Appendix C: The Intersection of Constraint Surfaces

In the proof of Theorem 1 in Section 3 (a) we rely on our ability, at least in theory, to eliminate x_1, x_2, x_3 from the three constraint equations shown in (3.3).

$$\begin{aligned} a(x_1, x_2, x_3, x_4, x_5, x_6) &= 0 \\ b(x_1, x_2, x_3, x_4, x_5, x_6) &= 0 \\ c(x_1, x_2, x_3, x_4, x_5, x_6) &= 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x_1 &= X^1(x_4, x_5, x_6) \\ x_2 &= X^2(x_4, x_5, x_6) \\ x_3 &= X^3(x_4, x_5, x_6) \ . \end{aligned} \qquad (3.3)$$

How does one know this is possible? The equations might be very complicated, some of the coordinates might not appear in some of the equations, and so on.

We look first at three examples in E^3 then in Example 4 we return to our question.

Example 1: Consider these two constraint equations in E^3 :

$$\begin{aligned} a(x_1,x_2,x_3) &= x_1^2 + x_2^2 + x_3^2 - 2^2 = 0 \\ b(x_1,x_2,x_3) &= (x_1-2)^2 + x_2^2 + x_3^2 - 2^2 = 0 \ . \end{aligned} \qquad (C.1)$$

Each constraint represents a sphere of (radius 2) and so is a 2-dimensional smooth surface in E^3. The first sphere is centered at the origin, while the second has its center at (2,0,0). Subtracting the first equation from the second gives $-4x_1+4 = 0$ so $x_1 = 1$, then inserting this into both equations gives

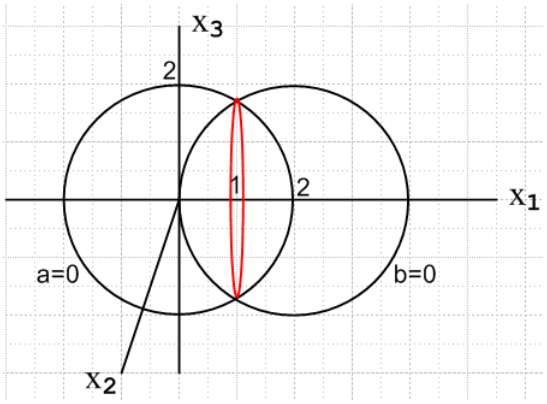
$$\begin{aligned} 1 + x_2^2 + x_3^2 - 2^2 &= 0 \\ 1 + x_2^2 + x_3^2 - 2^2 &= 0 \end{aligned} \quad \Rightarrow \ x_2^2 + x_3^2 = 3 \ . \qquad (C.2)$$

The intersection of the two constraint surfaces is a circle, a 1-dimensional smooth surface in E^3.

Given $a(x_1,x_2,x_3) = 0$ and $b(x_1,x_2,x_3) = 0$, is it possible to eliminate x_1 and x_2 from the two constraint equations? The answer is yes, as follows,

$$\begin{aligned} x_1 &= X^1(x_3) = 1 \\ x_2 &= X^2(x_3) = \pm\sqrt{3 - x_3^2} \ . \end{aligned} \qquad (C.3)$$

The following schematic drawing shows the intersection surface in red,



(C.4)

Notice in this example that there are two possible solutions for x_2. One can regard the intersection surface (the circle $x_2^2 + x_3^2 = 3$) as having two pieces (front half, back half) and the ± sign selects one of these pieces.

It is easy to imagine more complicated examples for two constraints in E^3 where perhaps one or both of the constraint surfaces contain disjoint pieces (e.g. a hyperboloid), and where the intersection surface also has several disjoint pieces.

Example 2: Consider these two constraint equations in E^3 :

$$\begin{aligned} a(x_1,x_2,x_3) &= x_1^2 + x_2^2 + x_3^2 - 2^2 = 0 \\ b(x_1,x_2,x_3) &= \quad x_2^2 + x_3^2 - 1^2 = 0 \,. \end{aligned} \tag{C.5}$$

