Analysis and Synthesis of Violin Arches

DAVID GOLBER

David Golber Violins, Boston, Massachusetts
dgolber@aol.com

Abstract

Modern methods provide extensive data on the shape of fine old violins. This should allow makers to copy the arches, for example, of old violins—except that a modern maker would not want to copy exactly the arch of an old instrument. The arches of old instruments are asymmetrical, and have asymmetrical outlines and bottom edges which are not flat. The modern maker wants to make a symmetrical arch, with a symmetrical outline, and flat bottom edges. So the problem is to use the data from an old asymmetrical instrument to somehow construct the arches of a new symmetrical instrument. In this article, we describe a construction method that starts from computed tomography (CT) scan data, and (1) is strongly based on data from an instrument, (2) completely defines the arch from edge to edge, (3) involves minimal human choice, and (4) allows controlled alteration of the arch. The method produces a 3D model of a violin arch. The model can be used, for example, to drive a CNC router or 3D printer, to cut cross-section templates ("quinte"), or to be part of modeling the acoustic behavior of an instrument.

OUTLINE

In this article, we describe the arch as built in three pieces: a central part, the edge, and a fill section (Fig. 1).

The edge can be designed very simply using standard computer-aided design (CAD) tools. The central part is the real concern. The fill section must smoothly bridge the gap between the central part and the edge. Designing the fill section is difficult, but has no great significance, and will not be discussed in detail.

An important concept is what we call the "low curve". The arch descends from the center toward the edge. Before it reaches the edge, it turns up slightly. Look at the lowest points (Fig. 2); the collection of all these lowest points is the "low curve". In our description, the low curve is the edge of the central part of the arch. Outside the low curve, the "fill section" begins.

Although our approach is not the only possible one, we think of the central part of the arch as built from cross-section curves sitting on top of the low curve. For readability, see Figure 3, which shows cross-section curves at 30-mm intervals. The actual spacing used was 2 mm.

Defining these cross-section curves is the real meat of the method. The method defines a family of curves with enough parameters (adjustabilities) to describe the cross-section data of a violin, but not too many to be manageable. We find these parameters for the old instrument ("analysis"), and build a new arch using these parameters ("synthesis").

THE DATA

We begin with a CT scan of a fine violin.[1] The scan data consists of X-ray densities of little cubes about half a millimeter on each side. Notice that we usually describe violin dimensions to 0.1 mm; the data we are starting with here

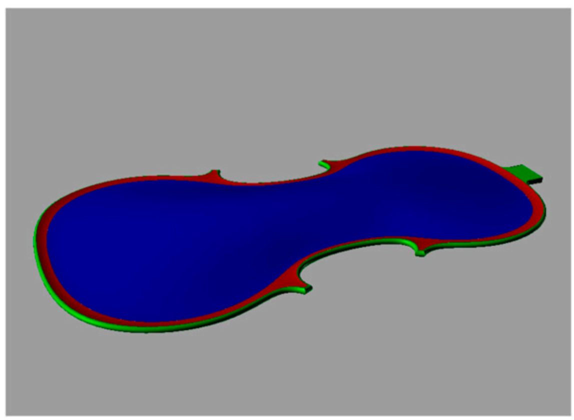

Figure 1. Edge, fill, and central parts.

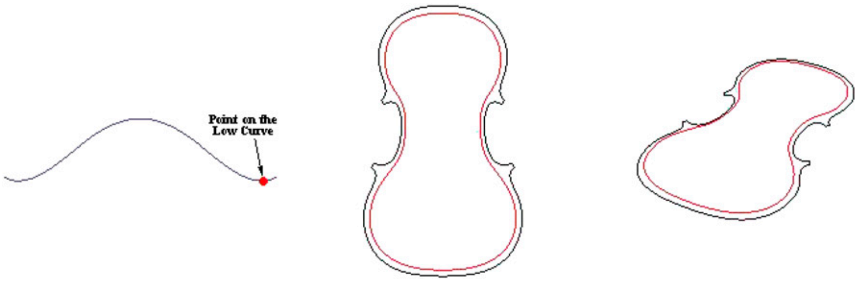

Figure 2. The "low curve".

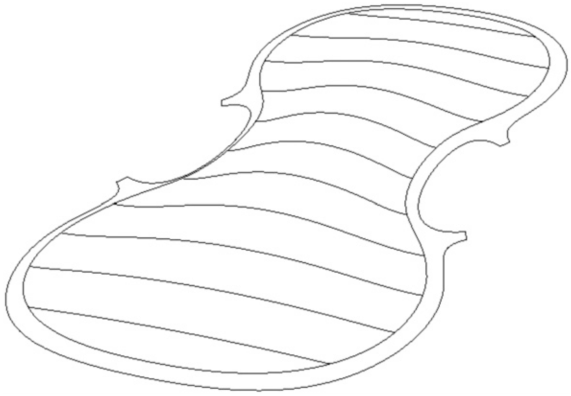

Figure 3. Cross sections sitting on the low curve.

is in some sense five times coarser! Figure 4 shows a cross section including the fingerboard. Figure 5 is a close-up of the data in the small rectangle. Each square is about 1/2 mm on a side. Figure 6 shows the values along the vertical line given in Figure 4. Adjacent points are separated by about 0.5 mm. Note the complicated behavior at the edges of the bottom, top, and fingerboard. The point here is that getting fine resolution measurements from CT data is not a simple matter.

The first step in using the CT data is to extract the surfaces of the instrument—inside and out. This is a substantial operation; it was performed for me by Biomedical Modeling Inc. of Boston, MA, USA. Biomedical Modeling chose to use the stereolithography (STL) format for the extracted surfaces.

The basic idea is that the surface is assumed to be at a certain X-ray density ("threshold"). Choosing a different threshold will cause the surface to move in or out, as shown in Figure 7.

The shape of an arch in the extracted surface will vary very little with change of threshold, because arches are fairly flat curves. For example, Figure 8 shows two curves. The upper curve is offset by 1/2 mm from the lower. But, if we pull the upper one down by 1/2 mm to make the ends coincide, then the maximum distance between the two curves is only 0.035 mm. In other words, the two curves have nearly the same shape.

