PATCH KESSLER

NOTES ON NURBS

MECHANICAL DUST

PUBLISHED BY MECHANICAL DUST

WWW.MECHANICALDUST.COM

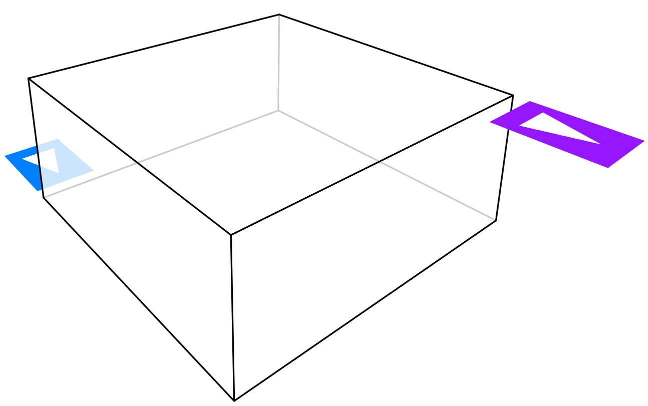
First printing, January 2020

Contents

Introduction

NURBS are widely used for designing beautiful curves and surfaces. Especially in Industrial Design, NURBS are universal. The casing of your phone, the hood of your car- these are all NURBS. These surfaces are crafted in a design process that is as intense and laborious as sculpting with clay. A designer nudges the surface this way or that to satisfy various constraints and boundary conditions. The space on one side of the surface may need to have a certain shape or volume (e.g., to enable a certain pattern of air flow). At the same time the optical reflections from the surface need to be just so.

NURBS are also used in engineering design. For instance NURBS can be used to solve for the road that connects two points in such a way as to optimize a given measure of passenger comfort.



The shape of a NURBS object is controlled by a collection of *control points* and *weights* in a way that is intuitive and local, so that moving a control point only affects nearby regions. Basically, NURBS are well behaved and easy to use. This is why they are so popular in a variety of disciplines.

These working notes track my progress learning about NURBS, from complete ignorance in April 2019 to somewhat better than that today. I hope these notes share some of the elegance, beauty, and joy I've discovered in the mathematics of NURBS and related constructions. These notes are less of a comprehensive review of NURBS, and more of an exploration of some of the interesting topics that come up when you begin working with NURBS.

- Chapters 1 & 2 consist of a simplex generalization of the Bezier construction, as well as a proof that this generalization can be used to build n-spheres. This is my own original work, however I'd be surprised if these topics haven't already been explored by others.
- Chapter 3 consists of a definition of the NURBS basis, including a convenient method of keeping track of its elements. Various properties of the NURBS basis are discussed and proved.
- Chapter 4 shows how to build curves, surfaces, and more using NURBS.
- Chapter 5 is a development of higher dimensional conics and conic sections.
- Chapter 6 consists of some basic ideas in projective geometry including Steiner conics. It was in my struggle to understand these topics that I began these notes in the first place.

It's hard to talk about something if you don't know what to call it. Observe that NURBS stands for

N*on* U*niform* R*ational* B*asis* S*pline*,

as well as

N*on* U*niform* R*ational* B*asis* S*plines*.

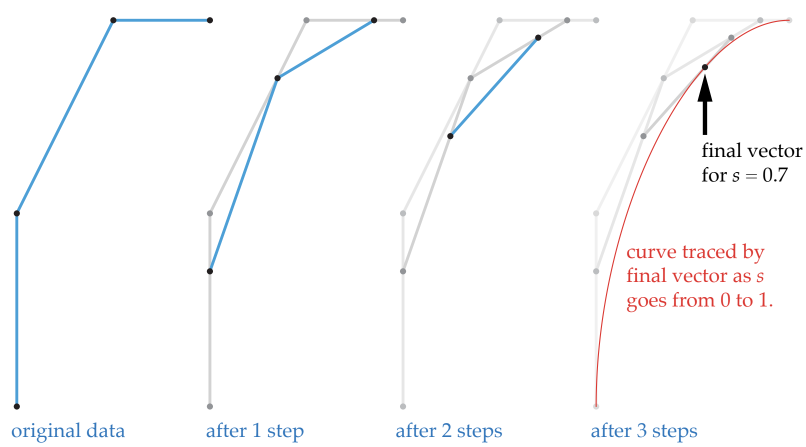
Because of this, we treat NURBS as an *irregular plural*, like fish. You can have one fish, or lots of fish. You can have one NURBS or lots of NURBS.

The Bezier Construction

A Bezier curve starts with a chain of n vectors $\mathbf{x}_1$, ..., $\mathbf{x}_n$. A chain of $n-1$ vectors $\mathbf{y}_1$, ..., $\mathbf{y}_{n-1}$, is then constructed with each $\mathbf{y}_i$ given by

$$\mathbf{y}_i = (1-s)\mathbf{x}_i + s\mathbf{x}_{i+1}$$

for $s \in [0,1]$. This construction step is repeated $n-1$ times with the same s, ending with a single vector. This single vector traces out a curve that follows the original chain as s goes from 0 to 1.

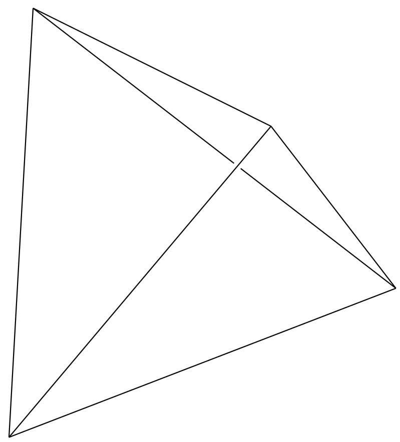


Higher Dimensions

We'd like to extend the Bezier construction to higher dimensional objects, such as surfaces and solids. The standard way of doing this is with Cartesian products. Instead of proceeding along the standard route, we explore moving into higher dimensions with simplexes.

Simplexes

Consider n vectors $\mathbf{u}_i$ as the vertices of a simplex. For instance if $n = 4$ we get a tetrahedron.



Barycentric Coordinates

Consider a set of *barycentric coordinates* $s_i \in \mathbb{R}$, which satisfy $\sum s_i = 1$. We map these s_i's to

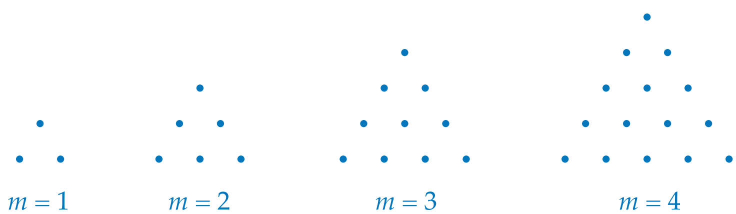
$$\mathbf{x} = s_1\mathbf{u}_1 + s_2\mathbf{u}_2 + \cdots + s_n\mathbf{u}_n.$$

Generally, $\mathbf{x}$ lives within an *affine* space defined by the $\mathbf{u}_i$'s. For instance, if there are two $\mathbf{u}_i$'s in $\mathbb{R}^3$, then $\mathbf{x}$ is restricted to the line through these two points, which doesn't necessarily pass through the origin. The sum $\mathbf{x} = \sum s_i\mathbf{u}_i$ is a blend of the $\mathbf{u}_i$'s, with $\mathbf{x} = \mathbf{u}_i$ when $s_k = \delta_{ki}$. The $\mathbf{u}_i$'s must all come from the same vector space of course, however this vector space can be anything. The $\mathbf{u}_i$'s can be temperatures in $\mathbb{R}$, positions in $\mathbb{R}^2$, colors in $\mathbb{R}^3$, continuous functions on $[0, 1]$, and so on.

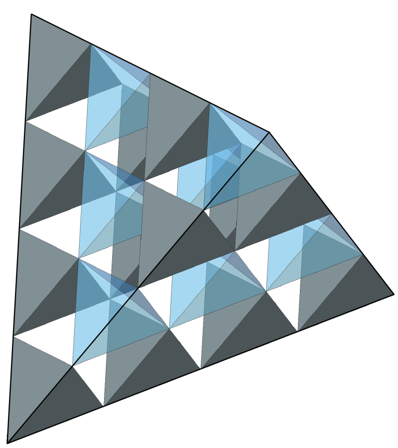
The symbol δ_{ki} is equal to 1 if $k = i$ and equal to 0 otherwise.

Simplex Grids

We form a *simplex grid* G_m by restricting the n coordinates s_i to discrete values k_i/m, where m is a positive integer and where each k_i is an integer between 0 and m. In addition, we inherit the requirement $\sum k_i = m$ from $\sum s_i = 1$. The k_i's are *integer coordinates* for the grid points. Here are some examples of simplex grids for $n = 3$.



Each point in a simplex grid G_m has integer coordinates $\mathbf{k} = (k_1, k_2, \ldots, k_n)$ with each $k_i \in \{0, 1, 2, \ldots, m\}$ and with $\sum k_i = m$. Consider the perturbations $\mathbf{p}_{ij} = (p_1, p_2, \ldots, p_n)$ with all entries zero except $p_i = -1$ and $p_j = 1$. If $\mathbf{k} + \mathbf{p}_{ij} \in G_m$ for fixed i and every $j \neq i$, then $\mathbf{k}$ is the vertex of a simplex defined by $\mathbf{k}$ and the points $\mathbf{k} + \mathbf{p}_{ij}$. We call this a *mini simplex* of G_m, and we say that $\mathbf{k}$ is vertex i of this mini simplex. In this way we label the vertices of each mini simplex 1 to n. These mini simplexes connect to each other at their vertices.



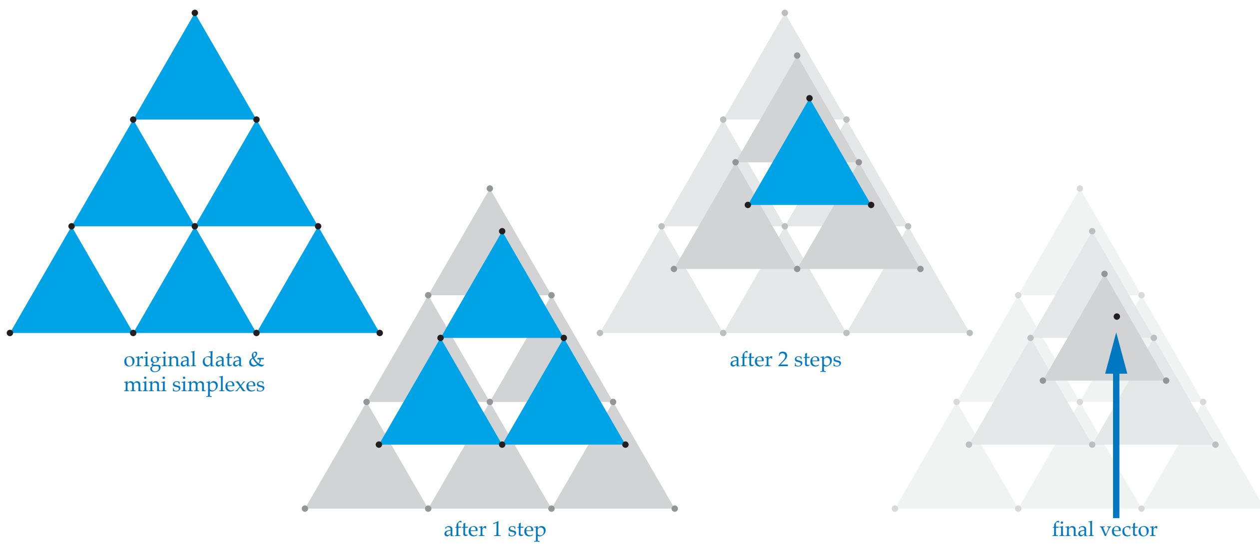
Although it is difficult to visualize a simplex grid for $n > 4$, there is no change in the mechanics of the bookkeeping as the dimension increases.

