Lecture 2: Continuation of equilibria

2.1 Algebraic Continuation Problems

Consider a system of ODEs depending on one parameter

$$\dot{u} = f(u, \alpha), \quad u \in \mathbb{R}^n, \ \alpha \in \mathbb{R}, \tag{2.1}$$

where $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is smooth. Looking at how its equilibria depend on the parameter, leads to computing the corresponding **equilibrium manifold**, i.e. set of points

$$\begin{pmatrix} u \\ \alpha \end{pmatrix} \in \mathbb{R}^{n+1}$$

satisfying $f(u, \alpha) = 0$. This is an example of a general **Algebraic Continuaton Problem (ALCP)**: Compute a solution set $M \subset \mathbb{R}^{N+1}$ of the smooth system

$$F(x) = 0, \quad F : \mathbb{R}^{N+1} \to \mathbb{R}^N, \tag{2.2}$$

starting form a given point $x_0 \in M$.

2.1.1 Regular points

A point $p \in M$ is called **regular** for ALCP (2.2) if rank $F_x(p) = N$. At such a point, the $N \times (N+1)$ matrix

$$J = F_x(p) = \left. \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \dfrac{\partial F_1}{\partial x_2} & \cdots & \dfrac{\partial F_1}{\partial x_N} & \dfrac{\partial F_1}{\partial x_{N+1}} \\ \dfrac{\partial F_2}{\partial x_1} & \dfrac{\partial F_2}{\partial x_2} & \cdots & \dfrac{\partial F_2}{\partial x_N} & \dfrac{\partial F_2}{\partial x_{N+1}} \\ \cdots & & & & \\ \dfrac{\partial F_N}{\partial x_1} & \dfrac{\partial F_N}{\partial x_2} & \cdots & \dfrac{\partial F_N}{\partial x_N} & \dfrac{\partial F_N}{\partial x_{N+1}} \end{pmatrix} \right|_{x=p}$$

has N linearly-independent rows and there exist N colums which are linearly-independent.

Lemma 2 *Near any regular point p, ALCP (2.2) defines a solution curve M that passes through p and is locally unique and smooth.*

Proof:

Let J_1 be the non-singular $N \times N$ matrix composed by the linearly-independent columns of J. Suppose that the j-th column of J

$$g = \frac{\partial F}{\partial x_j} = \begin{pmatrix} \dfrac{\partial F_1}{\partial x_j} \\ \dfrac{\partial F_2}{\partial x_j} \\ \vdots \\ \dfrac{\partial F_N}{\partial x_j} \end{pmatrix}$$

is their linear combination. The Implicit Function Theorem implies that (locally to p) M is the graph of a smooth function $\mathbb{R} \to \mathbb{R}^N$:

$$\begin{cases} x_1 & = & \varphi_1(x_j), \\ x_2 & = & \varphi_2(x_j), \\ \dots & & \\ x_{j-1} & = & \varphi_{j-1}(x_j), \\ x_{j+1} & = & \varphi_{j+1}(x_j), \\ \dots & & \\ x_{N+1} & = & \varphi_{N+1}(x_j) \end{cases}$$

Taking $s = x_j - p_j$, we get a smooth local parametrization of $M : x = x(s)$. One can use any other smooth local parametrization with $x(0) = p$, i.e. by the archlength. $\square$

Lemma 3 *If p is a regular point of ALCP (2.2) then the linear equation $Jv = 0$ with $J = F_x(p)$ has a unique (modulo scaling) solution $v \in \mathbb{R}^{N+1}$, i.e. the kernel of J is one-dimensional.*

Proof:

$$Jv = 0 \Leftrightarrow J_1 \begin{pmatrix} v_1 \\ \vdots \\ v_{j-1} \\ v_{j+1} \\ \vdots \\ v_{N+1} \end{pmatrix} = -v_j g,$$

where J_1 is non-singular. Thus

$$\begin{pmatrix} v_1 \\ \vdots \\ v_{j-1} \\ v_{j+1} \\ \vdots \\ v_{N+1} \end{pmatrix} = -v_j J_1^{-1} g$$

with arbitrary scaling factor $v_j \in \mathbb{R}$. □

Lemma 4 *A tangent vector v to M at a regular point $p \in M$ satisfies $Jv = 0$.*

Proof:
Consider a smooth parameterization of M near p: $x = x(s)$ with $x(0) = p$. By definition,

$$v = \dot{x}(0) = \left.\frac{dx(s)}{ds}\right|_{s=0}.$$

Notice that one can always select a parameterization such that $\|v\| = 1$. Differentiating the identity $F(x(s)) = 0$ w.r.t. s at $s = 0$ gives:

$$F_x(x(0))\dot{x}(0) = 0$$

or $Jv = 0$. □

The following result is used to compute the kernel of $F_x(x)$ near a regular point p.

