one of $z_{\mathbf{a}}$ or $z_{\mathbf{b}}$ can exceed $\epsilon/2\ell$.

Let $z\in X_{\mathcal{A},w}$. Then there is some facet $\mathcal{F}$ of $\mathcal{T}$ such that if $\mathbf{a}\notin\mathcal{F}$, then $0\le z_{\mathbf{a}}<\epsilon/2\ell$. Suppose that $\mathcal{F}=\{\mathbf{a}_0,\mathbf{a}_1,\ldots,\mathbf{a}_d\}$ and set

$$z_{\mathcal{F}} \;:=\; (1-z_{\mathbf{a}_1}-\cdots-z_{\mathbf{a}_d})e_{\mathbf{a}_0}+z_{\mathbf{a}_1}e_{\mathbf{a}_1}+\cdots+z_{\mathbf{a}_d}e_{\mathbf{a}_d} \;\; ,$$

which is a point of the facet $|\mathcal{F}|$ of the geometric realization $|\mathcal{T}|\subset\Delta^{\mathcal{A}}$. Then

$$\begin{aligned}\|z-z_{\mathcal{F}}\| &= \left\| \sum_{\mathbf{a}\in\mathcal{A}} z_{\mathbf{a}}e_{\mathbf{a}} \;-\; (1-z_{\mathbf{a}_1}-\cdots-z_{\mathbf{a}_d})e_{\mathbf{a}_0} - z_{\mathbf{a}_1}e_{\mathbf{a}_1}-\cdots-z_{\mathbf{a}_d}e_{\mathbf{a}_d} \right\| \\ &\le \sum_{\mathbf{a}\notin\mathcal{F}} z_{\mathbf{a}} \;+\; (1-z_{\mathbf{a}_0}-\cdots-z_{\mathbf{a}_d}) \;=\; 2\sum_{\mathbf{a}\notin\mathcal{F}} z_{\mathbf{a}} \\ &< 2\ell\frac{\epsilon}{2\ell} \;=\; \epsilon\,,\end{aligned}$$

as $1=\sum_{\mathbf{a}} z_{\mathbf{a}}$.

We say that two subsets X and Y of Euclidean space are within a distance ϵ if for every point x of X there is some point y of Y within a distance ϵ of x, and vice-versa.

Theorem 17. *Let $\mathcal{A} \subset \mathbb{R}^d$ be a finite set of exponents. Suppose that $|\mathcal{T}| \subset \triangle^{\mathcal{A}}$ is the geometric realization of a triangulation $\mathcal{T}$ of $\mathcal{A}$ and there is a weight w such that the distance between $X_{\mathcal{A},w}$ and $|\mathcal{T}|$ is less than $1/2(d+1)$. Then $\mathcal{T}$ is the regular triangulation induced by the lifting function $\lambda(\mathbf{a}) = \log(w_{\mathbf{a}})$.*

Proof. To show that $\mathcal{T}$ is the regular triangulation induced by the lifting function λ whose value at $\mathbf{a}$ is $\log(w_{\mathbf{a}})$, we must show that if a segment $\overline{\mathbf{a},\mathbf{b}}$ between two points of $\mathcal{A}$ does not lie in the triangulation $\mathcal{T}$, then its lift by λ lies below the lift of some face $\mathcal{F}$ of $\mathcal{T}$.

Set $\epsilon := 1/2(d{+}1)$. For each face $\mathcal{F}$ of $\mathcal{T}$, let $x_{\mathcal{F}} \in \triangle^{\mathcal{A}}$ be the barycenter of $\mathcal{F}$,

$$x_{\mathcal{F}} \;:=\; \sum_{\mathbf{a}\in\mathcal{F}} \frac{1}{\#\mathcal{F}} e_{\mathbf{a}} \;\;,$$

where $\#\mathcal{F}$ is the number of points of $\mathcal{A}$ in $\mathcal{F}$, which is at most $d{+}1$. If z is a point of $X_{\mathcal{A},w}$ within a distance ϵ of $x_{\mathcal{F}}$, so that $\|z - x_{\mathcal{F}}\| < \epsilon$, then in particular no component of the vector $z - x_{\mathcal{F}}$ has absolute value exceeding ϵ. Thus we have the dichotomy

$$\begin{array}{ll} z_{\mathbf{a}} < \quad \epsilon \;=\; 1/2(d{+}1) & \text{if } \; \mathbf{a} \notin \mathcal{F}\,, \\ z_{\mathbf{a}} > 1/\#\mathcal{F} - \epsilon > 1/2(d{+}1) & \text{if } \; \mathbf{a} \in \mathcal{F}\;. \end{array} \tag{20}$$

Now suppose that the segment $\overline{\mathbf{a},\mathbf{b}}$ does not lie in the triangulation $\mathcal{T}$. Then there is a face $\mathcal{F}$ of the triangulation whose interior meets the interior of this segment. That is, there is an equality of convex combinations (18) and a corresponding equation that holds for points $z \in X_{\mathcal{A},w}$,

$$z_{\mathbf{a}}^{\mu} z_{\mathbf{b}}^{1-\mu} \cdot \prod_{\mathbf{c}\in\mathcal{F}} w_{\mathbf{c}}^{\nu_{\mathbf{c}}} \;=\; w_{\mathbf{a}}^{\mu} w_{\mathbf{b}}^{1-\mu} \cdot \prod_{\mathbf{c}\in\mathcal{F}} z_{\mathbf{c}}^{\nu_{\mathbf{c}}} \;\;.$$