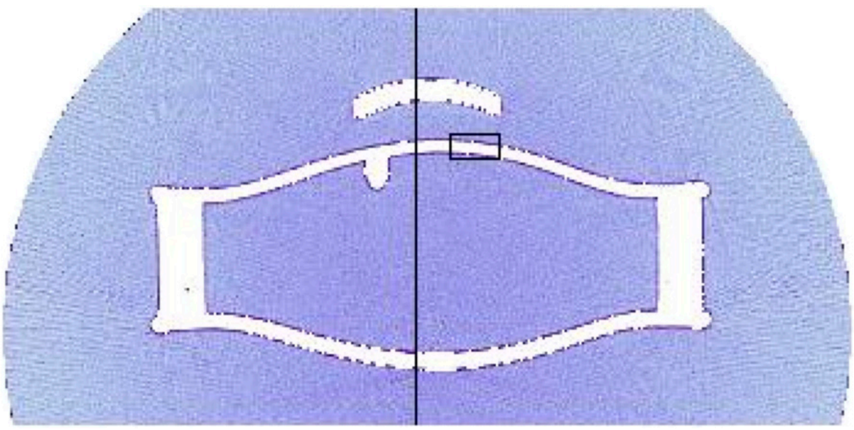

Figure 4. Cross section including the fingerboard.

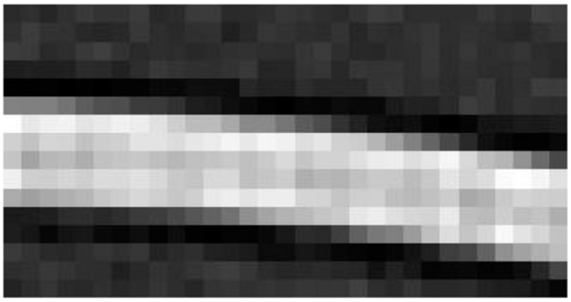

Figure 5. Close-up of the data.

On the other hand, a thickness found from the extracted surfaces will depend on the threshold, and will typically require a correction to make it match reality.

The following analysis works entirely off the STL file. In one place, a thickness (graduation) is involved, and a correction will be explicitly made. Otherwise, only shapes are involved, and no correction is made.

ANALYSIS OF CROSS SECTIONS

Figure 9 shows the outside cross section of the back, 75 mm above the tail end. The view is looking up from the tail end toward the scroll. The instrument is turned over, with its back up, so the bass side is on the right. The points are taken from the STL file at 1-mm intervals, and have been moved so that the lowest points are on the base line, equally spaced on the two sides.

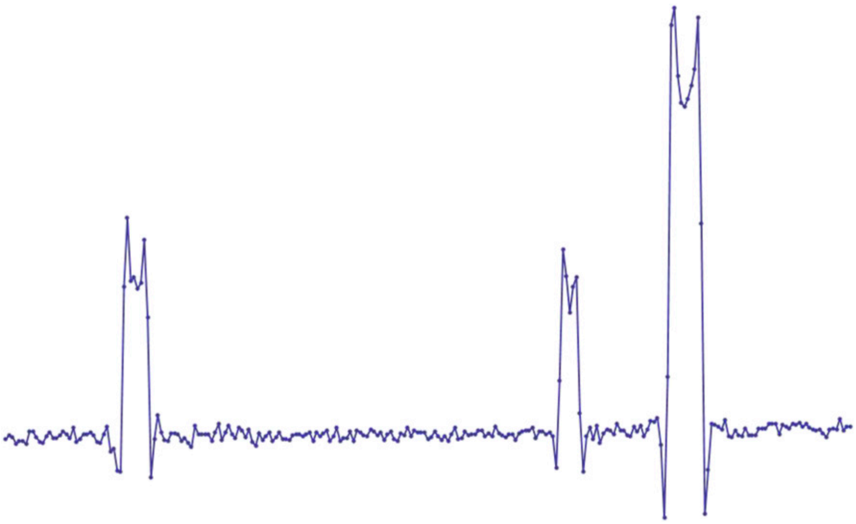

Figure 6. Values along the vertical line.

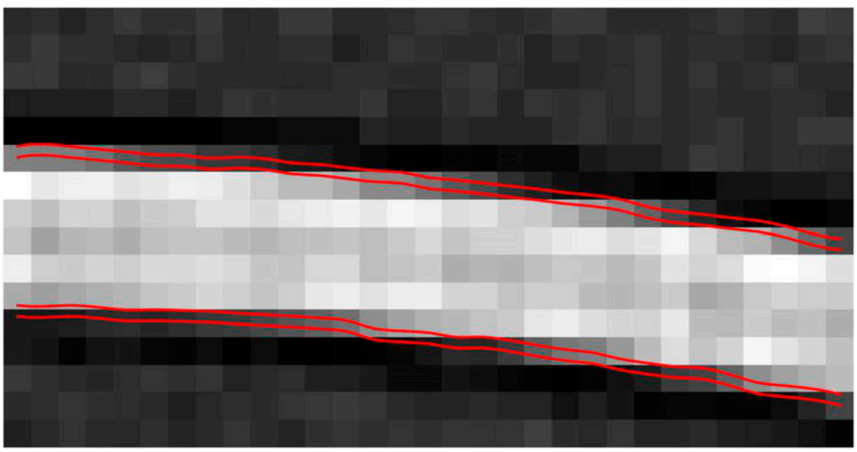

Figure 7. Different surfaces from different thresholds.

Figure 8. Offset curves.