Lemma 5 (Keller-Lemma) *The $(N+1) \times (N+1)$ matrix*

$$B = \begin{pmatrix} J \\ v^{\mathrm{T}} \end{pmatrix},$$

where v satisfies $Jv = 0$ and $\|v\| = 1$, is non-singular at any regular point.

Proof:
Suppose that $Bw = 0$ for some $v \in \mathbb{R}^{N+1}$ with $w \neq 0$. This is equivalent to the system of equations

$$\left\{ \begin{array}{rcl} Jw & = & 0, \\ v^{\mathrm{T}}w & = & 0. \end{array} \right.$$

By Lemma 3, the first equation implies that $w = Cv$ with some constant $C \in \mathbb{R}$. Then, the second equation gives

$$0 = Cv^{\mathrm{T}}v = C\|v\|^2 = C,$$

i.e. $C = 0$. This implies $w = Cv = 0$, a contradiction. □

2.1.2 Limit points

A regular point $p \in M$ is a **limit point** for ALCP (2.2) with respect to a coordinate x_j if $v_j = 0$, where v is a normalized tangent vector to M at p.

Lemma 6 *If p is a limit point of ALCP (2.2) w.r.t. x_{N+1}, then the $N \times N$ matrix*

$$A = \left.\left(\frac{\partial F_i(p)}{\partial x_j}\right)\right|_{i,j=1,2,\ldots,N}$$

is singular.

Proof:

Let $x = x(s)$ be a smooth parametrization of M near p such that $x(0) = p$ and

$$\dot{x}(0) = v = \begin{pmatrix} w \\ 0 \end{pmatrix} \in \mathbb{R}^{N+1}$$

with $w \neq 0$. Then

$$J = (A \;\; g), \quad g_i = \left.\frac{\partial F_i(x)}{\partial x_{N+1}}\right|_{x=p} \; (i = 1, 2, \dots, N)$$

and

$$0 = Jv = Aw + v_{N+1} g = Aw,$$

so that $w \in \mathbb{R}^N$ is a nontrivial null-vector of A. $\square$

Since J has rank N at the limit point w.r.t. x_{N+1}, matrix A must have rank $N-1$ (not less!). Therefore, there exists $\psi \in \mathbb{R}^N$ such that $\psi^{\mathrm{T}} A = 0$, or

$$A^{\mathrm{T}} \psi = 0.$$

The vector is unique modulo scaling.

Consider a limit point

$$p = \begin{pmatrix} u_0 \\ \alpha_0 \end{pmatrix}$$

of the equilibrium manifold of (2.1)

$$f(u, \alpha) = 0, \quad f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n,$$

w.r.t. the parameter α. Let

$$x(s) = \begin{pmatrix} u(s) \\ \alpha(s) \end{pmatrix}$$

be a smooth parametrization of the manifold near the limit point such that $u(0) = u_0,\ \alpha(0) = \alpha_0$. The tangent vector to the equilibrium manifold at $x(s)$ will be

$$\dot{x}(s) = \begin{pmatrix} \dot{u}(s) \\ \dot{\alpha}(s) \end{pmatrix}$$

where $\dot{\alpha}(0) = 0$ and $w = \dot{u}(0) \neq 0$ by definition.

Differentiating the identity

$$f(u(s), \alpha(s)) = 0$$

twice w.r.t. s, we obtain

$$\begin{aligned} f_u(u(s), \alpha(s))\dot{u}(s) + f_\alpha(u(s), \alpha(s))\dot{\alpha}(s) &= 0, \\ f_{uu}(u(s), \alpha(s))[\dot{u}(s), \dot{u}(s)] + f_u(u(s), \alpha(s))\ddot{u}(s) + 2f_{\alpha u}(u(s), \alpha(s))[\dot{\alpha}(s), \dot{u}(s)] & \\ + f_\alpha(u(s), \alpha(s))\ddot{\alpha}(s) + f_{\alpha\alpha}(u(s), \alpha(s))\dot{\alpha}(s)\dot{\alpha}(s) &= 0. \end{aligned}$$

Here $f_{uu}(u, \alpha)[w, w] = B(u, \alpha; w, w)$ where

$$B_i(u, \alpha; w, w) = \sum_{j,k=1}^{n} \frac{\partial^2 f_i(u, \alpha)}{\partial u_j \partial u_k} w_j w_k$$

and

$$(f_{\alpha u}(u,\alpha)[\beta,w])_i = \sum_{k=1}^{n} \frac{\partial^2 f_i(u,\alpha)}{\partial\alpha\,\partial u_k}\beta w_k$$

for $i = 1, 2, \dots, n$.