The Bezier Construction

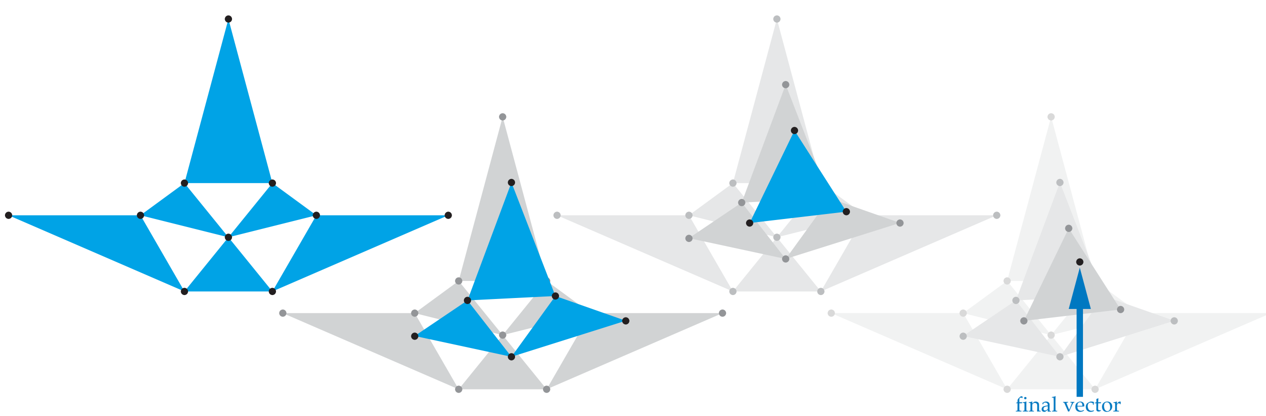
Consider a vector function defined on a simplex grid (i.e., a vector is assigned to each point in G_m). The Bezier construction consists of repeatedly performing the following Bezier step.

> Compute $\sum s_i \mathbf{u}_i$ for each mini simplex of G_k, where $\mathbf{u}_i$ are the vectors at the (labeled) vertices of the mini simplex. This results in a single vector if $k = 1$. Otherwise, the vectors $\sum s_i \mathbf{u}_i$ are associated with a new (smaller) simplex grid G_{k-1}.

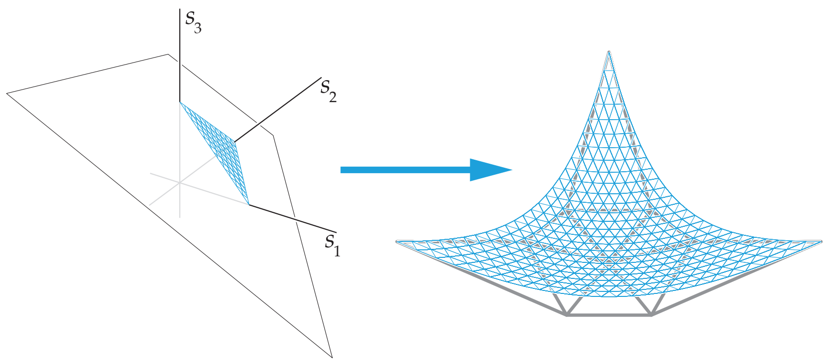
If we start with a simplex grid G_m, then m Bezier steps results in a single vector. This vector is a linear combination of the $\mathbf{u}_i$'s from the original function defined on G_m. The weights in this linear combination are polynomial functions of the s_i's. Here's an example of G_3 with $n = 3$, with $(s_1, s_2, s_3) = (0.1, 0.3, 0.6)$.



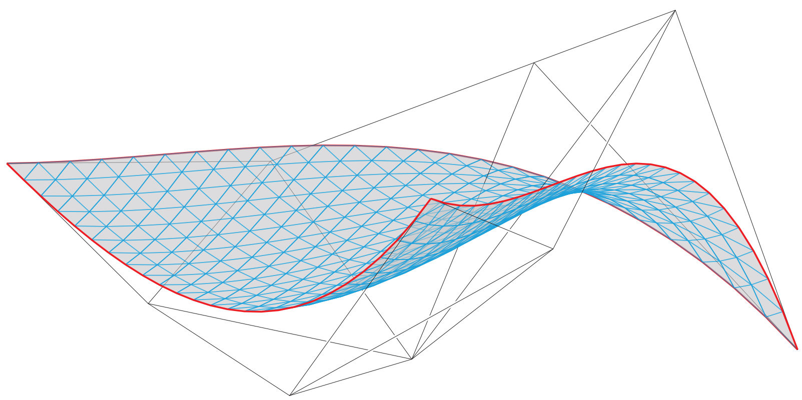
Here we are a little more adventurous with the original data.



Next, instead of a single point, consider the *set* of points with barycentric coordinates s_i that satisfy $s_i \geq 0$. The Bezier construction maps this set to a region that depends on the $\mathbf{u}_i$'s.



If the $\mathbf{u}_i$'s are points in $\mathbb{R}^3$, we get a surface in $\mathbb{R}^3$.



Notice how the surface is guided by the background grid. There is tangency at the corners, and the interior points seem to be attracting the surface. This intuitive connection between the surface and its control points makes it easy for a designer (or an algorithm) to adjust the surface shape.

Although we have illustrated the Bezier construction with low dimensional examples ($n = 3$ gives triangles while $n = 4$ gives tetrahedra), our construction techniques work for *any* value of n. Also, the dimensionality of the vertex data $\mathbf{u}_i$ is independent of n. The $\mathbf{u}_i$'s can come from *any* vector space, even one that is infinite dimensional.

Direct Expression

It is useful to have a direct (single step) expression for the output $\mathbf{x}$ of the Bezier construction. We find that the $\mathbf{x}$ generated by Barycentric coordinates $(s_1, s_2, \ldots, s_n)$ is given by

$$\mathbf{x} = \sum_{(k_1,k_2,\ldots,k_n)\in\mathcal{K}} s_1^{k_1} s_2^{k_2} \cdots s_n^{k_n} \mathbf{u}_{k_1,k_2,\ldots,k_n}$$

where $\mathbf{u}_{k_1,k_2,\ldots,k_n}$ is the vector associated with the point $(k_1, k_2, \ldots, k_n)$ in the simplex grid G_m, and where $\mathcal{K}$ is the set of all points in G_m. That is, $\mathcal{K}$ is the set of length n lists of integers k_i, with each k_i taken from $\{0, 1, 2, \ldots, m\}$ and with $\sum k_i = m$.

In the case $m = 2$, we can arrange the terms in a 2D array, and deal with the required multiplications and additions using the formalism of matrix multiplication.

$$\mathbf{x} = \begin{bmatrix} s_1 & s_2 & \cdots & s_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_{11} & \mathbf{v}_{12} & \cdots & \mathbf{v}_{1n} \\ \mathbf{v}_{21} & \mathbf{v}_{22} & \cdots & \mathbf{v}_{2n} \\ \vdots & \vdots & & \vdots \\ \mathbf{v}_{n1} & \mathbf{v}_{n2} & \cdots & \mathbf{v}_{nn} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

The vectors $\mathbf{v}_{ij}$ stand for the vectors $\mathbf{u}_{k_1,k_2,\ldots,k_n}$ as follows:

- When $i \neq j$, $\mathbf{v}_{ij}$ is the $\mathbf{u}_{k_1,k_2,\ldots,k_n}$ for which $k_i = k_j = 1$.
- When $i = j$, $\mathbf{v}_{ii}$ is the $\mathbf{u}_{k_1,k_2,\ldots,k_n}$ for which $k_i = 2$.

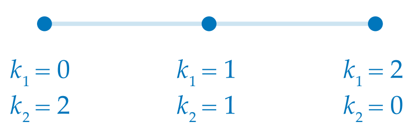
This covers all cases, because when $m = 2$, either two of the k_i's are 1 and the rest are zero, or one of the k_i's is 2 and the rest are zero.

Building A Sphere

It turns out that G_2 Bezier constructions exist that can be projected to form n-spheres for any n. What versatility! In the previous chapter we used Bezier constructions to create whimsical surfaces- here we use them to create perfect spheres.

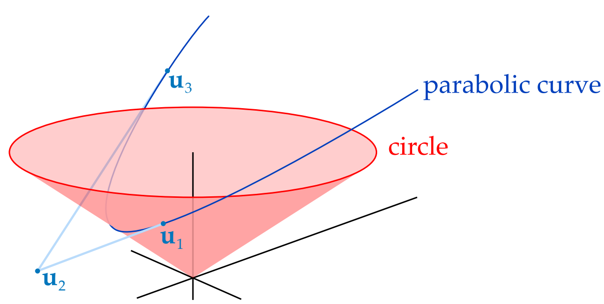
Start With A Circle

A simplex with $n = 2$ vertices generates a simplex grid G_2 that consists of a chain of three points.



If we assign a spatial position $\mathbf{u}_i \in \mathbb{R}^3$ to each point, the Bezier construction returns a parabolic space curve. We can build a general cone in $\mathbb{R}^3$ which has a given parabolic curve as one of its sections, and a general cone in $\mathbb{R}^3$ can always be sectioned to obtain a circle.

An example of a general cone in $\mathbb{R}^3$ is given by $\{(x, y, z)|\alpha x^2 + \beta y^2 = z^2\}$, with $\alpha, \beta > 0$. If $\alpha \neq \beta$, then there are two distinct planar orientations that result in circular sections. This leads to the *circles map to circles* property of stereographic projection.



Suppose the cone apex is at the origin and the $z = 1$ plane intersects the cone in a circle. Algebraically, if the parabolic curve is given by $(x(s), y(s), z(s))$, then the projection to the $z = 1$ plane is given by

$$(\frac{x(s)}{z(s)}, \frac{y(s)}{z(s)}, 1).$$

Thus a 1-sphere can be built as a ratio of Bezier construction components, (i.e., polynomial functions of the s_i's).

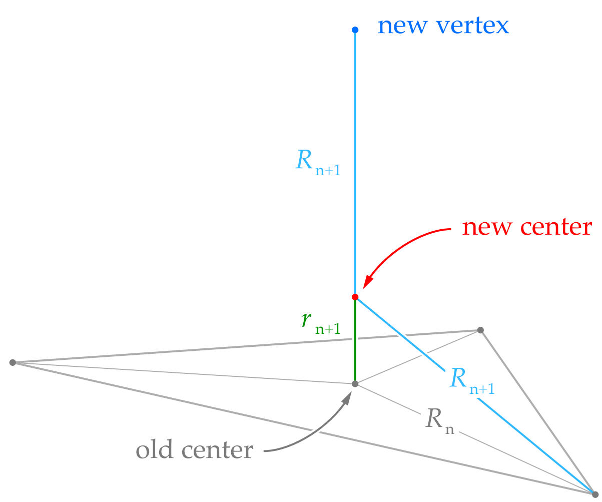
The situation is more involved when $n > 2$. In higher dimensions, many G_2 Bezier constructions are not analogous to well behaved parabolic space curves. For instance when $n = 3$, it is easy to generate a G_2 Bezier construction that intersects itself. In spite of this, we show that for each n, there exists at least one G_2 Bezier construction which projects to a perfect n-sphere.