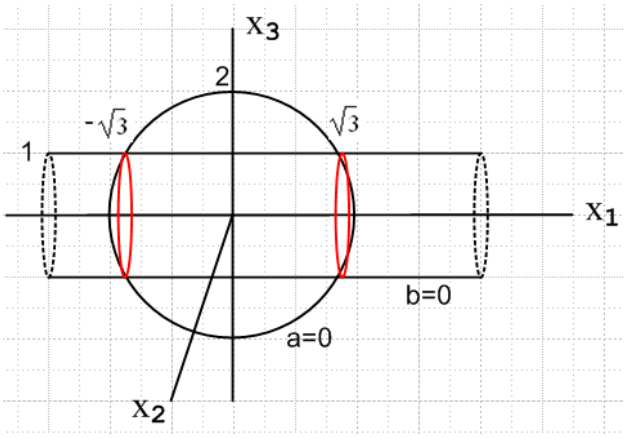
The second equation is missing the x_1 coordinate. The first equation describes the same origin-centered radius 2 sphere of the previous example. The second equation is that of a circle of radius 1, but as a function of three variables it is in fact a cylinder whose axis is the x_1 axis. If x_2 and x_3 lie on the circle, any value of x_1 satisfies the second equation. Missing coordinates result in surfaces which are "extruded" in the dimensions of the missing coordinates. Subtracting the second equation from the first gives $x_1^2 - 3 = 0$ so $x_1 = \pm\sqrt{3}$, then inserting this into both equations gives

$$\begin{aligned} 3 + x_2^2 + x_3^2 - 2^2 &= 0 \\ x_2^2 + x_3^2 - 1^2 &= 0 \quad \Rightarrow \; x_2^2 + x_3^2 = 1 \,. \end{aligned} \tag{C.6}$$

Given $a(x_1,x_2,x_3) = 0$ and $b(x_1,x_2,x_3) = 0$, is it possible to eliminate x_1 and x_2 from the two constraint equations? The answer is yes, as follows,

$$\begin{aligned} x_1 &= X^1(x_3) = \pm\sqrt{3} \\ x_2 &= X^2(x_3) = \pm\sqrt{1 - x_3^2} \,. \end{aligned} \tag{C.7}$$

The following schematic drawing shows the intersection surface in red,



(C.8)

In this example there are two possible solutions for x_1 and two for x_2 so overall there are four solutions. The intersection surface consists of the two circles each having a front half and a back half.

Example 3: Consider these two constraint equations in E^3 :

$$\begin{aligned} a(x_1,x_2,x_3) &= x_1^2 + x_2^2 + x_3^2 - 2^2 = 0 \\ b(x_1,x_2,x_3) &= x_1^2 + x_2^2 - 1^2 = 0 \,. \end{aligned} \tag{C.9}$$

Now x_3 is missing from the second equation instead of x_1. Subtracting the second equation from the first gives $x_3^2 - 3 = 0$ so $x_3 = \pm\sqrt{3}$, then inserting this into both equations gives

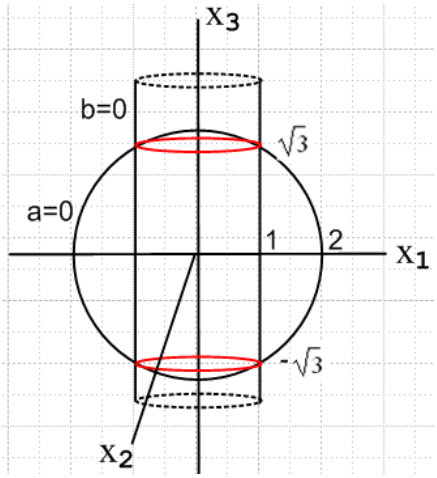
$$\begin{aligned} & x_1^2 + x_2^2 + 3 - 2^2 = 0 \\ & x_1^2 + x_2^2 - 1^2 = 0 \quad \Rightarrow \; x_1^2 + x_2^2 = 1 \,. \end{aligned} \tag{C.10}$$

Given $a(x_1,x_2,x_3) = 0$ and $b(x_1,x_2,x_3) = 0$, is it possible to eliminate x_1 and x_2 from the two constraint equations? The answer is yes, as follows, where α is an arbitrary real parameter in [-1,1],

$$\begin{aligned} & x_1 = X^1(x_3 = \pm\sqrt{3}) = \alpha \qquad -1 \le \alpha \le 1 \\ & x_2 = X^2(x_3 = \pm\sqrt{3}) = \pm\sqrt{1 - \alpha^2} \,. \quad \text{// the two } \pm \text{ signs are independent} \end{aligned} \tag{C.11}$$

In the previous two examples the argument x_3 was a free parameter whose variation mapped out the smooth constraint intersection surface piece(s). In Example 3 x_3 is fixed at $\pm\sqrt{3}$ and a new parameter α must be introduced to map out the intersection surfaces. There are again four solution half-circles.