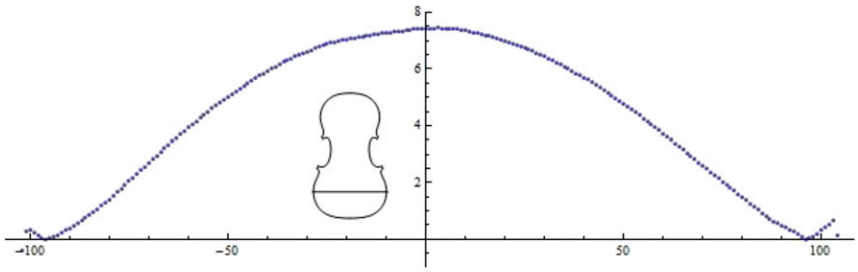

Figure 9. Seventy-five millimeter above the end.

The scales in the two directions are different so that the curve can be seen well.

In the following analysis, only the data on the bass (right) side is used, in the hope that it is less distorted by the sound post.

The Curtate Cycloid

One candidate that has been proposed to describe the cross-section curves is the curtate cycloid. Figure 10 is the result when we approximate the data (bass side only) with a curtate cycloid.

In each case, the cycloid is centered with low ends at the low ends of the data, and is adjusted to give the best possible fit to the data (the bass side data).

We see that the cycloid gives a reasonable fit to the data near the bridge and in the upper bout, but not a good fit in the lower bout.

Also, as we said in the abstract, we want a method that allows alteration of the arch. But, the cycloid is completely determined by its height and width; once the height and width are set, the curve cannot, for example, be made more or less "plump".

More Flexible Curves

This section defines a family of curves, which we call "B3 curves". A B3 curve has three parameters (adjustabilities), rather than the two of curtate cycloids, and is therefore more "flexible".

This section is the only section of this article containing actual algebra. This section needs to be read only if the reader actually wants to implement the method of the article. Otherwise, the reader can skip this section; the article is intended to be readable without this section.

First, we need to define the coordinates we are using (Fig. 11). The curves we are defining are varieties of Bèzier curves. Rather than referring to the literature, we define the curves completely here. Let

$$
\begin{aligned}
b_1(t) &= (1-t)^4 \\
b_2(t) &= 4t(1-t)^3 \\
b_3(t) &= 6t^2(1-t)^2 \\
b_4(t) &= 4t^3(1-t) \\
b_5(t) &= t^4,
\end{aligned}
$$

then, given five points (P_1, P_2, P_3, P_4, and P_5) in the xz plane and five weights (w_1, w_2, w_3, w_4, and w_5), we have the 5-point parametric Bèzier curve:

$$
\begin{aligned}
\text{Curve}(t) = {} & [w_1\, P_1\, b_1(t) + w_2\, P_2\, b_2(t) + w_3\, P_3\, b_3(t) \\
& + w_4\, P_4\, b_4(t) + w_5\, P_5\, b_5(t)] \\
& \div [w_1\, b_1(t) + w_2\, b_2(t) + w_3\, b_3(t) \\
& + w_4\, b_4(t) + w_5\, b_5(t)].
\end{aligned}
$$

The curves we are looking at are symmetrical around $x = 0$, based on the line $z = 0$, and flat at the ends, where they come down to the line $z = 0$. For such curves, we have

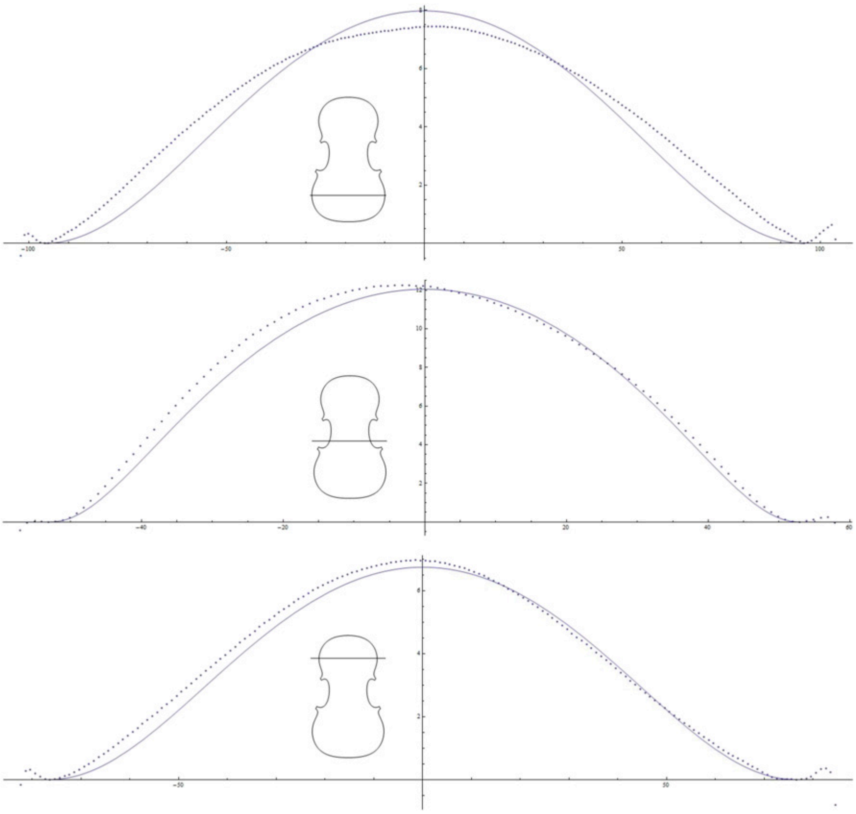

Figure 10. Approximation by a cycloid at 75, 166, and 288 mm.