The Unit Simplex

A *unit simplex* is a simplex with unit distance between all pairs of vertices. Given a unit simplex with n vertices,

- Let R_n be the center-to-vertex distance (i.e., the radius of the smallest ball containing the simplex).
- Let r_n be the center-to-face distance (i.e., the radius of the largest ball contained within the simplex).

Given a unit simplex with n vertices, we create one with $n+1$ vertices by adding a new vertex orthogonally offset from the center of the existing one.



From the Pythagorean theorem, we see that $R_n^2 + (R_{n+1} + r_{n+1})^2 = 1$, and $R_n^2 + r_{n+1}^2 = R_{n+1}^2$. Eliminating r_{n+1}, we find that R_n^2 satisfies the recursion

$$R_{n+1}^2 = \frac{1}{4(1 - R_n^2)}.$$

We know that $R_2^2 = 1/4$, and solving for the next couple of entries, we suspect that $R_n^2 = \frac{n-1}{2n}$. This is easily verified by induction. Next, let $\mathbf{u}_i$ be the vector from the center of the simplex to the ith vertex. Using basic trigonometry, we find that $\mathbf{u}_i \cdot \mathbf{u}_j = R_n^2 - \frac{1}{2} = \frac{-1}{2n}$ for $i \neq j$.

Normal Vectors

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, and consider the surface $S \subset \mathbb{R}^n \times \mathbb{R}$ consisting of points $(\mathbf{x}, f(\mathbf{x}))$ where $\mathbf{x} \in \mathbb{R}^n$. We'd like to find a vector $\mathbf{u}$ which is normal to the surface at a given point $(\mathbf{x}, f(\mathbf{x}))$. If we nudge $\mathbf{x}$ by $\partial\mathbf{x}$, we get a new point in the surface which is vanishingly close to $(\mathbf{x} + \partial\mathbf{x}, f(\mathbf{x}) + \nabla f\ \partial\mathbf{x})$. Thus $\mathbf{u}$ needs to be perpendicular to

$$\begin{bmatrix} \partial x_1 & \partial x_2 & \cdots & \partial x_n & A \end{bmatrix},$$

where $A = \frac{\partial f}{\partial x_1}\partial x_1 + \frac{\partial f}{\partial x_2}\partial x_2 + \cdots + \frac{\partial f}{\partial x_n}\partial x_n$. Clearly $\mathbf{u}$ is aligned with

$$\begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} & -1 \end{bmatrix}.$$

As an example, if $f(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x}$, then the normal to the corresponding surface at $(\mathbf{x}, f(\mathbf{x}))$ is given by $\mathbf{u} = (2\mathbf{x}, -1)$.

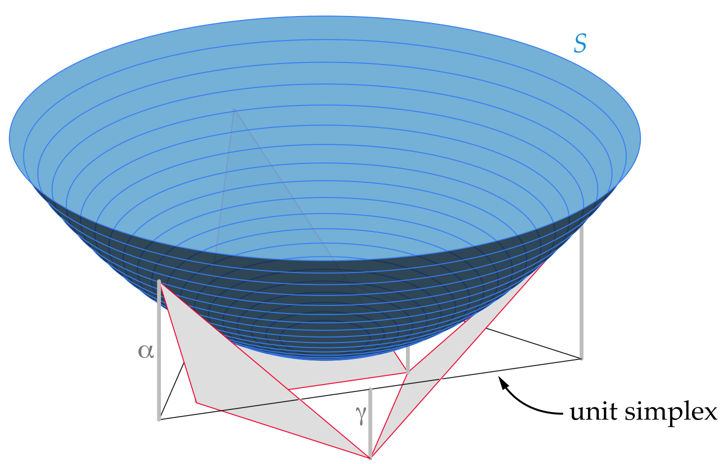
A Parabola in Higher Dimensions

Consider the subspace E of $\mathbb{R}^n$ consisting of vectors which are perpendicular to some unit vector $\mathbf{E}$. Let S be the parabolic surface in $\mathbb{R}^n$ consisting of points $\mathbf{x} + (\mathbf{x} \cdot \mathbf{x})\mathbf{E}$, where $\mathbf{x} \in E$.

We create this parabolic surface as a G_2 Bezier construction. Each node of this construction is indexed by a list $\mathbf{k} = (k_1, k_2, \ldots, k_n)$ of integers, where each $k_i \in \{0, 1, 2\}$ and where the sum of the k_i's is 2. We assign a vector $\mathbf{v}$ to each node depending on the $\mathbf{k}$ for that node. We build the $\mathbf{v}$'s out of the n vertices $\mathbf{u}_i$ of a dimension $n-1$ unit simplex contained in E and centered at the origin.

$$\mathbf{v} = \begin{cases} \mathbf{u}_i + \alpha\mathbf{E} & \text{if the } i\text{th element of } \mathbf{k} \text{ equals } 2. \\ \beta(\mathbf{u}_i + \mathbf{u}_j) - \gamma\mathbf{E} & \text{if elements } i \text{ and } j \text{ of } \mathbf{k} \text{ equal } 1. \end{cases}$$

This covers all possible cases, because either one of the k_i's is 2 and the rest are zero, or two of the k_i's are 1 and the rest are zero. The idea is that the $\mathbf{v}$'s assigned to the *vertices* of the unit simplex are coincident with S, while the other $\mathbf{v}$'s result in planes that are tangent to S (provided α, β, and γ are chosen properly). Here's what it looks like when $n = 3$.



Because the $\mathbf{u}_i$'s satisfy $||\mathbf{u}_i|| = R_n$, we have $\alpha = R_n^2 = (n-1)/(2n)$. This takes care of $\mathbf{v}_{ii} = \mathbf{u}_i + \alpha\mathbf{E}$. To find $\mathbf{v}_{ij} = \beta(\mathbf{u}_i + \mathbf{u}_j) - \gamma\mathbf{E}$, we observe two additional requirements.

First, the G_2 Bezier construction needs to pass through the origin. That is, we need for there to be a collection of s_i's that map to zero. Because of symmetry in the $\mathbf{v}_{ij}$'s, we can immediately get to zero in every direction but $\mathbf{E}$ by making all the s_i's equal (to $1/n$). To also get to zero in the $\mathbf{E}$ direction, note that $(s_1, s_2, \ldots, s_n)$ with each $s_i = 1/n$ gets mapped by the Bezier construction to $\mathbf{x} = \sum \mathbf{v}_{ij}$. Considering the $\mathbf{E}$ direction we see that $\gamma = 1/2n$ causes $\mathbf{x} = \mathbf{0}$.

Our second requirement is tangency between the parabola and the planes formed by the $\mathbf{v}_{ij}$'s. The parabola surface normal at $\mathbf{v}_{ii}$ is given by $\mathbf{n}_{ii} = 2\mathbf{u}_i - \mathbf{E}$, and we must have $(\mathbf{v}_{ii} - \mathbf{v}_{ij}) \cdot \mathbf{n}_{ii} = 0$. With $\gamma = 1/2n$, this condition causes $\beta = 1/2$. Thus the $\mathbf{v}_{ij}$'s are given by

$$\begin{aligned} \mathbf{v}_{ii} &= \mathbf{u}_i + \frac{n-1}{2n}\mathbf{E}, \\ \mathbf{v}_{ij} &= \frac{1}{2}(\mathbf{u}_i + \mathbf{u}_j) - \frac{1}{2n}\mathbf{E}. \end{aligned}$$

At this point we have a fully specified G_2 Bezier construction, but we don't know for sure that it is parabolic. We prove this now. Let $\{\mathbf{e}_i\}$ be an orthonormal basis for $\mathbb{R}^n$ with $\mathbf{e}_n = \mathbf{E}$. Let $\mathbf{s}$ be a list of Barycentric coordinates $(s_1, s_2, \ldots, s_n)$ with $\sum s_i = 1$, and let $\mathbf{x}(\mathbf{s})$ be the output of the G_2 Bezier construction using the $\mathbf{v}_{ij}$'s as input. Recall from the end of the previous chapter that

$$\mathbf{x} = \begin{bmatrix} s_1 & s_2 & \cdots & s_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_{11} & \mathbf{v}_{12} & \cdots & \mathbf{v}_{1n} \\ \mathbf{v}_{21} & \mathbf{v}_{22} & \cdots & \mathbf{v}_{2n} \\ \vdots & \vdots & & \vdots \\ \mathbf{v}_{n1} & \mathbf{v}_{n2} & \cdots & \mathbf{v}_{nn} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix},$$

and note that

$$\begin{aligned} \mathbf{x} - \mathbf{E}(\mathbf{x} \cdot \mathbf{E}) &= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} s_i s_j (\mathbf{u}_i + \mathbf{u}_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} s_i s_j \mathbf{u}_i \\ &= \sum_{i=1}^{n} s_i \mathbf{u}_i, \end{aligned} \qquad (1)$$

where to obtain the last equality we use the fact that $\sum s_i = 1$.

We now prove our main result, which is that

$$\sum_{i=1}^{n-1} (\mathbf{x} \cdot \mathbf{e}_i)^2 = \mathbf{x} \cdot \mathbf{e}_n. \tag{2}$$

Evaluating the right side of (2), we find that

$$\mathbf{x} \cdot \mathbf{e}_n = \frac{1}{2} \sum_{i=1}^{n} s_i^2 - \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{n} s_i s_j. \tag{3}$$

The left side of (2) can be expanded with the help of (1).

$$\begin{aligned} \sum_{i=1}^{n-1} (\mathbf{x} \cdot \mathbf{e}_i)^2 &= ||\mathbf{x} - \mathbf{E}(\mathbf{x} \cdot \mathbf{E})||^2 \\ &= (\sum_{i=1}^{n} s_i \mathbf{u}_i) \cdot (\sum_{j=1}^{n} s_j \mathbf{u}_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} s_i s_j \mathbf{u}_i \cdot \mathbf{u}_j \end{aligned}$$

This final term is equal to (3) because $\mathbf{u}_i \cdot \mathbf{u}_i = R_n^2 = 1/2 - 1/2n$, and $\mathbf{u}_i \cdot \mathbf{u}_j = -1/2n$ for $i \neq j$. Thus we have proven (2).

Is it a Conic Section?