Evaluating the first equation at $s = 0$ and taking into account that $\dot\alpha(0) = 0$, we see that

$$f_u^0 \dot u(0) = 0,$$

where upper index ^{0} indicates the value at (u_0, α_0). Thus (in accordance with Lemma 6)

$$Aw = 0$$

where $A = f_u^0 = f_u(u_0, \alpha_0)$. Evaluation of the second equation at $s = 0$ leads to

$$f_{uu}^0[\dot u(0), \dot u(0)] + f_u^0 \ddot u(0) + f_\alpha^0 \ddot\alpha(0) = 0.$$

Taking the scalar product of the last equation with non-zero vector $\psi \in \mathbb{R}^N$ satisfying $\psi^{\mathrm{T}} A = 0$, we get the following expression:

$$\ddot\alpha(0) = -\frac{\psi^{\mathrm{T}} f_{uu}(p)[w,w]}{\psi^{\mathrm{T}} f_\alpha(p)}.$$

Here $\psi^{\mathrm{T}} f_\alpha(p) \neq 0$ (otherwise $\psi^{\mathrm{T}} J = \psi^T (A\ f_\alpha) = 0$ and rank $J \leq N-1$). A limit point of the equilibrium manifold of (2.1) is called **quadratic** if

$$a = \frac{1}{2}\langle \psi, B(p; w, w)\rangle \neq 0.$$

Locally, $f(u,\alpha) = 0$ looks like a parabola, implying the collision and disappearance of two equilibria as the parameter α passes the limit point value.

2.2 Numerical solutions of ALCP

Solving ALCP (2.2) numerically means: Given an initial point $x^{(0)}$ close to $x_0 \in M$, find a sequence of points

$$x^{(1)}, x^{(2)}, x^{(3)}, \dots$$

such that the union of line segments connecting consequent points approximates M with given accuracy.

This is usually achieved by with a **predictor-corrector method**:

- tangent prediction $X^0 = x^{(i)} + hv^{(i)}$, where h is the **stepsize** and $v^{(i)}$ is tangent to M at $x^{(i)}$; $\|v^{(i)}\| = 1$;
- Newton-like corrections (their type determines the continuation algorithm);
- adaptive step-size control (convergence-dependent).

All defined below corrections converge quadratically to a point $x^{(i+1)}$ in the curve M near $x^{(i)}$, provided the step size h is sufficiently small.

2.2.1 Natural continuation

Apply the standard Newton method to

$$G(x) = \begin{pmatrix} F(x) \\ x_j - X_j^0 \end{pmatrix} = 0,$$

where $|v_j^{(j)}|$ is maximal in absolute value component of $v^{(i)}$. It is equavalent to the Newton corrections in the hyperplane through X^0 orthogonal to the x_j-axis (see Figure 2.1(a)). We have

$$G_x = \begin{pmatrix} F_x \\ [e_j]^{\mathrm{T}} \end{pmatrix},$$

where e_j is the unit vector along the x_j-axis.

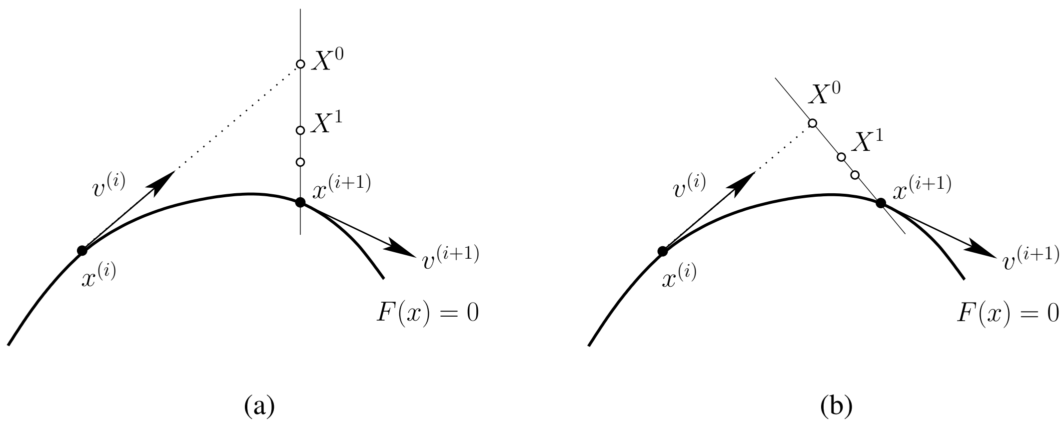

Figure 2.1: Simplest continuation methods: (a) natural continuation (the x_j-axis is assumed to be horizontal); (b) pseudo-arclength continuation.