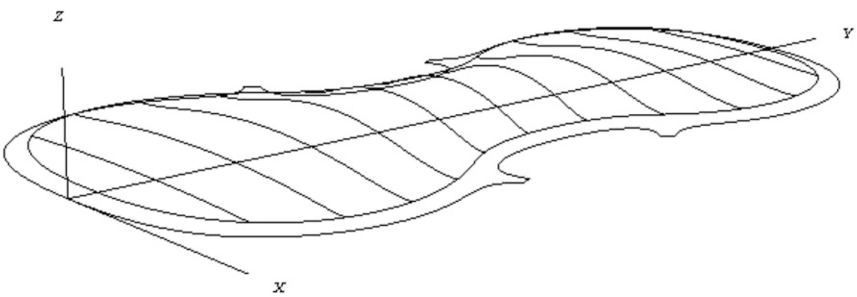

Figure 11. Definition of coordinates.

$$w_1 = w_5 = 1$$
$$w_2 = w_4$$
$$P_1 = (-\text{width},0)$$
$$P_2 = (-x_4,0)$$
$$P_3 = (0,z_3)$$
$$P_4 = (x_4,0)$$
$$P_5 = (\text{width},0).$$

So, here, we have a family of curves with the right general shape, with five parameters: width, x_4, z_3, w_3, and w_4. After considerable trials, I realized that this was too many parameters. It turned out that taking $w_3 = 1$ and $w_4 = 5/4$ reduced the number of parameters to three and still left enough flexibility to give a good representation of the cross-section data. This choice of w_3 and w_4 also makes z_3 = three times the height of the curve, which is very convenient. The x_4 parameter in some sense controls the bulginess of the curve, so we define bulge = x_4/width. Legitimate values of bulge are between 0 and 1.

Thus, we have a family of curves, which I call B3 curves. For parameters height, width, and bulge, we set $z_3 = 3 \times$ height, x_4 = bulge × width, and define P_1, P_2, P_3, P_4, and P_5 and $w_1 = 1$, $w_2 = 5/4$, $w_3 = 1$, $w_4 = 5/4$, and $w_5 = 1$ according to the previous formulas. Then, the five-point Bèzier defined previously is the curve.

Using the B3 Curves

The previous section defines a family of "B3 curves". For a given height, width, and a third number called "bulge", we get a curve. Figure 12 shows how these curves look for various values of bulge. We compare the fit of the B3 curves and the cycloid in Figure 13

In each case, the width of the B3 curve is fixed, and the height and bulge are adjusted to give the best possible fit to the data. Similarly, the width of the cycloid is fixed and the height is adjusted. The best-fit B3 and cycloid curves are shown. We see that the B3 curve fits the data better than the cycloid. So in the rest of this work, we use only B3 curves.

ANALYSIS

We now begin the actual "analysis" of the data: For cross sections from 30 to 324 mm, measured

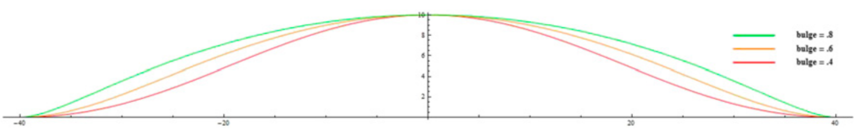

Figure 12. Various values of "bulge."

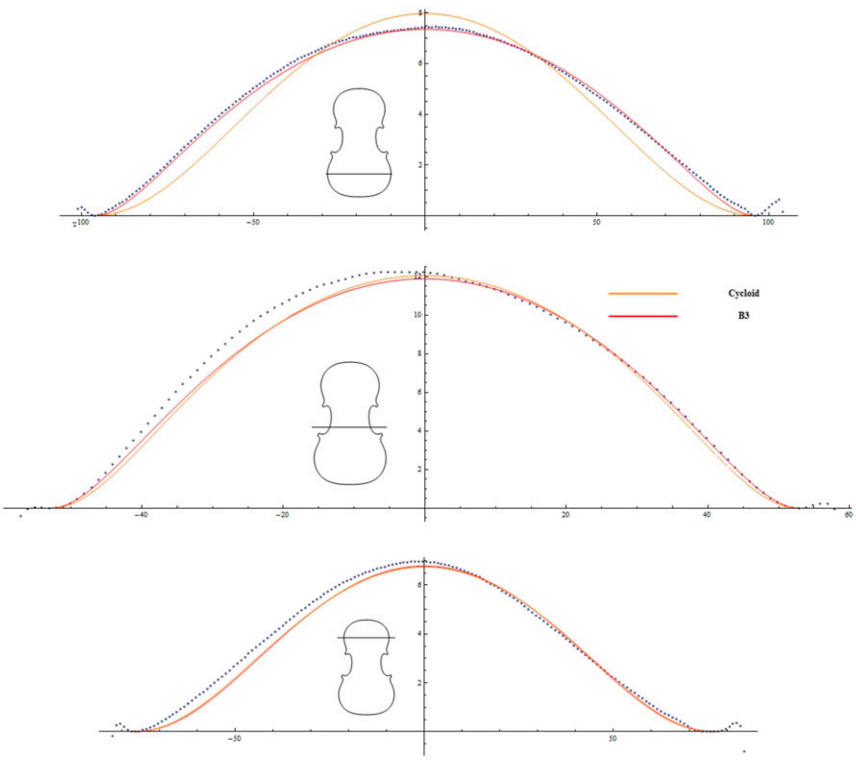

Figure 13. B3 and cycloid curves.

from the tail of the instrument at intervals of 2 mm, we fit a B3 curve to the bass side cross-section data. (For about a dozen cross sections, the edge of the instrument is worn enough that there is no "lowest point" at the edge, and so, it is not possible to perform the analysis in the way we are doing it. So, there are gaps in the data for these cross sections.) For each cross section, we have a width, a height, and a bulge (Fig. 14).

Figure 15 shows the height, width, and bulge of the cross sections (shown against the outline of the instrument). Note that the low curve (width of the cross sections) pushes slightly into the corners.

The cross-section curves sit on the "low curve". So, we need data about the low curve. The widths of the cross-section curves, shown previously, give the location of the low curve, as seen from above. The height of the low curve is the thickness of the plate under these width points. In Figure 16 we are measuring a thickness from the STL file, so a thickness correction is applied.

The red points are the thicknesses read from the STL file. Subtracting a correction[2] of 0.6 mm gives the green points; these are the heights of the low curve. The outline of the instrument is shown for comparison. The figure shows that the thickness at the low curve is about 1 mm greater in the C bout than in the outer bouts.

Getting the B3 parameters and the low curve heights constitutes the "analysis" of the old violin.