We have created a general parabola consisting of points $\mathbf{y} + (\mathbf{y} \cdot \mathbf{y})\mathbf{e}_n$, where $\mathbf{y} \in \text{span}(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-1})$, and where the $\mathbf{e}_i$'s comprise an orthonormal basis for $\mathbb{R}^n$. Let $\mathbf{e}_{n+1}$ be a new unit vector, orthogonal to all the rest, and let

$$\mathbf{x} = \mathbf{y} + (\mathbf{y} \cdot \mathbf{y})\mathbf{e}_n + \gamma \mathbf{e}_{n+1}. \tag{4}$$

The points $\mathbf{x}$ comprise a conic section in $\mathbb{R}^{n+1}$ if we can find a γ and a symmetric linear operator $\mathbf{A}$ on $\mathbb{R}^{n+1}$ with one negative eigenvalue and the rest positive, such that $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$. Drawing on our intuition from the case $n = 2$, (see beginning of this chapter), we consider the $\mathbf{A}$ which does nothing to $\mathbf{y}$, and which interchanges $\mathbf{e}_n$ and $\mathbf{e}_{n+1}$.

See the chapter on Conic Sections.

$$\begin{aligned} \mathbf{A}\mathbf{e}_i &= \mathbf{e}_i \text{ for } i = 1, 2, \ldots, n-1 \\ \mathbf{A}\mathbf{e}_n &= \mathbf{e}_{n+1} \\ \mathbf{A}\mathbf{e}_{n+1} &= \mathbf{e}_n \end{aligned}$$

This $\mathbf{A}$ has eigenvalues -1 (multiplicity 1) and $+1$ (multiplicity n), and $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$ iff $\gamma = -1/2$. This shows that our general parabola is indeed the section of a cone.

Projection

We know that the points $\mathbf{x}$ from (4) comprise a parabolic section of the cone given by $\mathbf{A}$. We now scale these points so that they are projected onto a plane perpendicular to the cone axis. The positive eigenvalues of $\mathbf{A}$ all equal $+1$, and so sectioning the cone with a plane perpendicular to the cone axis gives us an n-sphere.

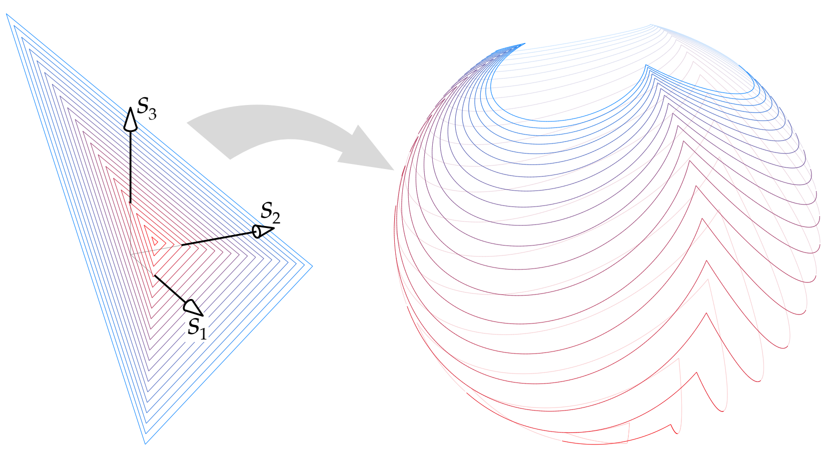
In detail, let $\mathbf{b}_i$ denote a basis of unit eigenvectors of $\mathbf{A}$, with $\mathbf{b}_i = \mathbf{e}_i$ for $i = 1$ to $n-1$, and with

$$\mathbf{b}_n = \frac{\mathbf{e}_n + \mathbf{e}_{n+1}}{\sqrt{2}} \text{ and } \mathbf{b}_{n+1} = \frac{\mathbf{e}_n - \mathbf{e}_{n+1}}{\sqrt{2}}.$$

The cone axis is aligned with $\mathbf{b}_{n+1}$. We scale each $\mathbf{x}$ by $1/(\mathbf{x} \cdot \mathbf{b}_{n+1})$, so that the resulting point $\mathbf{w}$ satisfies $\mathbf{w} \cdot \mathbf{b}_{n+1} = 1$. That is, we scale $\mathbf{x}$ so that it is on the plane perpendicular to the cone axis, and passing through the point $\mathbf{b}_{n+1}$. If w_i are the components of $\mathbf{w}$ with respect to the $\mathbf{b}_i$ basis, then $w_{n+1} = 1$. Also, note that $\mathbf{w}$ remains on the cone defined by $\mathbf{A}$, and so $\mathbf{w}^T\mathbf{A}\mathbf{w} = 0$. Combining these results, we get

$$w_1^2 + w_2^2 + \cdots + w_n^2 = 1,$$

and so the projected points $\mathbf{w}$ do indeed lie on a sphere. As an example, the following image shows a computation when $n = 3$. The $\sum s_i = 1$ plane on the left is mapped to a 2-sphere on the right.



Stereographic Projection

Although it's a diversion, we can't help but notice that straight lines in the plane get mapped to circles that pass through the north pole of the sphere, suspiciously like stereographic projection. How surprising! It's like running into an old friend in the middle of a remote city on the other side of the world!

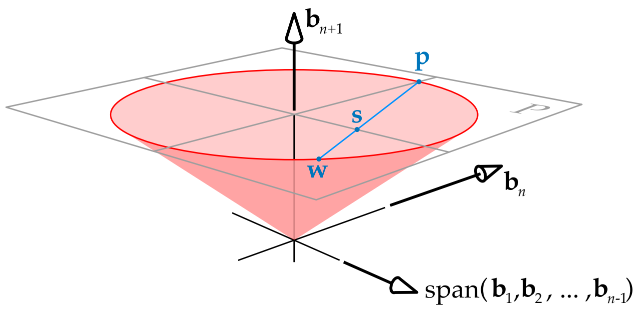
I explore various properties of stereographic projection in my *Notes On Stereographic Projection* which can be found at `http://www.mechanicaldust.com/`

Expanding $\mathbf{w}$ from the previous section, we find that

$$\mathbf{w} = \sqrt{2} \cdot \frac{\mathbf{u} + (\mathbf{u} \cdot \mathbf{u})\mathbf{e}_n - (1/2)\mathbf{e}_{n+1}}{\mathbf{u} \cdot \mathbf{u} + 1/2},$$

where $\mathbf{u} \in \text{span}(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-1})$. From (1) note that $\mathbf{u} = \sum s_i \mathbf{u}_i$ is the position located by the Barycentric coordinates s_i with respect to the Barycentric basis $\mathbf{u}_i$. The point $\mathbf{w}$ is on a sphere within plane P consisting of points $\mathbf{x}$ such that $\mathbf{x} \cdot \mathbf{b}_{n+1} = 1$. We stereographically project with respect to the point $\mathbf{p} = \mathbf{b}_n + \mathbf{b}_{n+1}$, which we think of as the *north pole* of the sphere.

The Barycentric coordinates s_i satisfy $\sum s_i = 1$, and are used to locate a point $\sum s_i \mathbf{u}_i$ in an affine space. We call the $\mathbf{u}_i$'s a Barycentric basis for this space.



The stereographic projection of $\mathbf{w}$ is given by $\mathbf{s} = \mathbf{p} + \eta(\mathbf{w} - \mathbf{p})$, with $\eta = 1/(1 - \mathbf{w} \cdot \mathbf{b}_n)$ chosen so that $\mathbf{s} \cdot \mathbf{b}_n = 0$. Expanding and canceling terms, we find that $\mathbf{s} = \mathbf{u}\sqrt{2} + \mathbf{b}_{n+1}$, as suspected! That is, we find that the stereographic projection of $\mathbf{w}$ is a scaled copy of $\mathbf{u}$. The inverse projection

$$\mathbf{w} = \mathbf{p} + \frac{2(\mathbf{s} - \mathbf{p})}{||\mathbf{s} - \mathbf{p}||^2},$$

allows us to circumvent all the steps from this chapter- the $\mathbf{v}_{ij}$'s, the parabola, the general cone, and the cone section. Of course, we needed these steps to know where we were going- but now that we're there, we have a much cleaner way to perform computations.

The Big Picture

It's time to take stock of what we've done. In the previous chapter we used simplexes to generalize the Bezier construction to higher dimensions. Then in this chapter we proved that a G_2 Bezier construction could be used to generate a perfect sphere in any dimension. Along the way, we encountered higher dimensional parabolas, cones, conic sections, and spheres, and we made a discovery that everything is related by stereographic projection. These connections are why all of this is worthwhile. If we only wanted to create a point on an n-sphere, we could take any $\mathbf{x} \in \mathbb{R}^{n+1}$, and then scale it by $1/||\mathbf{x}||$.

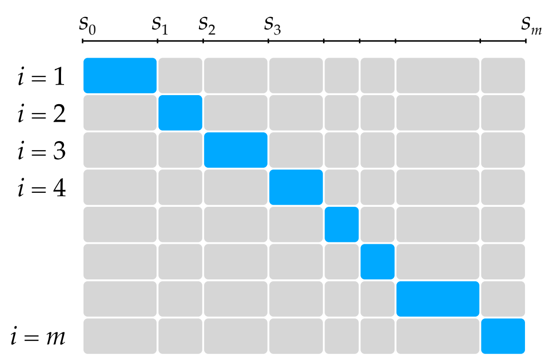
The NURBS Basis

B-Spline Basis Functions

We start by partitioning an interval of $\mathbb{R}$ with a sequence of non-decreasing values s_i which are referred to as *knots*.

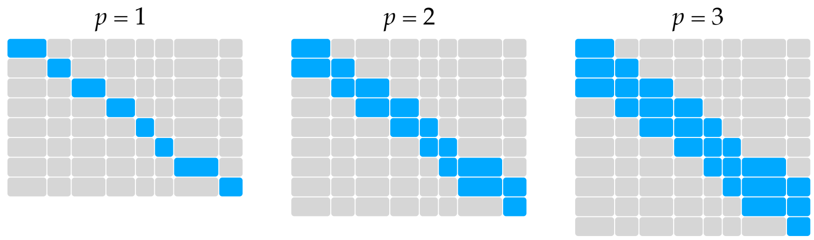
$$s_0 \le s_1 \le s_2 \le \cdots \le s_m$$

We define $N_{i,1}(s)$ to be the function with value 1 on $[s_{i-1}, s_i]$, and value 0 everywhere else. There are m of these functions, and we keep track of them with the following *box diagram*.



The ith row corresponds to the function $N_{i,1}(s)$. The gray boxes show intervals over which $N_{i,1}(s)$ is the zero function while the blue boxes show intervals over which $N_{i,1}(s)$ has non-zero values.

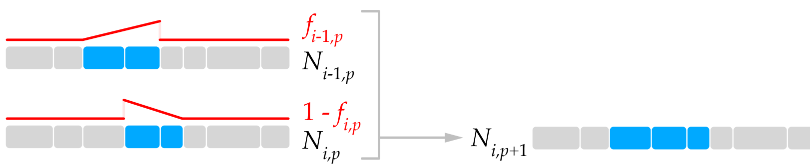
We generate additional $N_{i,p}$'s by recursion. With each recursive step, a portion of each row dribbles into the row below. This increases the number of rows in the box diagram, as well as the number of blue boxes in each row. The total number of $N_{i,p}$'s for a given p and m is the number of rows, which is $m + p - 1$. Also, p is the number of $N_{i,p}$'s that are non-zero between adjacent knots.