2.2.2 Pseudo-arclength continuation

Apply Newton's method to

$$G(x) = \begin{pmatrix} F(x) \\ \langle x - X^0, v^{(i)} \rangle \end{pmatrix} = 0.$$

It is equivalent to the Newton corrections in the plane through X^0 orthogonal to $v^{(i)}$ (see Figure 2.1(b)). The linearization matrix

$$G_x = \begin{pmatrix} F_x \\ [v^{(i)}]^{\mathrm{T}} \end{pmatrix}$$

at each iterate is close to the matrix B computed at $x^{(i)}$ and is nonsingular due to Keller-Lemma.

2.2.3 Moore-Penrose continuation

Take $V^0 \in \mathbb{R}^{N+1}$ satisfying $F_x(X^0)V^0 = 0$ and $\|V^0\| = 1$. Make one Newton correction for

$$G(x) = \begin{pmatrix} F(x) \\ \langle x - X^0, V^0 \rangle \end{pmatrix} = 0.$$

The linearization of this system about X^0 is

$$\begin{cases} F(X^0) + F_x(X^0)(X - X^0) &= 0, \\ [V^0]^{\mathrm{T}}(X - X^0) &= 0, \end{cases} \tag{2.3}$$

implying

$$\begin{pmatrix} F_x(X^0) \\ [V^0]^{\mathrm{T}} \end{pmatrix} (X - X^0) = - \begin{pmatrix} F(X^0) \\ 0 \end{pmatrix}.$$

Therefore, define

$$X^1 = X^0 - \begin{pmatrix} F_x(X^0) \\ [V^0]^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} F(X^0) \\ 0 \end{pmatrix}.$$

Then compute V^1 satisfying

$$F_x(X^1)V^1 = 0, \quad \|V^1\| = 1,$$

and set

$$X^2 = X^1 - \begin{pmatrix} F_x(X^1) \\ [V^1]^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} F(X^1) \\ 0 \end{pmatrix},$$

etc.

In general, the Moore-Penrose corrections are defined by

$$X^{k+1} = X^k - \begin{pmatrix} F_x(X^k) \\ [V^k]^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} F(X^k) \\ 0 \end{pmatrix}, \tag{2.4}$$

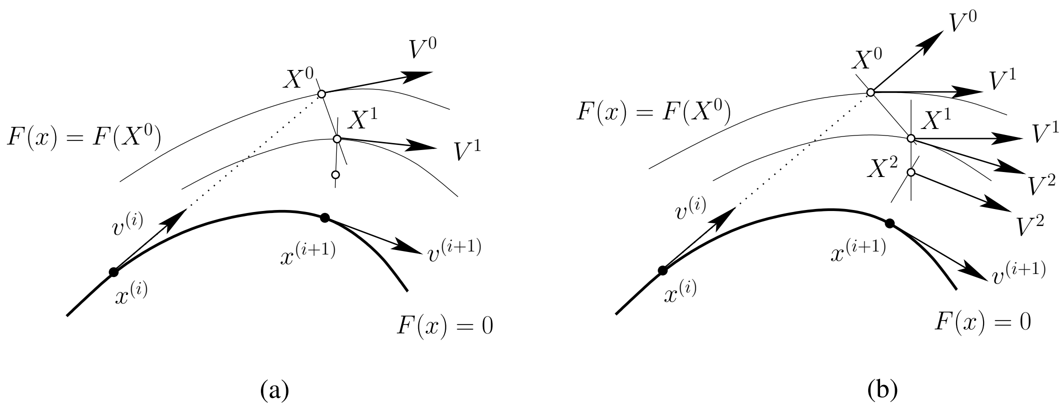

Figure 2.2: Advanced continuation methods: (a) Moore-Penrose continuation; (b) approximate Moore-Penrose continuation.

where

$$F_x(X^k)V^k = 0, \quad \|V^k\| = 1. \tag{2.5}$$

Each correction occurs within the plane orthogonal to the kernel of $F_x(X^k)$ at X^k (see Figure 2.2(a)). If the corrections converge to $x^{(i+1)}$, the corresponding vectors V^k converge to the next tangent vector $v^{(i+1)}$.