SYNTHESIS

Smoothing

For each cross section, from 30 to 324 mm above the end of the instrument, at 2-mm intervals (with some gaps), we have a height, width, and bulge. These define a B3 curve. And, we have the low curve that the B3 curves sit on. So

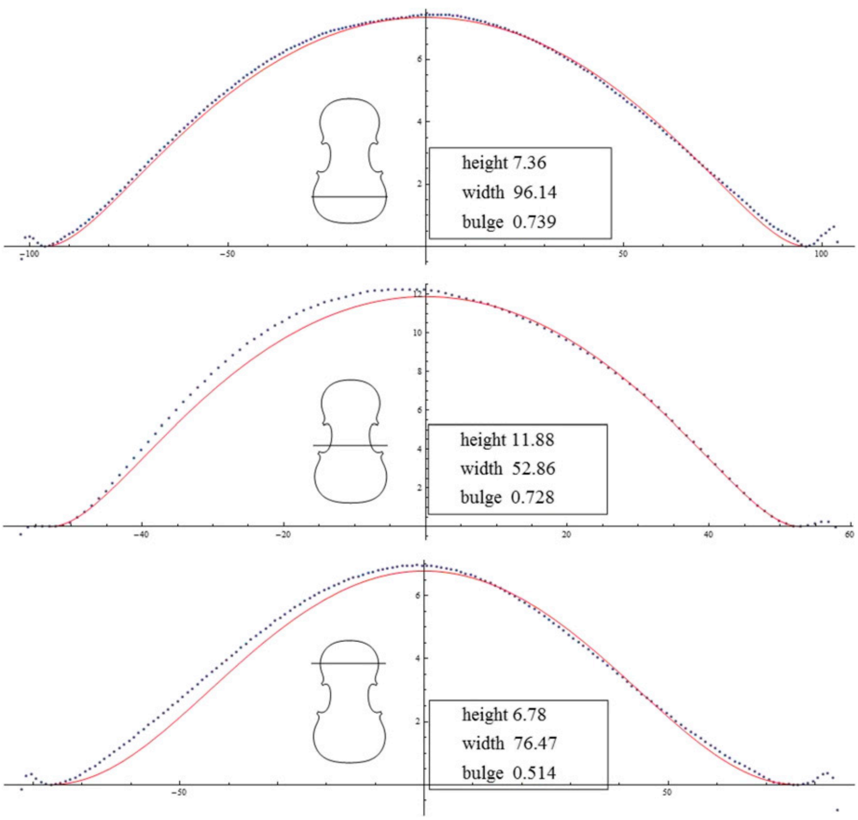

Figure 14. B3 Curves fitted to cross-section data.

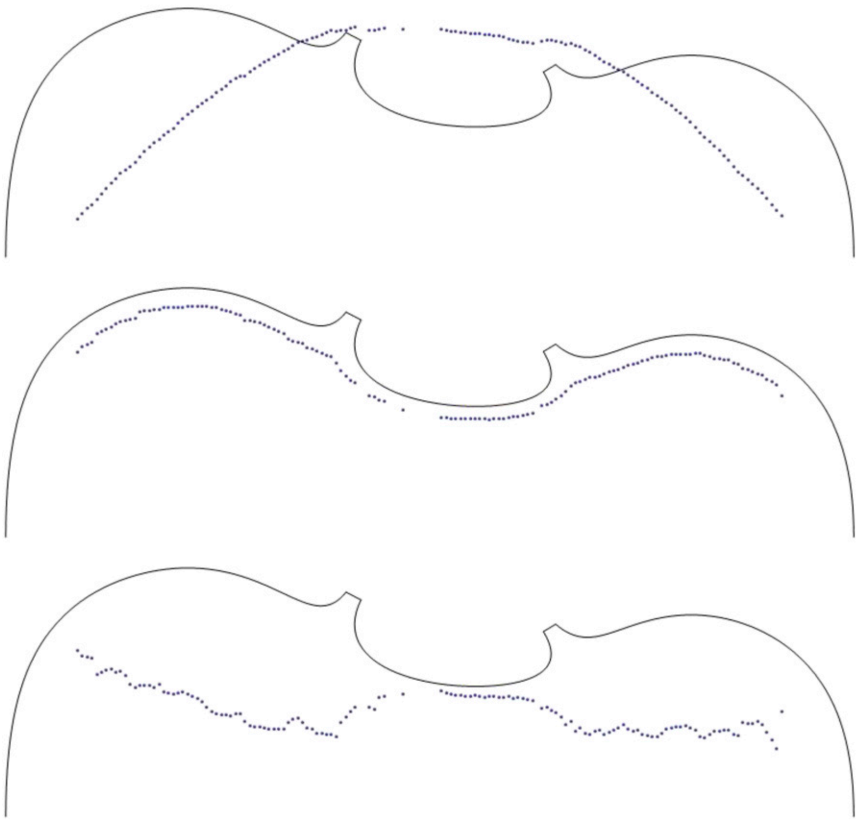

Figure 15. Height, width, and bulge of the cross sections.

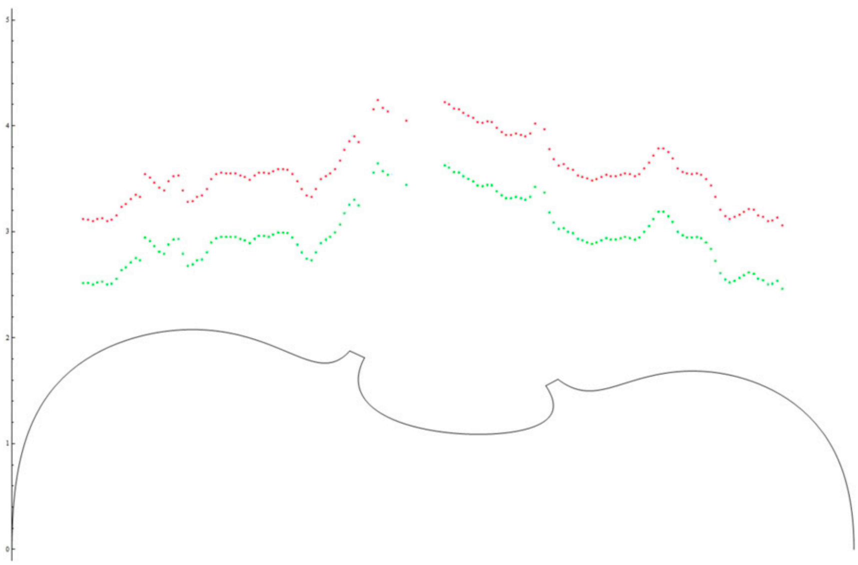

Figure 16. Low curve heights with correction.

we could just put all these B3 curves in place to form an arch. But, if we do this, we get a jagged surface. The center line appears as shown in Figure 17.