The recursive step is

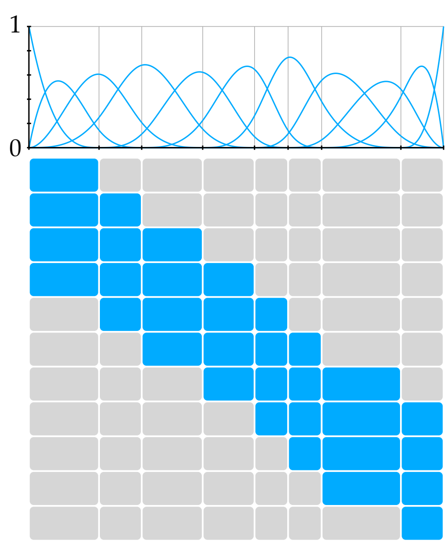
$$N_{i,p+1} = f_{i-1,p} N_{i-1,p} + (1 - f_{i,p}) N_{i,p},$$

where $f_{i,p}(s)$ increases linearly from 0 to 1 over the support for $N_{i,p}$, and where $N_{0,p}$ and $f_{0,p}$ equal zero for all p.



The recursive index p is the *order* of each of the polynomial components of $N_{i,p}$. Here's a graph of the $N_{i,p}$ functions when $p = 4$, together with the corresponding box diagram.

A degree k polynomial has order $k + 1$. The degree is the highest power; the order is the number of coefficients, e.g., $f(s) = a_0 + a_1 s + a_2 s^2 + a_3 s^3 + \cdots + a_k s^k$



Partition of Unity

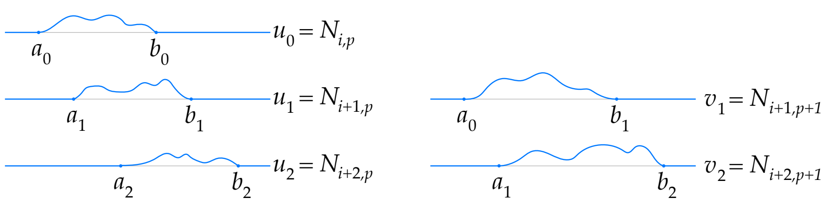
The $N_{i,p}$'s are a partition of unity, which is a fancy way of saying that they sum to one. To prove this, note that during a recursive step, each $N_{i,p}$ is divided into two pieces over its support, and one of these pieces is moved down to the next row in the box diagram. The two pieces have the same sum as the original, and so the overall sum of the $N_{i,p}$'s is unchanged by the recursive step. This sum is 1 by definition when $p = 1$, and so it is 1 for all p.

Continuity

Away from the origin, the product $x \cdot f(x)$ isn't any smoother than $f(x)$. As an example, suppose that f is C^{n-1}, but that $f^{[n]}$ is discontinuous at some nonzero x_0. Note that

$$(xf)^{[n]} = nf^{[n-1]} + xf^{[n]}, \tag{5}$$

and so $(xf)^{[n]}$ also has a discontinuity at x_0. Because of this, we are surprised that the $N_{i,p}$'s get *smoother* with each recursive step. Consider the following functions.



Suppose the piecewise polynomial u_i's are C^{n-1} but not C^n. That is, suppose the piecewise linear $u_i^{[n-1]}$'s are continuous, while the piecewise constant $u_i^{[n]}$'s are discontinuous at the connecting points. Let $f_i = (x - a_i)/(b_i - a_i)$, and recall that

$$v_i = f_{i-1}u_{i-1} - (1 - f_i)u_i.$$

Using (5), we see that

$$v_1^{[n]} = n(\alpha_0 u_0^{[n-1]} - \alpha_1 u_1^{[n-1]}) + f_0 u_0^{[n]} + (1 - f_1)u_1^{[n]}, \tag{6}$$

where $\alpha_i = 1/(b_i - a_i)$. In order for $v_1^{[n]}$ to be continuous at knot s, we need $v_{1,s-}^{[n]} = v_{1,s+}^{[n]}$. We achieve this with

We use f_{s-} for the limit of $f(x)$ as $x < s$ approaches s. Likewise we use f_{s+} for the limit of $f(x)$ as $x > s$ approaches s.

$$f_0(s) \cdot (u_{0,s+}^{[n]} - u_{0,s-}^{[n]}) + (1 - f_1(s)) \cdot (u_{1,s+}^{[n]} - u_{1,s-}^{[n]}) = 0. \tag{7}$$

This condition couples together the discontinuities in the $u_i^{[n]}$ functions at each knot value. We prove (7) with an inductive argument.

When $n = 0$, the u_i's are the constant $N_{i,p}$'s (with $p = 1$), and the condition follows immediately. For the inductive step, we assume (7) for all the $u_i^{[n]}$ functions at all the knot values. We need the same condition to hold for the $v_i^{[n+1]}$ functions. From (6), we see that

$$\begin{aligned} v_2^{[n+1]} &= (n+1)(\alpha_1 u_1^{[n]} - \alpha_2 u_2^{[n]}) \\ v_3^{[n+1]} &= (n+1)(\alpha_2 u_2^{[n]} - \alpha_3 u_3^{[n]}) \end{aligned}$$

Condition (7) for the $v_i^{[n+1]}$ functions is

$$F = g_2(s) \cdot (v_{2,s+}^{[n+1]} - v_{2,s-}^{[n+1]}) + (1 - g_3(s)) \cdot (v_{3,s+}^{[n+1]} - v_{3,s-}^{[n+1]}),$$

where $g_i = (x - a_{i-1})/(b_i - a_{i-1})$. Expanding terms, we obtain

$$\begin{aligned} F = (n+1)\Bigg[& \frac{s - a_1}{b_2 - a_1}\,\frac{1}{b_1 - a_1}\,(u_{1,s+}^{[n]} - u_{1,s-}^{[n]}) - \frac{s - a_1}{b_2 - a_1}\,\frac{1}{b_2 - a_2}\,(u_{2,s+}^{[n]} - u_{2,s-}^{[n]}) \\ & + \frac{b_3 - s}{b_3 - a_2}\,\frac{1}{b_2 - a_2}\,(u_{2,s+}^{[n]} - u_{2,s-}^{[n]}) - \frac{b_3 - s}{b_3 - a_2}\,\frac{1}{b_3 - a_3}\,(u_{3,s+}^{[n]} - u_{3,s-}^{[n]}) \Bigg] \end{aligned}$$

Notice that the coefficients of $u_{2,s+}^{[n]} - u_{2,s-}^{[n]}$ satisfy

$$\frac{b_3 - s}{b_3 - a_2} - \frac{s - a_1}{b_2 - a_1} = \frac{b_2 - s}{b_2 - a_1} - \frac{s - a_2}{b_3 - a_2},$$

thus allowing us to write

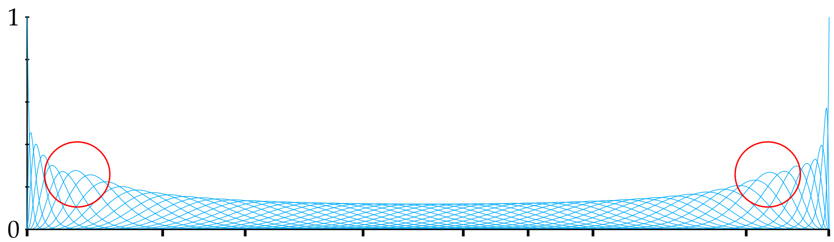
$$\begin{aligned} F = (n+1)\Bigg[& \frac{1}{b_2 - a_1} \boxed{\frac{s - a_1}{b_1 - a_1}\,(u_{1,s+}^{[n]} - u_{1,s-}^{[n]})} - \frac{1}{b_3 - a_2} \boxed{\frac{s - a_2}{b_2 - a_2}\,(u_{2,s+}^{[n]} - u_{2,s-}^{[n]})} \\ & + \frac{1}{b_2 - a_1} \boxed{\frac{b_2 - s}{b_2 - a_2}\,(u_{2,s+}^{[n]} - u_{2,s-}^{[n]})} - \frac{1}{b_3 - a_2} \boxed{\frac{b_3 - s}{b_3 - a_3}\,(u_{3,s+}^{[n]} - u_{3,s-}^{[n]})} \Bigg] \end{aligned}$$

zero by hypothesis   zero by hypothesis

It follows that $F = 0$ as desired, and so the v_i's are C^{n+1}. Translating over to the $N_{i,p}$'s we see that each $N_{i,p}$ is C^{p-2}.

The above development breaks down if any knot has multiplicity greater than 1. When this happens, continuity at the knot decreases. If a knot s has multiplicity k, then the $N_{i,p}$'s are C^{p-k-1} at s.

In the following image we consider the $N_{i,p}$ functions when $p = 50$, using the same irregular knot values as before. As p increases, it seems that the $N_{i,p}$ functions become more even. What is the envelope that these are approaching? And what's the reason for the irregularity circled in red? Is it numerical?



Derivatives

It turns out that the derivative of each $N_{i,p}$ can be expressed as a linear combination of $N_{i,p-1}$'s. Taking $p-2$ derivatives, we obtain a linear combination of $N_{i,2}$ functions, which are continuous. This provides us with a different way to prove that the $N_{i,p}$'s are C^{p-2}.

Basis of What?

The $N_{i,p}$'s span the space of piecewise polynomial functions that consist of order p polynomials on the intervals $[s_{i-1}, s_i]$, and that are C^{p-2} at each of the internal knots s_1 to s_{m-1}. Summing up degrees of freedom we find that this space has dimension $m+p-1$, which is the number of $N_{i,p}$ functions. The $N_{i,p}$'s comprise a basis for this space.

NURBS Constructions

Building Curves

We use the $N_{i,p}$ functions to build a curve $\mathbf{x}(s)$ that follows a chain of $n = m + p - 1$ control points $\mathbf{u}_i$ (i.e., one $\mathbf{u}_i$ for each $N_{i,p}$).

$$\mathbf{x}(s) = N_{1,p}(s)\mathbf{u}_1 + N_{2,p}(s)\mathbf{u}_2 + \cdots + N_{n,p}(s)\mathbf{u}_n$$

As s goes from s_0 to s_m, the point $\mathbf{x}(s)$ follows the control points $\mathbf{u}_i$ in a way that depends on p. If $p = 0$, then $\mathbf{x}(s)$ jumps from each control point to the next. If $p = 1$, we get linear interpolation. As p increases, the curve $\mathbf{x}(s)$ gets smoother and smoother- still influenced by the control points however no longer necessarily passing through them.

The control points $\mathbf{u}_i$ and resulting path $\mathbf{x}(s)$ can live in *any* vector space- for instance they could live in the space of 15 megapixel digital photographs, in which case our path would consist of a morphing photograph, transitioning smoothly from one image to the next.