The following schematic drawing shows the intersection surface in red,



(C.12)

Example 4: Now consider these three constraint equations in E^6 :

$$a(x_1, x_2, x_3, x_4, x_5, x_6) = 0$$
$$b(x_1, x_2, x_3, x_4, x_5, x_6) = 0$$
$$c(x_1, x_2, x_3, x_4, x_5, x_6) = 0 \ . \qquad (3.2) \qquad \text{(C.13)}$$

We assume that each equation describes a smooth surface of dimension 5 in E^6 (each is a hypersurface) and each surface may have several pieces. In a problem with constraints, one *assumes* that there is some non-null surface of intersection of all the constraint surfaces. This intersection surface may have multiple pieces and each piece (barring pathological cases) is itself a smooth surface of dimension 3 in E^6. A candidate solution point **r** must lie on one of these intersection surface pieces. In general, each constraint knocks down the dimension of the intersection surface by one degree of freedom.

So let us assume that x_4, x_5, x_6 are the last three components of some point(s) on the overall intersection surface. Each such point of course has some x_1, x_2, x_3 components. If x_4, x_5, x_6 are varied slightly, x_1, x_2, x_3 will also vary slightly, and this is what we mean by writing the functions

$$x_1 = X^1(x_4, x_5, x_6)$$
$$x_2 = X^2(x_4, x_5, x_6)$$
$$x_3 = X^3(x_4, x_5, x_6) \ . \qquad (3.3) \qquad \text{(C.14)}$$

If the intersection surface has multiple pieces which have the same x_4, x_5, x_6 value, then any of the three functions X^n above may be multi-valued, as occurred in our earlier examples. Conversely, if we select a triplet x_4, x_5, x_6 for which there are no points on the intersection surface, the solution x_1, x_2, x_3 does not exist. We don't care about such points in E^6.

So this then is what we mean in Section 3 (a) when we say above (3.3) that one can use the three constraint equations to eliminate the three variables x_1, x_2, x_3. In the concluding equations (3.11) one can regard the functions X^1, X^2, X^3 as being any of the multi-valued functions just discussed if the intersection surface has multiple pieces. For example, we had in (3.11) that

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} f_1 & f_2 & f_3 & f_4 \\ a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{pmatrix} \begin{pmatrix} X^1{}_4 \\ X^2{}_4 \\ X^3{}_4 \\ 1 \end{pmatrix} \qquad \textit{part of (3.11)} \qquad \text{(C.15)}$$

and here $X^1{}_4$ refers to $(\partial/\partial x_4)X^1(x_4, x_5, x_6)$ evaluated at coordinates x_4, x_5, x_6 for some candidate solution point **r** on the intersection surface. The fact that the matrix in (C.15) must have zero determinant is not affected by the possible existence of multiple solution X^1, X^2, X^3 functions.

Notice that the derivatives appearing in the matrix are unaffected by the possibility of multi-valued functions for X^1, X^2, X^3.

In the same manner as above, we can "eliminate" any triplet of variables x_i, x_j, x_k ($i \neq j \neq k$) from the three constraint equations (C.13).

In the general case where there are S-1 constraints and **r** has N components $x_1, x_2 ... x_N$ with N > S, the dimensionality of the intersection surface is N - (S-1) in E^N. The conclusions of the Theorem 1 proof outlined in Section 3 (b) are similarly not affected by the possible existence of an intersection surface having multiple pieces and the functions X^n possibly having multiple values.

Footnote: Why is $a(x_1, x_2, x_3, ...x_N) = 0$ an (N-1)-dimensional surface in E^N ? (C.16)

This fact probably seems obvious to the reader and is certainly obvious in the first three examples above. Here we attempt a simple "engineering" explanation.