Let J be a $N \times (N+1)$ matrix with rank $J = N$. Its **Moore-Penrose pseudo-inverse** is the $(N+1) \times N$ matrix

$$J^+ = J^{\mathrm{T}}(JJ^{\mathrm{T}})^{-1}.$$

Since J has N linearly-independent rows, the corresponding Gramm-determinant $\det(JJ^{\mathrm{T}}) > 0$, so $(JJ^{\mathrm{T}})^{-1}$ exists.

Consider the non-singular linear system for $x \in \mathbb{R}^{N+1}$ with a given $b \in \mathbb{R}^N$

$$\left\{ \begin{array}{rcl} Jx & = & b, \\ v^{\mathrm{T}}x & = & 0, \end{array} \right. \tag{2.6}$$

where $v \in \mathbb{R}^{N+1}$ satisfies $Jv = 0$ and $\|v\| = 1$.

Lemma 7 *The solution to* (2.6) *is given by* $x = J^+b$.

Proof:

$$Jx = JJ^+b = JJ^{\mathrm{T}}(JJ^{\mathrm{T}})^{-1}b = I_N b = b$$

and

$$v^{\mathrm{T}}x = v^{\mathrm{T}}J^+b = v^{\mathrm{T}}J^{\mathrm{T}}(JJ^{\mathrm{T}})^{-1}b = (Jv)^{\mathrm{T}}(JJ^{\mathrm{T}})^{-1}b = 0,$$

since $Jv = 0$. □

Therefore, the first Moore-Penrose correction can be written as

$$X^1 = X^0 - F_x^+(X^0)F(X^0).$$

In general, the corrections defined by (2.4) and (2.5) can be written as

$$X^{k+1} = X^k - F_x^+(X^k)F(X^k), \quad k = 0, 1, 2, \ldots.$$

One can motivate the Moore-Penrose corrections as follows. If a point $x^{(i)} \in M$ is known, one can try to solve the following optimization problem

$$\min_x \{\|x - X^0\| \mid F(x) = 0\}$$

to obtain the next point $x = x^{(i+1)} \in M$. For X^0 close to $x^{(i)}$, this problem is equivalent to solving the system

$$\left\{ \begin{array}{rcl} F(x) & = & 0, \\ v^{\mathrm{T}}(x - X^0) & = & 0, \end{array} \right.$$

where $F_x(x)v = 0$ and $\|v\| = 1$. The linearization of this system about X^0 gives (2.3).

2.2.4 Approximate Moore-Penrose continuation

A disadvantage of the described Moore-Penrose correction algorithm is that one needs to compute the null-vector V^k by setting up and solving (2.5) at each X^k. One can avoid this by looking for X^{k+1} within the plane through X^k that is orthogonal to the *previous* kernel, i.e. V^{k-1} (see Figure 2.2(b)).

Let $V^0 = v^{(i)}$ with $\|V^0\| = 1$. As in the exact Moore-Penrose algorithm, set

$$X^1 = X^0 - \begin{pmatrix} F_x(X^0) \\ [V^0]^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} F(X^0) \\ 0 \end{pmatrix}.$$

To find V^1 satisfying $F_x(X^0)V^1 = 0$, compute first

$$W = \begin{pmatrix} F_x(X^0) \\ [V^0]^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

which amounts to solving a linear system

$$\begin{cases} F_x(X^0)W &= 0, \\ \langle V^0, W \rangle &= 1, \end{cases}$$

with exactly the same matrix as used to compute X^1. The vector W spans the kernel of $F_x(X^0)$. Now we can set

$$V^1 = \frac{W}{\|W\|}$$

and repeat the procedure.

This leads to the following approximate Moore-Penrose corrections:

$$\begin{cases} X^{k+1} &= X^k - \begin{pmatrix} F_x(X^k) \\ [V^k]^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} F(X^k) \\ 0 \end{pmatrix}, \\ W^{k+1} &= \begin{pmatrix} F_x(X^k) \\ [V^k]^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; V^{k+1} = \dfrac{W^{k+1}}{\|W^{k+1}\|}. \end{cases}$$

As in the exact Moore-Penrose case, the vectors V^k converge to the next tangent vector $v^{(i+1)}$. Notice that

$$W^{k+1} = V^k - \begin{pmatrix} F_x(X^k) \\ [V^k]^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} F_x(X^k)V^k \\ 0 \end{pmatrix}.$$