Projection

We now project the curve $\mathbf{x}(s)$ onto a hyperplane. There are several reasons for doing this, for instance it enables us to form perfect circles. Let $\mathbf{e}$ be a unit vector, and consider the hyperplane E consisting of vectors $\mathbf{x}$ such that $\mathbf{x}^T\mathbf{e} = 1$. Typically we work in a vector space consisting of $n \times 1$ arrays of numbers, and we choose $\mathbf{e}$ to be the unit vector with all zeros and 1 in the last entry. The control points $\mathbf{u}_i$ can be written as

$$\mathbf{u}_i = w_i \begin{bmatrix} \mathbf{v}_i \\ 1 \end{bmatrix},$$

where $w_i = \mathbf{e}^T\mathbf{u}_i$, and $w_i\mathbf{v}_i$ is everything else. The $\mathbf{e}$ component of $\mathbf{x}(s)$ is given by

$$W(s) = N_{1,p}(s)w_1 + N_{2,p}(s)w_2 + \cdots + N_{n,p}(s)w_n,$$

and so $\mathbf{x}(s)$ can be projected onto the hyperplane E by multiplying it by $1/W(s)$.

$$\frac{1}{W(s)}\mathbf{x}(s) = \frac{1}{W(s)}\sum_{i=1}^{n} N_{i,p}(s) w_i \begin{bmatrix} \mathbf{v}_i \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{y}(s) \\ 1 \end{bmatrix}$$

The function $\mathbf{y}(s)$ is a NURBS curve.

$$\mathbf{y}(s) = \frac{1}{W(s)}\sum_{i=1}^{n} N_{i,p}(s) w_i \mathbf{v}_i = \frac{\sum_{i=1}^{n} N_{i,p}(s) w_i \mathbf{v}_i}{\sum_{i=1}^{n} N_{i,p}(s) w_i}$$

Higher Dimensions

We'd like to develop something like the NURBS basis elements for the simplex grids we developed in a previous chapter.

Conic Sections

A cone in $\mathbb{R}^n$ consists of vectors $\mathbf{x}$ which satisfy $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$, where $\mathbf{A}$ is symmetric with one eigenvalue negative and the rest positive. The Spectral Theorem guarantees an orthonormal basis $(\mathbf{e}_i)$ of corresponding eigenvectors. The components of $\mathbf{x}$ with respect to this basis satisfy

This chapter relies on a grounding in linear algebra. For this I recommend

Sheldon Axler. *Linear Algebra Done Right*. Springer, 1996. ISBN 0387982582

$$\alpha_1 x_1^2 + \alpha_2 x_2^2 + ... + \alpha_{n-1} x_{n-1}^2 = \alpha_n x_n^2,$$

where all the α_i's are positive, and where elements are labeled so that $\mathbf{e}_n$ corresponds to the negative eigenvalue $-\alpha_n$. The cone axis is aligned with $\mathbf{e}_n$, and we note that a plane perpendicular to $\mathbf{e}_n$ only sections the cone in a circle (or n-sphere) if $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ are equal.

Conic Sections in General

Suppose that $\mathbf{x}$ belongs to a cone, and that $\mathbf{x}$ also lies on a plane in $\mathbb{R}^n$ given by $\mathbf{x}^T\mathbf{u} = \gamma$, where $\mathbf{u}$ is a unit vector and $\gamma \in \mathbb{R}$. Motivated by conic sections in $\mathbb{R}^3$, we expect to decompose $\mathbf{x}$ as the sum $\mathbf{w} + \mathbf{y}$, where $\mathbf{w}$ is fixed and where $\mathbf{y}^T\mathbf{u} = 0$ with $\mathbf{y}$ a point on a dimension $n - 1$ ellipsoid (or perhaps hyperboloid or paraboloid).

Let $f(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T\mathbf{A}\mathbf{v}$ so that vectors $\mathbf{x}$ on the cone satisfy $f(\mathbf{x}, \mathbf{x}) = 0$. Note that

$$f(\mathbf{w} + \mathbf{y}, \mathbf{w} + \mathbf{y}) = f(\mathbf{w}, \mathbf{w}) + f(\mathbf{y}, \mathbf{y}) + 2f(\mathbf{w}, \mathbf{y}).$$

The last term on the right vanishes if we choose $\mathbf{w}$ such that $\mathbf{A}\mathbf{w} = \mathbf{u}$, because $\mathbf{y}^T\mathbf{u} = 0$ by hypothesis. We now have two classes to consider, depending on whether $\mathbf{u}$ and $\mathbf{w}$ are orthogonal.

Class 1
If $\mathbf{u}$ and $\mathbf{w}$ are not orthogonal, we scale $\mathbf{w}$ so that $\mathbf{w}^T\mathbf{u} = \gamma$. Observe that vectors $\mathbf{y}$ satisfy

$$f(\mathbf{y}, \mathbf{y}) = -f(\mathbf{w}, \mathbf{w}) = \text{constant.} \tag{8}$$

The function f is capable of acting on all vectors in $\mathbb{R}^n$, however we only give it vectors $\mathbf{y}$ that are perpendicular to $\mathbf{u}$. These vectors comprise an $n-1$ dimensional subspace B. Thus $f : B \times B \longrightarrow \mathbb{R}$.

Fixing one of the arguments of f equal to some vector $\mathbf{u}$ leaves us with a linear map from B to $\mathbb{R}$ (i.e., a *functional*). B inherits an inner product from $\mathbb{R}^n$, and so a functional $\varphi : B \longrightarrow \mathbb{R}$ is equivalent to taking an inner product with some vector in B. That is, $\varphi(\mathbf{x}) = \mathbf{v} \cdot \mathbf{x}$ for some $\mathbf{v} \in B$. Thus we used f and the inner product on B to move from $\mathbf{u} \in B$ to $\mathbf{v} \in B$. This mapping is linear and we call it F.

The mapping $F : B \longrightarrow B$ is also symmetric (i.e., $F(\mathbf{u}) \cdot \mathbf{v} = F(\mathbf{v}) \cdot \mathbf{u}$), and so it follows from the Spectral Theorem that F has real eigenvalues λ_i, and eigenvectors $\mathbf{u}_i$ that form an orthonormal basis for B. Expressing $\mathbf{y}$ with respect to this basis as $\sum y_i \mathbf{u}_i$, we obtain $f(\mathbf{y}, \mathbf{y}) = F(\mathbf{y}) \cdot \mathbf{y} = \sum \lambda_i y_i^2$. This allows us to write (8) as

$$\frac{\lambda_1}{w} y_1^2 + \frac{\lambda_2}{w} y_2^2 + \cdots + \frac{\lambda_{n-1}}{w} y_{n-1}^2 = 1$$

where $w = -f(\mathbf{w}, \mathbf{w})$. This results in an ellipsoid if the λ_i / w terms are positive. We can write this as $\mathbf{y}^T\mathbf{B}\mathbf{y} = 1$, where $\mathbf{B}$ is $n-1 \times n-1$ and symmetric.

Class 2
If $\mathbf{u}$ and $\mathbf{w}$ are orthogonal, we decompose $\mathbf{x}$ as

$$\mathbf{x} = \gamma\mathbf{u} + w\mathbf{w} + \mathbf{y},$$

where $\mathbf{y}$ is orthogonal to $\mathbf{u}$ and $\mathbf{w}$. Expanding $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$, and noticing that $\mathbf{w}^T\mathbf{u} = \mathbf{w}^T\mathbf{A}\mathbf{w} = 0$, we obtain

$$(\gamma\mathbf{u} + \mathbf{y})^T\mathbf{A}(\gamma\mathbf{u} + \mathbf{y}) + 2\gamma w = 0, \tag{9}$$

which combines a quadratic form in $\mathbf{y} \in \mathbb{R}^{n-2}$ with a term that is linear in w, thus giving a *parabolic* conic section. For instance if we're working in $\mathbb{R}^3$, in which $\mathbf{y} = y\mathbf{e}$ for some vector $\mathbf{e}$, then (9) becomes

$$w = Ay^2 + By + C.$$

Linear Transformation of a Cone

If $\mathbf{x}$ belongs to a cone defined by $\mathbf{A}$, and $\mathbf{T}$ is an invertible linear map, then $\mathbf{y} = \mathbf{Tx}$ satisfies $\mathbf{y}^T\mathbf{By} = 0$, where $\mathbf{B} = \mathbf{T}^{-T}\mathbf{AT}^{-1}$. The operator $\mathbf{B}$ is symmetric, and so from the Spectral Theorem we know there exists an orthonormal basis $(\mathbf{u}_i)$ of eigenvectors of $\mathbf{B}$. From Sylvester's Law of Inertia we know that $\mathbf{B}$ has $n-1$ positive eigenvalues and 1 negative eigenvalue, just the same as $\mathbf{A}$. It follows that the vectors $\mathbf{y} = \mathbf{Tx}$ comprise a cone. Thus cones get mapped to cones by invertible linear maps.

For a discussion and proof of Sylvester's Law of Inertia, see:

Charles F. Van Loan Gene H. Golub. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996. ISBN 0801854148

Linear Transformation of a Conic Section

We now know that a linear transformation maps a cone to a cone, and of course it maps a plane to a plane. It follows that conic sections gets mapped to a conic sections. However we can do better than this! We can show that ellipsoids get mapped to ellipsoids, hyperboloids get mapped to hyperboloids, and so on. To do this, recall that we have partitioned all conic sections into two classes, labeled 1 and 2.

Class 1 conic sections consists of vectors $\mathbf{x}$ which satisfy $\mathbf{x}^T\mathbf{Bx} = 1$, where $\mathbf{B}$ is symmetric. Within this class, the further refinement of section type depends on the inertia of the spectrum of $\mathbf{B}$ (that is, the numbers of eigenvalues of different signs). For instance, if all the eigenvalues are positive, then the section is an ellipsoid. Using machinery from the previous paragraph (i.e., Sylvester's Law of Inertia), we see that class 1 conic sections of a particular type get transformed by invertible linear maps to conic sections of exactly the same type.

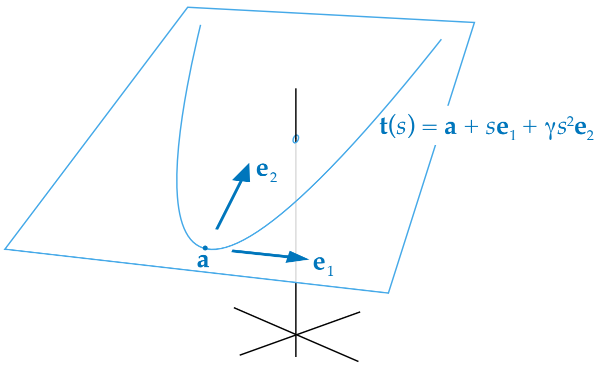
The preservation of conic sections of class 2 is an immediate consequence of the above. For if a conic section of class 2 was transformed by an invertible linear operator to class 1, then the *inverse* of this operator would transform a section of class 1 into a section of class 2, in violation of the previous paragraph.

Parabolas & Cones in $\mathbb{R}^3$

In our work on projective geometry we use the following

Lemma: If the plane of a parabolic curve in $\mathbb{R}^3$ does not contain the origin, then the curve is a conic section.

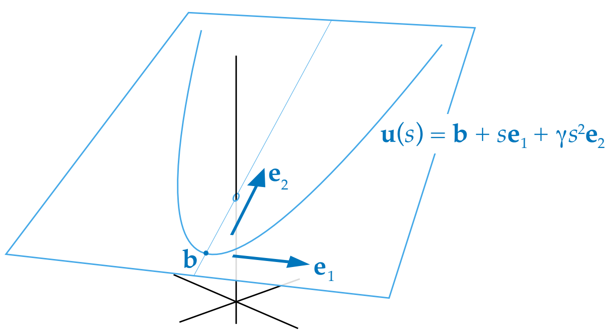
Proof: Let the parabolic curve be given by $\mathbf{t}(s) = \mathbf{a} + s\mathbf{e}_1 + \gamma s^2 \mathbf{e}_2$, where $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthonormal, and $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{a}$ are independent.