We *assume* that a constraint equation is a "smooth" equation, meaning it is continuous and differentiable in all its arguments except perhaps at certain isolated locations which can be handled on an *ad hoc* basis either by fiat or by taking limits. Consider that, since a(**r**) = constant (namely 0), da = 0 so

$$0 = da = \Sigma_{i=1}^{N} a_i dx_i \ , \qquad // \ a_i = a_i(\mathbf{r}) = a_i(x_1, x_2, x_3, ...x_N) \qquad \text{(C.17)}$$

where a_i as usual means $\partial a/\partial x_i$. Let **r** be some point for which a(**r**) = 0. At this point we shall assume that there exists at least one $a_s(\mathbf{r})$ which is non-zero. If *all* $a_i(\mathbf{r}) = 0$, we show below that **r** is a point on a surface which has a null normal vector, which is not possible, so such an **r** value could not lie on a(**r**) = 0. Then (C.17) may be written

$$dx_s = (1/a_s) \Sigma_{i \neq s} a_i dx_i \ . \qquad \text{(C.18)}$$

Create a coordinate system whose origin is located at this point **r** for which a(**r**) = 0, and whose axes are aligned with those of E^N. Now imagine an arbitrary tiny displacement of all the dx_i other than dx_s. For this set of N-1 dx_i there is only one possible dx_s which causes the point **r** + d**r** to satisfy the equation a(**r**+d**r**) = a(**r**) + da = 0 + 0 = 0, where d**r** ≡ $(dx_1, dx_2 .....dx_N)$. That one possible dx_s value is that given by (C.18). Imagine repeating this process for a continuum of values for the dx_i other than dx_s, and in each case we obtain the unique solution dx_s from (C.18). We can think of the x_s axis as being "vertical" and all the other axes being "horizontal" inasmuch as they are all perpendicular to the x_s axis. The vectors d**r**

generated in this manner comprise a tiny patch of a "plane" of dimension N-1 (a hyperplane) in E^N which contains the point **r**. This is so because $\Sigma_{i=1}^{N} a_i dx_i = \mathbf{a} \bullet d\mathbf{r} = 0$ is the equation of a plane in E^N passing through our origin at **r** with **a** being that plane's normal vector. More generally, $\hat{\mathbf{n}} \bullet \mathbf{r} = d$ is the equation of a hyperplane in E^n having a unit normal $\hat{\mathbf{n}}$ and whose closest approach to the origin is distance d.

Thus at a point **r** satisfying a(**r**) = 0 we have constructed a tiny neighborhood of nearby points **r** which also satisfy a(**r**) = 0. This neighborhood is a patch of a hyperplane in E^N which is certainly a piece of "surface" in E^N having dimension N-1. By repeating this process using a mesh of points $\mathbf{r}_i$ satisfying a($\mathbf{r}_i$) = 0, we then map out a triangulated surface of dimension N-1 in E^N. In the limit the mesh size goes to 0, we arrive at a smooth surface of dimension N-1 in E^N and that surface is a(**r**) = 0. The tiny planar patch at any point on the surface is part of the "tangent plane" to the surface at that point.

The constraint a(**r**) = 0 is assumed to be "locally smooth" in the immediate region of any point **r** on the operational constraint surface, so a small local planar region (an open set) can be constructed around any such point. This is the basic idea of a surface being a manifold M, and the set of vectors in the tangent plane of a point **r** on M comprise the "tangent space" of M at **r**, usually denoted by $T_{\mathbf{r}}M$.

****

What happens if x_1 is missing from one or more equations?

$a(x_1,x_2,x_3) = x_1^2 + x_2^2 + x_3^2 - 2^2 = 0$
$b(x_1,x_2,x_3) = x_2^2 + x_3^2 - 2^2 = 0$.

First is a sphere, second is a circle? Well you can regard the first as a sphere dim = 2 in E^3 and the second as a cylinder dim=2 in E^3.

Stop.

These functions are evaluated at some

and this is what we mean by saying we can in theory eliminate x_1, x_2, x_3

, and the problem solution point **r** will lie on one of these pieces.

More generally, if there are S-1 constraint equations, it is later implied that one can eliminate any subset of S-1 variables from the full set of N equations and this is how we demonstrate that the relevant submatrices of the R matrix all have zero determinant.

We shall comment here only on the viability of eliminating x_1, x_2,x_3 in the demonstration case of Section 3 (a), and the reader can easily generalize this viability to the general case.

Reasonableness of the constraint surfaces.