Suppose that z is a point of $X_{\mathcal{A},w}$ that lies within a distance of $1/2(d{+}1)$ of the barycenter $X_{\mathcal{F}}$ of the face $\mathcal{F}$. Then, by the estimates (20), we have

$$\frac{1}{2(d{+}1)} \prod_{\mathbf{c}\in\mathcal{F}} w_{\mathbf{c}}^{\nu_{\mathbf{c}}} \;>\; z_{\mathbf{a}}^{\mu} z_{\mathbf{b}}^{1-\mu} \cdot \prod_{\mathbf{c}\in\mathcal{F}} w_{\mathbf{c}}^{\nu_{\mathbf{c}}} \;=\; w_{\mathbf{a}}^{\mu} w_{\mathbf{b}}^{1-\mu} \cdot \prod_{\mathbf{c}\in\mathcal{F}} z_{\mathbf{c}}^{\nu_{\mathbf{c}}} \;>\; w_{\mathbf{a}}^{\mu} w_{\mathbf{b}}^{1-\mu} \cdot \frac{1}{2(d{+}1)} \;\;.$$

canceling the common factor of $1/2(d{+}1)$ and taking logarithms, we obtain

$$\sum_{\mathbf{c}\in\mathcal{F}} \nu_{\mathbf{c}} \log(w_{\mathbf{c}}) \;>\; \mu \log(w_{\mathbf{a}}) + (1-\mu)\log(w_{\mathbf{b}}) \;\;,$$

which implies that the point (18) common to the segment $\overline{\mathbf{a},\mathbf{b}}$ and the face $\mathcal{F}$ of $\mathcal{T}$ is lifted higher in the face $\mathcal{F}$ than in the segment $\overline{\mathbf{a},\mathbf{b}}$, and so the lift of the segment $\overline{\mathbf{a},\mathbf{b}}$ by λ lies below the lift of the face $\mathcal{F}$. As this is true for all segments, we see that $\mathcal{T}$ is the triangulation induced by the lifting function λ.

References

1. Dahmen, W.: Convexity and Bernstein-Bézier polynomials. In: Curves and surfaces (Chamonix-Mont-Blanc, 1990). Academic Press, Boston, MA (1991) 107–134
2. Craciun, G., Feinberg, M.: Multiple equilibria in complex chemical reaction networks. I. The injectivity property. SIAM J. Appl. Math. **65**(5) (2005) 1526–1546 (electronic)
3. Krasauskas, R.: Toric surface patches. Adv. Comput. Math. **17**(1-2) (2002) 89–133 Advances in geometrical algorithms and representations.
4. Brown, L.D.: Fundamentals of statistical exponential families with applications in statistical decision theory. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 9. Institute of Mathematical Statistics, Hayward, CA (1986)
5. Darroch, J.N., Ratcliff, D.: Generalized iterative scaling for log-linear models. Ann. Math. Statist. **43** (1972) 1470–1480
6. Sottile, F.: Toric ideals, real toric varieties, and the moment map. In: Topics in algebraic geometry and geometric modeling. Volume 334 of Contemp. Math. Amer. Math. Soc., Providence, RI (2003) 225–240
7. Garcia-Puente, L.D., Sottile, F.: Linear precision for parametric patches (2009) Advances in Computational Mathematics, to appear.
8. Karčiauskas, K., Krasauskas, R.: Comparison of different multisided patches using algebraic geometry. In Laurent, P.J., Sablonniere, P., Schumaker, L., eds.: Curve and Surface Design: Saint-Malo 1999, Vanderbilt University Press, Nashville (2000) 163–172
9. Cox, D.: What is a toric variety? In: Topics in algebraic geometry and geometric modeling. Volume 334 of Contemp. Math. Amer. Math. Soc., Providence, RI (2003) 203–223
10. DeRose, T., Goldman, R., Hagen, H., Mann, S.: Functional composition algorithms via blossoming. ACM Trans. on Graphics **12** (1993) 113–135
11. Fulton, W.: Introduction to toric varieties. Volume 131 of Annals of Mathematics Studies. Princeton University Press, Princeton, NJ (1993) The William H. Roever Lectures in Geometry.
12. Agresti, A.: Categorical Data Analysis. Wiley series in Probability and Mathematical Statistics. Wiley, New York (1990)
13. Graf von Bothmer, H.C., Ranestad, K., Sottile, F.: Linear precision for toric surface patches (2007) `ArXiv:math/0806.3230`.
14. Pachter, L., Sturmfels, B., eds.: Algebraic statistics for computational biology. Cambridge University Press, New York (2005)
15. Keller, O.: Ganze cremonatransformationen. Monatschr. Math. Phys. **47** (1939) 229–306
16. Pinchuk, S.: A counterexample to the strong real Jacobian conjecture. Math. Z. **217**(1) (1994) 1–4
17. Dahmen, W., Micchelli, C.A.: Convexity of multivariate Bernstein polynomials and box spline surfaces. Studia Sci. Math. Hungar. **23**(1-2) (1988) 265–287
18. Leroy, R.: Certificats de positivité et minimisation polynomiale dans la base de Bernstein multivariée. PhD thesis, Institut de Recherche Mathématique de Rennes (2008)
19. Sturmfels, B.: Gröbner bases and convex polytopes. American Mathematical Society, Providence, RI (1996)