Let $\mathbf{T}$ be the linear operator on $\mathbb{R}^3$ defined by

$$\mathbf{T}\mathbf{e}_1 = \mathbf{e}_1,\ \mathbf{T}\mathbf{e}_2 = \mathbf{e}_2,\ \mathbf{T}\mathbf{a} = \frac{\mathbf{a} - \mathbf{e}_1(\mathbf{a}\cdot\mathbf{e}_1)}{||\mathbf{a} - \mathbf{e}_1(\mathbf{a}\cdot\mathbf{e}_1)||}.$$

Let $\mathbf{b}$ denote $\mathbf{T}\mathbf{a}$, and let $\mathbf{u}(s)$ denote $\mathbf{T}\mathbf{t}(s)$. The idea is that $\mathbf{T}$ shears and translates vectors so that $\mathbf{b}\cdot\mathbf{e}_1 = 0$ and $||\mathbf{b}|| = 1$. Nothing within the plane of the parabola is affected.



Our job now is to come up with a cone that has this parabola as one of its sections. That is, we must build a cone operator $\mathbf{A}$ which satisfies $\mathbf{u}(s)^T\mathbf{A}\mathbf{u}(s) = 0$ for all s. We know that $\mathbf{b}$ is a unit vector on the cone surface, and from high school geometry we know that parabolic slices of cones are parallel to the cone edge. Thus $\mathbf{e}_2$ is also a unit vector on the cone surface, and on the opposite side from $\mathbf{b}$. It must be that the cone axis is aligned with $\mathbf{b} + \mathbf{e}_2$, and so we make this one of the eigenvectors of $\mathbf{A}$. It seems like $\mathbf{e}_1$ should also be an eigenvector, and so we propose to build an $\mathbf{A}$ with eigenvectors as follows.

$$\begin{aligned}\mathbf{a}_1 &= \mathbf{e}_1\\ \mathbf{a}_2 &= \mathbf{a}_3 \times \mathbf{a}_1\\ \mathbf{a}_3 &= \mathbf{b} + \mathbf{e}_2\end{aligned}$$

Letting λ_i denote the eigenvalue corresponding to $\mathbf{a}_i$, we find that

$$\begin{aligned}\mathbf{u}^T\mathbf{A}\mathbf{u} =& \lambda_1 s^2\\ &+\lambda_2 A^2(\gamma s^2 - 1)^2 \mathbf{a}_2 \cdot \mathbf{a}_2\\ &+\lambda_3 B^2(\gamma s^2 + 1)^2 \mathbf{a}_3 \cdot \mathbf{a}_3,\end{aligned}$$

where $A(\mathbf{a}_2 \cdot \mathbf{a}_2) = (\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{b}$ and $B(\mathbf{a}_3 \cdot \mathbf{a}_3) = 1 + \mathbf{b} \cdot \mathbf{e}_2$. It follows that $\mathbf{u}^T\mathbf{A}\mathbf{u} = 0$ for all s if

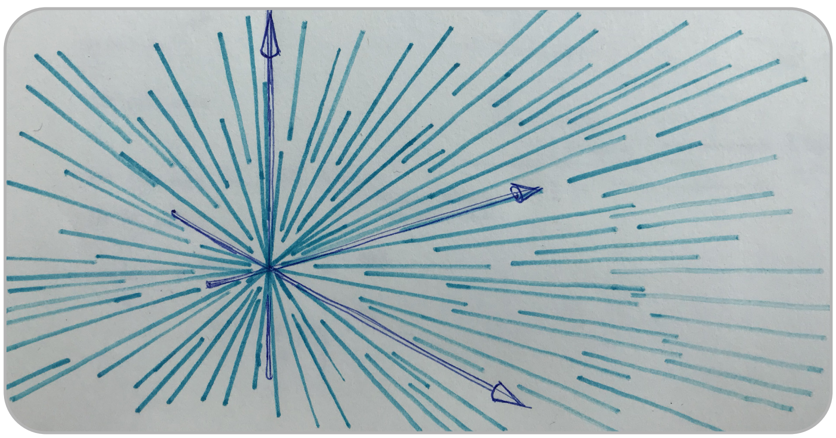
$$\begin{aligned}\lambda_3 &= -1,\\ \lambda_2 &= \frac{\mathbf{a}_2 \cdot \mathbf{a}_2}{\mathbf{a}_3 \cdot \mathbf{a}_3} \frac{(1 + \mathbf{b} \cdot \mathbf{e}_2)^2}{((\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{b})^2},\\ \lambda_1 &= \frac{4\gamma(1 + \mathbf{b} \cdot \mathbf{e}_2)^2}{\mathbf{a}_3 \cdot \mathbf{a}_3}.\end{aligned}$$

Thus we have constructed a cone that has the parabola $\mathbf{u}(s)$ as one of its sections. Then, from earlier in this chapter, it follows that the original parabola $\mathbf{t}(s)$ is also the section of a cone.

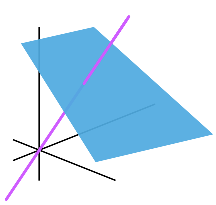
The Projective Plane

The projective plane consists of $\mathbb{R}^3$ with every line through the origin considered as a single item- as a single *point*. The projective plane is denoted by $\mathbb{P}^2$.

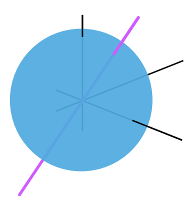
Presumably the lines through the origin of $\mathbb{R}^n$ comprise $\mathbb{P}^{n-1}$. In these notes we focus on the case $n = 3$.



The diagrams of classical projective geometry come from intersecting lines through the origin with a plane, such as the $z = 1$ plane. We refer to this as *planar projection*. Another visualization of $\mathbb{P}^2$ consists of intersecting lines through the origin with a sphere centered at the origin, each line becoming two antipodal points on the sphere's surface. We refer to this as *spherical projection*.



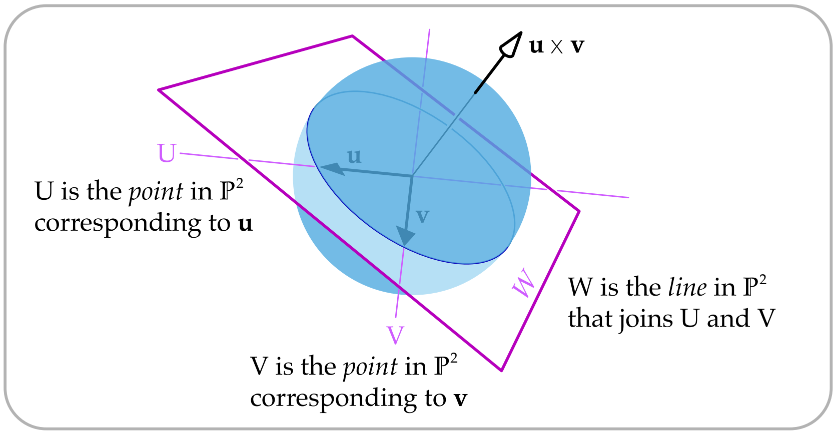
planar projection



spherical projection

Spherical projection is my personal favorite because it avoids the points at infinity that encumber planar projection. We keep in mind that spherical projection, planar projection, and other visualizations are all shadows of the same **thing**, which is $\mathbb{P}^2$.

Every nonzero vector $\mathbf{u} \in \mathbb{R}^3$ corresponds to a point in $\mathbb{P}^2$, and every nonzero multiple of $\mathbf{u}$ corresponds to this same point in $\mathbb{P}^2$. If vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ represent distinct points in $\mathbb{P}^2$, then $\mathbf{u} \times \mathbf{v}$ is perpendicular to a plane through the origin in $\mathbb{R}^3$ that contains $\mathbf{u}$ and $\mathbf{v}$. We call this plane the projective *line* joining the projective points associated with $\mathbf{u}$ and $\mathbf{v}$. In the planar visualization of $\mathbb{P}^2$, this plane projects to a straight line. In the spherical visualization of $\mathbb{P}^2$, this plane projects to a great circle.



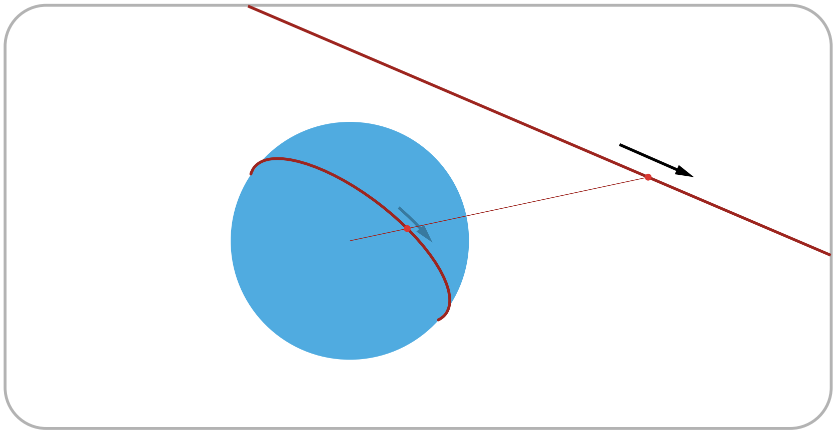
In $\mathbb{R}^3$, planes and lines passing through the origin are associated by perpendicularity. There is a bijection from the set of planes through the origin to the set of lines through the origin which assigns the perpendicular line to each plane. Planes and lines through the origin in $\mathbb{R}^3$ are projective lines and points in $\mathbb{P}^2$, and so these are associated in the same way. Thus, every projective line can be uniquely represented by a projective point, and vice versa. This *duality* between points and lines is pervasive in projective geometry. Every construction has a dual, obtained by interchanging "point" with "line", etc. For instance, distinct points are joined by a unique line, and distinct lines meet at a unique point.

Space Curves

A constant velocity space curve in $\mathbb{R}^3$ is a line, such as

$$\mathbf{x}(s) = \mathbf{a} + s\mathbf{u}.$$

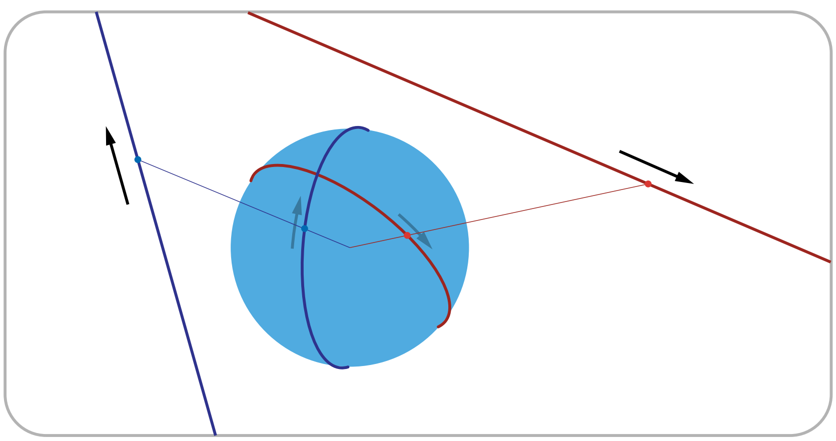
As s goes from $-\infty$ to ∞, the point $\mathbf{x}(s)$ traces out a line, and the spherical projection of $\mathbf{x}(s)$ traces out a great circle A. Let $x(s)$ denote the corresponding projective point.