The problem is that each cross section, and each low curve height, was found independently, with noise and imprecision from the CT data, the STL conversion, and so on. They do not fit together into a smooth surface. So, we have to smooth things out.

Figure 18 shows smoothed versions of the B3 height, width, and bulge, and the height of the low curve. The outline of the instrument has been included for orientation.

Notice that the width data projects slightly into the corners, whereas the smooth width curve does not. I can produce a smooth width curve that does project into the corners like the data, but if I do, these methods, including various modifications and extensions, produce arches with unacceptable lumpiness.

Building the Arch

We now have the following:

- smoothed low curve
- smoothed B3 width
- smoothed B3 height
- smoothed B3 bulge

We can now build an arch. From the smoothed B3 height, width and bulge, we build cross sections from 10 to 342 mm (above the tail of the instrument) at intervals of 2 mm. We put these on the smoothed low curve. This forms the central part of the arch.

We look at the lengthwise sections as a check. The section is 45 mm from the center line as shown in Figure 19.

The sections, from the center line all the way out to the edge, are as expected, without any unacceptable lumpiness.

The edge and fill parts are added to the central part to form the complete arch of the plate.

POSSIBLE DIFFICULTIES

Slight variations in the smoothing of the curves sometimes lead to a surface with an unacceptable kink. One can check for a kink by looking at the lengthwise sections. Figure 20 shows an example of a kink. This is a lengthwise section 45 mm from the center line, with points every 2 mm. The remedy for this problem is to fit a smooth curve to the section, as demonstrated in Figure 21. The bulge is then modified to force the cross sections to pass through this smoothed curve.

ALTERING AN ARCH

This method allows for altering an arch. Suppose we have an arch built with these methods and we want it to be slightly plumper in the upper bout, about 288 mm above the tail of the instrument. We increase the bulge at 288, enough to add about 1/2 mm there, and taper the alteration off to zero (Fig. 22).

Using the modified bulge, with the other pieces unchanged, we get a new arch. Figure 23 is a comparison between the altered and the original arch.

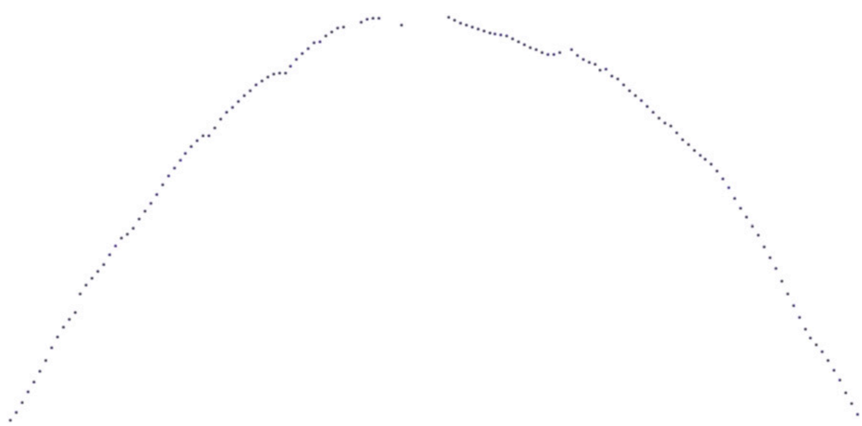

Figure 17. Jagged center line from un-smoothed data.