As was noted in ***, each constraint like $a(x_1, x_2, ...x_N) = 0$ describes an N-1 dimensional surface in E^N. This claim itself is perhaps not totally obvious, so we differentiate to get

$$0 = da = a_1 dx_1 + a_2 dx_2 + ... + a_N\, dx_N \qquad (C.1)$$

where a_i means $\partial_i a$. If all N of the a_i vanish, then $a(x_1, x_2, ...x_N)$ is a function of none of its variables and is some equation like $a = \pi$. We do not regard this as a meaningful "constraint", so there must be some $a_s = \partial_s a$ which does not vanish. One can then solve the above equation for dx_s ,

$$dx_s = (a_1 dx_1 + a_2 dx_2 + ... + a_N\, dx_N) / a_s \ . \qquad // \ a_s dx_s \text{ missing from the sum} \quad (C.2)$$

Let $\mathbf{r} = (x_1,x_2...x_N)$ be some point in E^N which satisfies the constraint equation. Suppose we select a set of small numbers for all the dx_i except dx_s. Then (C.2) above gives *the value* for dx_s which sets da = 0. We have then created a displacement d**r** such that **r** + d**r** still satisfies the constraint equation a = 0. In this way we can construct a continuum of such d**r** vectors whose tails lie at **r**, and for all of which a(**r**+d**r**) = 0. These vectors paint out a tiny patch of the constraint surface near point **r**. Continuing this process, we can paint out the entire constraint surface in E^N. The surface has N-1 degrees of freedom since that is how many dx_i are on the right side of (C.2).

When there are multiple constraints, the "legal operating region" for a solution point **r** consists of the intersection of all the constraint surfaces. Each new such constraint surface lowers by one the number of degrees of freedom (the dimensionality) of this surface of intersection. If there is no intersection, then the problem is not well-posed!

So we have to make some assumptions about the set of constraint equations. First of all, they are all "smooth" equations (continuous, differentiable) for which no derivatives like a_i are infinite. At any intersection point of the set of constraint surfaces, each surface can be modeled in the neighborhood of that point as a "plane" through the point having the same dimension as the surface (the tangent plane), allowing us to analyze things in terms of these linearized surfaces. In our region of interest, all the surfaces are formally manifolds being locally smooth at any point.

Consider then the case of three constraint surfaces,

$$a(x_1, x_2, x_3, x_4, x_5, x_6) = 0$$
$$b(x_1, x_2, x_3, x_4, x_5, x_6) = 0$$

$$c(x_1, x_2, x_3, x_4, x_5, x_6) = 0 \qquad (3.2) \qquad \text{(C.3)}$$

each of which is a smooth surface of dimension 5 in E^6. These surfaces intersect in some smooth surface of dimension 3 lying in the same space E^6. It is possible that this intersection surface might have several disjoint pieces, in which case we assume our "operating region" is on one of these pieces. Suppose we select a point **r** on the intersection surface which has certain values x_4, x_5, x_6.

Differentiating gives

$$\begin{aligned} 0 = da &= a_1dx_1 + a_2dx_2 + a_3dx_3 + a_4dx_4 + a_5dx_5 + a_6dx_6 \\ 0 = db &= b_1dx_1 + b_2dx_2 + b_3dx_3 + b_4dx_4 + b_5dx_5 + b_6dx_6 \\ 0 = dc &= c_1dx_1 + c_2dx_2 + c_3dx_3 + c_4dx_4 + c_5dx_5 + c_6dx_6 \end{aligned} \qquad \text{(C.4)}$$

which rewrite as

$$\begin{aligned} a_1dx_1 + a_2dx_2 + a_3dx_3 &= -a_4dx_4 - a_5dx_5 - a_6dx_6 \\ b_1dx_1 + b_2dx_2 + b_3dx_3 &= -b_4dx_4 - b_5dx_5 - b_6dx_6 \\ c_1dx_1 + c_2dx_2 + c_3dx_3 &= -c_4dx_4 - c_5dx_5 - c_6dx_6 \quad . \end{aligned} \qquad \text{(C.5)}$$

Consider a coordinate system whose origin lies at some point **r** on the intersection of the constraint surfaces and whose axes are parallel to the axes of the original coordinate system. In this system, we consider some small vectors of the form $d\mathbf{r} = (dx_1, dx_2, dx_3, dx_4, dx_5, dx_6)$. If we are given values for the components dx_4, dx_5, dx_6, can we compute values for dx_1, dx_2, dx_3 ? Using Cramer's Rule on (C.5) we the answer is yes as long as

$$\det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \neq 0 \ .$$

If this determinant were to vanish, it would mean that the three row vectors are not linearly independent. Since these three vectors are

****

Early in our proof of Theorem 1 in Section 3 (a) we rely on our ability, at least in theory, to eliminate x_1, x_2, x_3 from the three constraint equations shown in (3.3). How does one know that this is even possible?

Each of the three constraint equations in (3.3) represents a smooth surface of dimension 5 in E^6. The intersection of the three surfaces represents a smooth surface of dimension 3 in E^6.