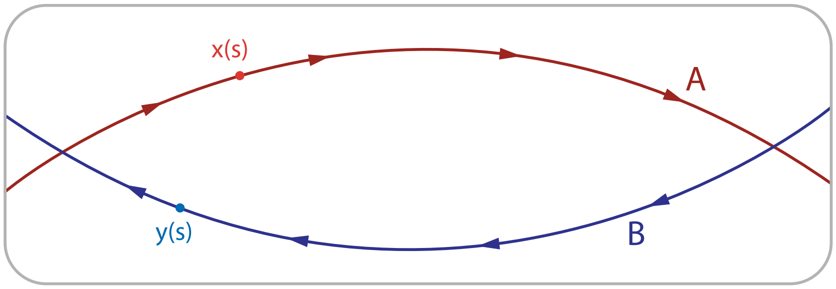
Now consider another constant velocity curve in $\mathbb{R}^3$, given by

$$\mathbf{y}(s) = \mathbf{b} + s\mathbf{v}.$$

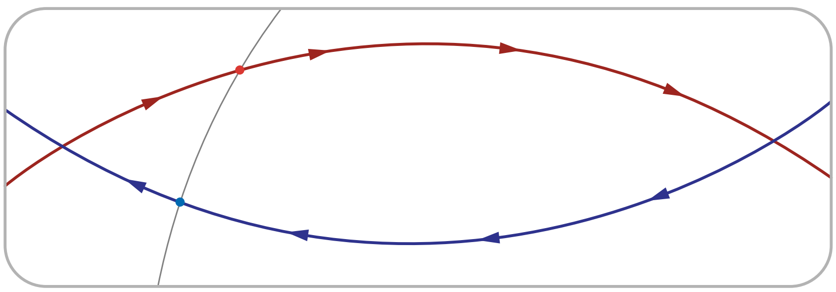
Let B denote the corresponding great circle, and $y(s)$ denote the corresponding projective point.



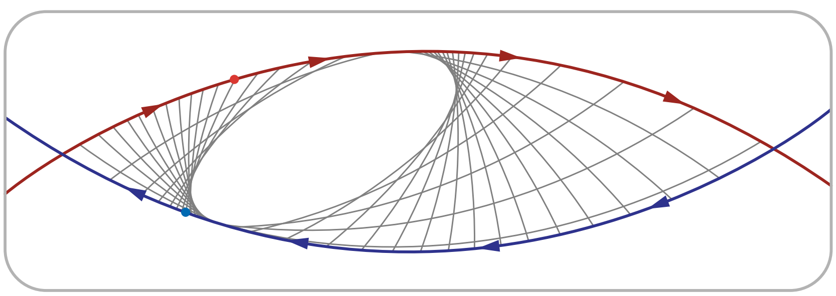
The two great circles A and B divide the sphere into four regions (or two if you want to be strict about identifying antipodal points). The boundary of each region includes a crossing point, and the boundary curves support a notion of flow- as s increases, points move along the boundary curves in a particular direction. Consider the region with boundary flow that both enters and exits the crossing point.



For each $s \in \mathbb{R}$, we have $x(s) \in A$ and $y(s) \in B$, both circling the region according to the flow shown above. For a given value of s, consider the projective line joining $x(s)$ to $y(s)$.



Now repeat the same construction for other values of s.



Our interest is in the envelope of these curves. Is this an actual ellipse? What does that even mean on the surface of a sphere?

Envelopes in $\mathbb{P}^2$

Let $\mathbf{x}(s)$ and $\mathbf{y}(s)$ be *arbitrary* parameterized curves in $\mathbb{R}^3$. Denote the corresponding curves in $\mathbb{P}^2$ by $x(s)$ and $y(s)$. For a given value of s, we join $x(s)$ and $y(s)$ by a projective line $xy(s)$.

We define the point $t(s)$ as the intersection (the meet) of the two lines $xy(s)$ and $xy(s+\epsilon)$ in the limit as ϵ goes to zero. These points $t(s)$ comprise the *envelope* curve associated with $x(s)$ and $y(s)$.

We develop an expression for $t(s)$ by working in $\mathbb{R}^3$, in which we find the intersection of $\text{span}(\mathbf{x},\mathbf{y})$ and $\text{span}(\mathbf{x}+\epsilon\mathbf{u},\mathbf{y}+\epsilon\mathbf{v})$, where

$\text{span}(\mathbf{a},\mathbf{b})$ is the plane through the origin that contains $\mathbf{a}$ and $\mathbf{b}$ (i.e., the set of all linear combinations of $\mathbf{a}$ and $\mathbf{b}$).

$$\mathbf{u} = \frac{d\mathbf{x}}{ds} \quad \text{and} \quad \mathbf{v} = \frac{d\mathbf{y}}{ds}.$$

We use cross products to construct normal vectors to each plane. A vector $\mathbf{T}$ in the intersection of the planes can be found by taking the cross product of their normal vectors.

$$\begin{aligned}\mathbf{T} &= (\mathbf{x}\times\mathbf{y})\times((\mathbf{x}+\epsilon\mathbf{u})\times(\mathbf{y}+\epsilon\mathbf{v})) \\ &= \epsilon\,(\mathbf{x}[\mathbf{x},\mathbf{y},\mathbf{v}]+\mathbf{y}[\mathbf{y},\mathbf{x},\mathbf{u}]) + \epsilon^2\,(\mathbf{y}[\mathbf{u},\mathbf{v},\mathbf{x}]-\mathbf{x}[\mathbf{u},\mathbf{v},\mathbf{y}])\end{aligned}$$

where $[\mathbf{a},\mathbf{b},\mathbf{c}]$ denotes $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$. In the limit as ϵ goes to zero, $\mathbf{T}$ becomes increasingly aligned with

$$\mathbf{t} = \mathbf{x}[\mathbf{x},\mathbf{y},\mathbf{v}]+\mathbf{y}[\mathbf{y},\mathbf{x},\mathbf{u}], \tag{10}$$

which corresponds to the desired projective point $t(s)$.

Cones

If $\mathbf{x}(s) = \mathbf{a}+s\mathbf{u}$ and $\mathbf{y}(s) = \mathbf{b}+s\mathbf{v}$, then (10) becomes

$$\mathbf{t}(s) = \mathbf{d}_0 + \mathbf{d}_1 s + \mathbf{d}_2 s^2, \tag{11}$$

where

$$\begin{aligned}\mathbf{d}_0 &= \mathbf{a}A + \mathbf{b}B, \\ \mathbf{d}_1 &= \mathbf{u}A + \mathbf{v}B \ldots \\ &\quad + \mathbf{a}U + \mathbf{b}V, \\ \mathbf{d}_2 &= \mathbf{u}U + \mathbf{v}V,\end{aligned}$$

where $A = [\mathbf{a},\mathbf{b},\mathbf{v}]$, $B = [\mathbf{b},\mathbf{a},\mathbf{u}]$, $U = [\mathbf{u},\mathbf{b},\mathbf{v}]$, and $V = [\mathbf{v},\mathbf{a},\mathbf{u}]$. The curve $\mathbf{t}(s)$ resides in a plane that is parallel to $\text{span}(\mathbf{d}_1,\mathbf{d}_2)$, and that passes through the point $\mathbf{d}_0$. We now show that $\mathbf{t}(s)$ is a parabola even when $\mathbf{d}_1$ and $\mathbf{d}_2$ are non-orthogonal.

The s at which $d\mathbf{t}/ds$ is perpendicular to $\mathbf{d}_2$ is given by

$$s_0 = -\frac{\mathbf{d}_1 \cdot \mathbf{d}_2}{2\mathbf{d}_2 \cdot \mathbf{d}_2}.$$

Let $\mathbf{u}(s) = \mathbf{t}(s + s_0)$, given by

$$\begin{aligned}\mathbf{u}(s) &= \mathbf{d}_0 + \mathbf{d}_1(s + s_0) + \mathbf{d}_2(s + s_0)^2, \\ &= \mathbf{e}_0 + \mathbf{e}_1 s + \mathbf{e}_2 s^2,\end{aligned}$$

where

$$\begin{aligned}\mathbf{e}_0 &= \mathbf{d}_0 + \mathbf{d}_1 s_0 + \mathbf{d}_2 s_0^2, \\ \mathbf{e}_1 &= \mathbf{d}_1 + \mathbf{d}_2 2 s_0, \\ \mathbf{e}_2 &= \mathbf{d}_2.\end{aligned}$$

The vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal, making it easy to see that $\mathbf{u}(s)$ traces out a parabola within the $\mathbf{e}_1$ $\mathbf{e}_2$ plane, with base located at $\mathbf{e}_0$. From the lemma in the chapter on cones, it follows that this parabola is a conic section. By projecting this cone onto different planes (e.g., the $z = 1$ plane), we get various conic sections: ellipse, parabola, and hyperbola.

A Different Foundation

Projective geometry is used by Farin[1] as a foundation for the development of NURBS, and it is through Farin's book that I first became aware of interesting topics such as Steiner conics. Ultimately however, I realized that NURBS and related topics can all be developed using standard linear algebra. Also, linear algebra is more familiar for more people than projective geometry. Thus projective geometry is interesting, but peripheral to most of the work we've done in these notes.

[1] Gerald E Farin. *NURBS*. A K Peters, second edition, 1999

References

The NURBS Book by Piegl and Tiller[2]
This is the book I should have started with. Comprehensive and well written. Accessible and complete. This book contains a very clean development of spline curves and surfaces from the ground up.

[2] Les Piegl and Wayne Tiller. *The NURBS Book*. Springer, second edition, 1997

An Introduction to NURBS by David Rogers[3]
This *extremely* introductory book makes a point of not having theorems or proofs.

[3] David F. Rogers. *An Introduction to NURBS: With Historical Perspective*. Morgan Kaufmann Publishers, 2001

NURBS by Gerald Farin[4]
Not my favorite book. Farin stimulates attentiveness by a large number of copy edits and typos. Generally difficult to follow. The book redeems itself a little by introducing important topics in the right order. For instance this was my first exposure to Steiner conics. Maybe it's not Farin's writing, and just classical projective geometry which is tedious and opaque. In my own notes I have circumvented this by using modern linear algebra to derive the results needed to understand things such as how to represent a perfect circle with a rational quadratic Bezier curve. Farin's book is like visiting a museum full of wonders, but where most of the lights are out, and where all the placards are written in a foreign language.

[4] Gerald E Farin. *NURBS*. A K Peters, second edition, 1999

Bibliography

Sheldon Axler. *Linear Algebra Done Right*. Springer, 1996. ISBN 0387982582.

Gerald E Farin. *NURBS*. A K Peters, second edition, 1999.

Charles F. Van Loan Gene H. Golub. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996. ISBN 0801854148.

Les Piegl and Wayne Tiller. *The NURBS Book*. Springer, second edition, 1997.

David F. Rogers. *An Introduction to NURBS: With Historical Perspective*. Morgan Kaufmann Publishers, 2001.