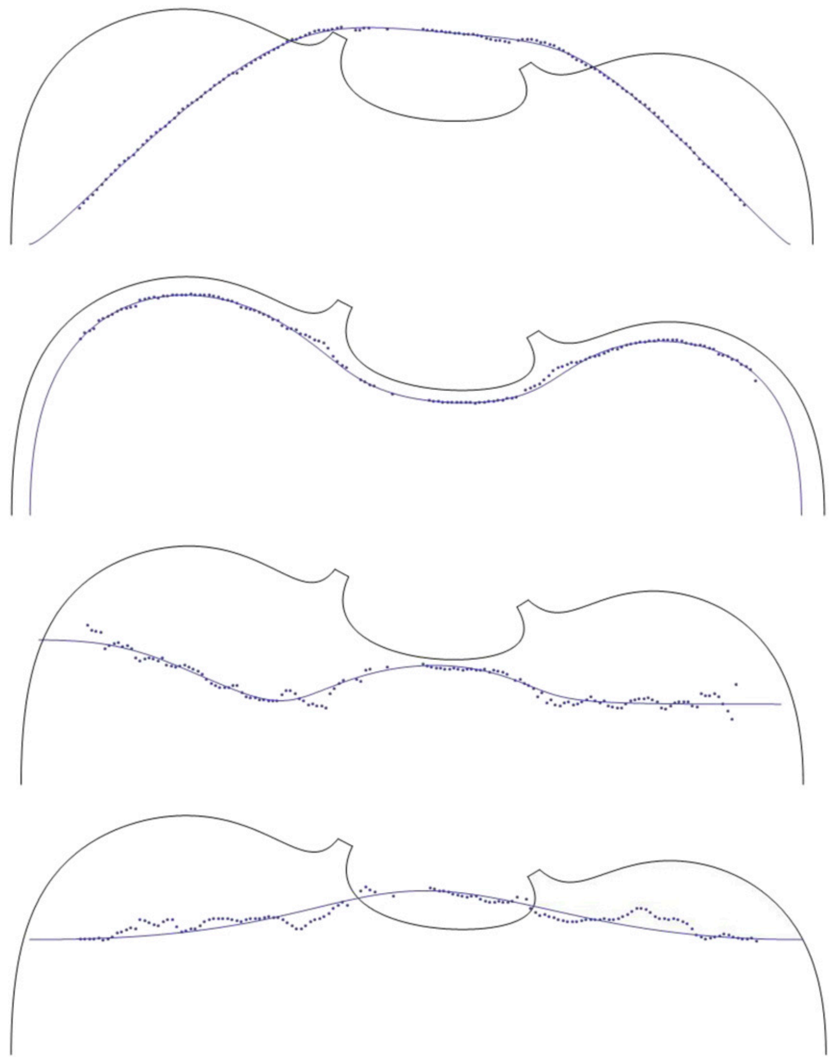

Figure 18. Smooth versions of the B3 height, width, and bulge, and the low curve height.

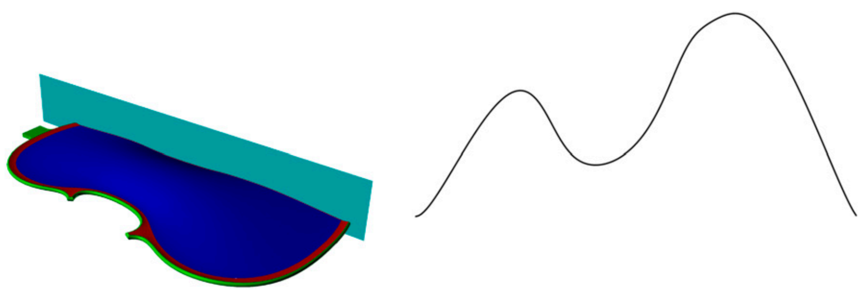

Figure 19. Section 45 mm from the center line.

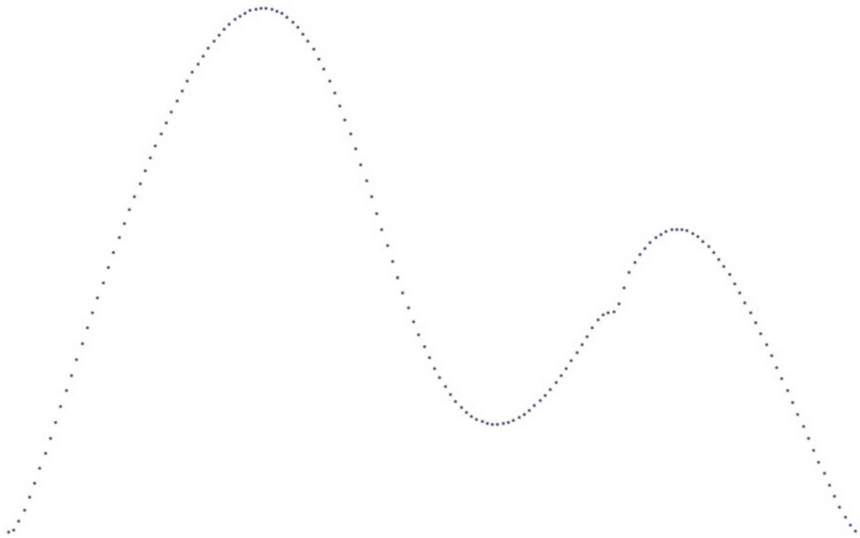

Figure 20. Section 45 mm from the center line showing kink.

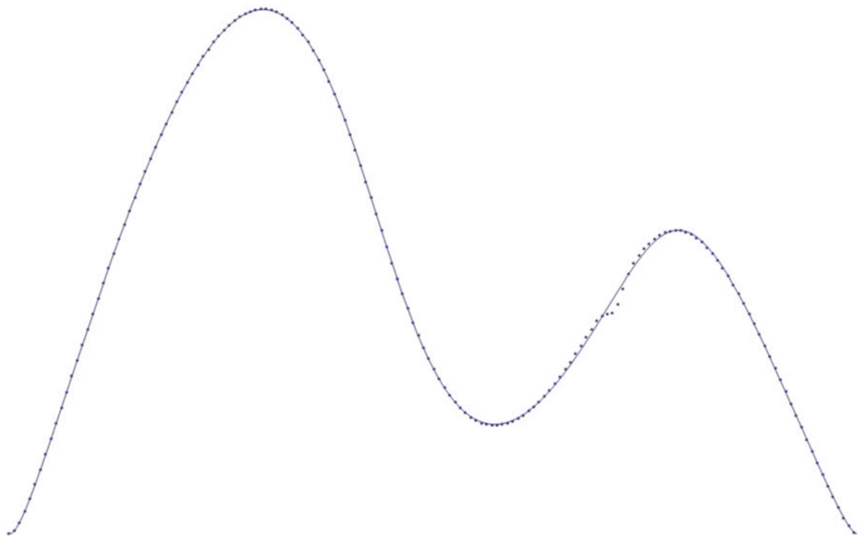

Figure 21. Smooth curve fitted to the kinked section.

ADAPTING TO A DIFFERENT OUTLINE

The previous procedure used the low curve from the original data (cleaned up and made smooth and symmetric). This goes along with using the outline from the original data (also cleaned up and made smooth and symmetric).

If one wants to use the arch data from the old instrument, but on a new instrument with a substantially different outline, the 2D low curve needs to be adapted to the new outline. Some examples of recipes that might be used to construct an appropriate 2D curve are as follows:

- The low curve lies over the inner edge of the linings;
- The low curve lies 6 mm from the edge;
- The low curve is 6 mm from the edge in the lower bout, 4 mm from the edge in the C bouts, and 8 mm from the edge in the upper bout.

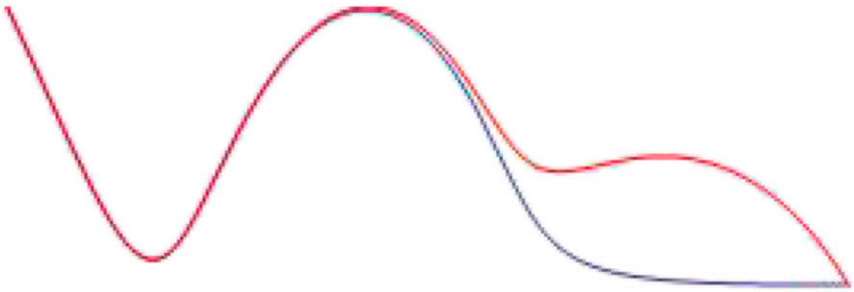

Figure 22. The altered bulge.

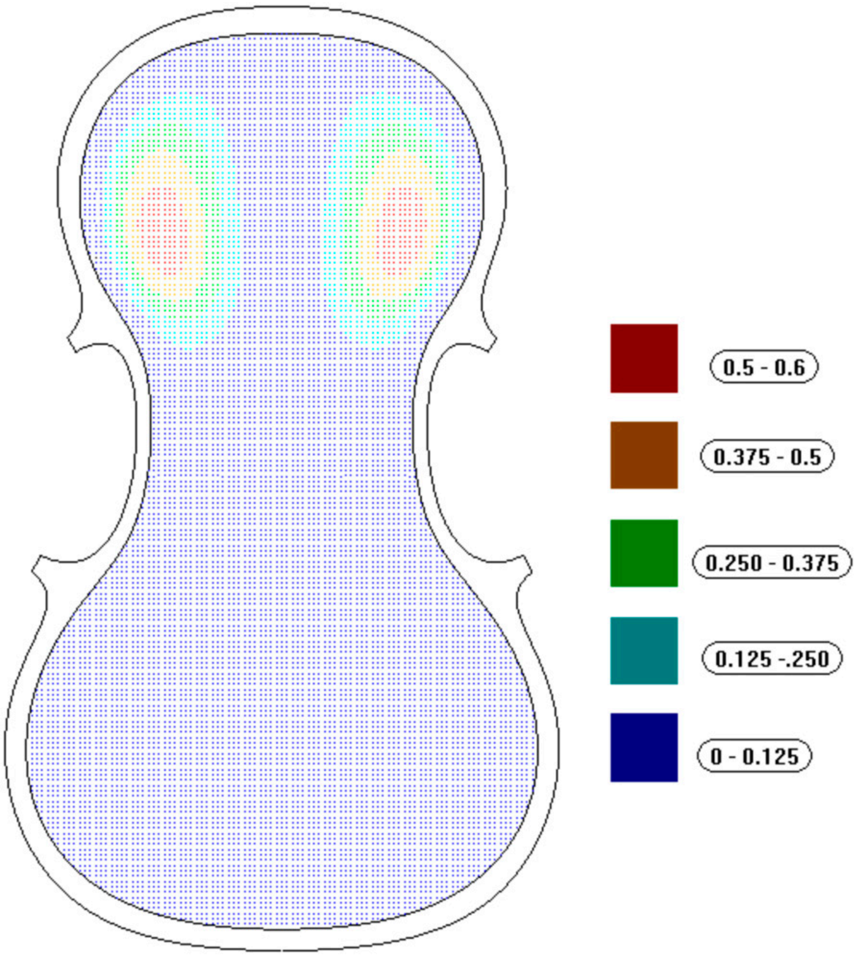

Figure 23. Differences between the altered arch and the first arch.

TOOLS AND TECHNIQUES

The major tools used were Mathematica and Rhino. Mathematica is a sophisticated and powerful tool for performing mathematical calculations, both symbolic and numerical. Unfortunately, it is difficult to use, partly because it is poorly documented. Rhino is a 3D CAD tool. An add-on, RhinoCAM, generates instructions for a computer numerical control (CNC) router.

I wrote functions in Mathematica for reading CT data (DICOM files) and STL files, and developed methods for transferring data back and forth between Mathematica and Rhino.

The calculations were performed in Mathematica. After the cross-section curves and low curve were found in Mathematica, they were transferred to Rhino, where the actual surface was produced. The edge was produced completely in Rhino. The basic idea of the fill surface is to invent a function something like a temperature function, moving from one temperature (height) at the low curve to a higher temperature (height) at the edge.

All curves are Bèzier curves of appropriate degree. A curve is "fitted" to data points by adjusting the Bèzier parameters (control points and weights) to minimize the sum of the squares of the distances from the data points to the curves. Specifically,

- The edge curves (upper, C, and lower bouts) are order 5 Bèzier curves (5 control points). The curves are fitted to data points consisting of the STL points where the normal (perpendicular to the surface) is close to the horizontal.
- The smoothed B3 width is an order 9 Bèzier curve (refer to Fig. 18). The data points in the corners were omitted, and the points in the C bout were emphasized by weighting them heavily.
- The smoothed low curve height is an order 5 Bèzier curve. It was not fitted to the data; the curve parameters were adjusted by hand ("eyeballed") instead.
- There is a tricky point about adjusting the B3 height: The curve we care about is not the B3 height itself, but the actual center line height of the arch, which is the sum of the B3 height and the low curve height, because the cross-section curves sit on the low curve. So, the process is as follows:
  - Add the B3 height data (the points in the first part of Fig. 18) to the height of the smoothed low curve (just found).
  - Fit an order 6 Bèzier curve to the resulting heights.
  - Subtract the height of the smoothed low curve, giving the smoothed B3 height.
- The smoothed bulge curve is an order 8 Bèzier curve. Some of the data points are deleted, to prevent the fitting process from producing a curve with very sharp turns.
- The smoothed curve fitted to the kinked section is an order 7 Bèzier curve.

If you are having difficulty implementing this method, you are invited to contact the author.

NOTES

1. Because these data have not been publicly released, the violin will not be identified.

2. This correction is based on an analysis of a CT scan of a sample violin (not the fine instrument which is the subject of the analysis) with marked locations of known thickness.