CHAPTER 2

Numerical homotopy continuation

2.1. Polynomial homotopy

In this chapter we consider $\mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_n]$, polynomials with complex coefficients. Moreover, we restrict ourselves to a *square* polynomial systems, $F = (f_1, \ldots, f_n) \in \mathbb{C}[x]^n$, that are 0-dimensional, i.e., $\mathbb{V}(F)$ is finite.

The main idea behind solving the system F is to use the homotopy

$$H_t = (1-t)G + tF \in \mathbb{C}[x]^n, \quad t \in [0,1], \tag{2.1.1}$$

that connects a *start system* $G = H_0$ with the *target system* $F = H_1$. Now we need to create a start system G such that

- solutions of G are readily available;
- as the *continuation parameter* t varies from 0 to 1 we get a smooth paths that lead from solutions of G to solutions of F.

2.1.1. Constructing a start system. The following start system results in the so-called *total-degree homotopy*:

$$G = \left(x_1^{d_1} - 1, \ldots, x_n^{d_n} - 1\right), \tag{2.1.2}$$

where $d_i = \deg f_i$ for $i \in [n]$. The number of solutions equals the *total degree* of the system, $|\mathbb{V}(G)| = d_1 \cdots d_n$. Indeed, the i-th coordinate of a solution is the d_i-th root of unity.

EXAMPLE 2.1.1. *For the target system*

$$F = \begin{pmatrix} x_2^3 - 6x_1^2 - 2x_1x_2 + 5x_2^2 + 2x_1 \\ x_1^2 + x_2^2 - 2 \end{pmatrix}$$

the total-degree start system is

$$G = \begin{pmatrix} x_1^3 - 1 \\ x_2^2 - 1 \end{pmatrix}$$

and the start solutions *are*

$$\left\{ \left(-\frac{1}{2} + \frac{\sqrt{3}}{2}\boldsymbol{i}, 1\right), \quad \left(-\frac{1}{2} - \frac{\sqrt{3}}{2}\boldsymbol{i}, 1\right), \quad (1,1), \right.$$
$$\left. \left(-\frac{1}{2} + \frac{\sqrt{3}}{2}\boldsymbol{i}, -1\right), \quad \left(-\frac{1}{2} - \frac{\sqrt{3}}{2}\boldsymbol{i}, -1\right), \quad (1,-1) \right\}.$$

EXERCISE 2.1.2. *Show that all points $\mathbb{V}(G)$ for G in (2.1.2) are regular.*

Once the homotopy H_t and start solutions are set up, we follow the *homotopy paths* initiating at the start solutions (at $t = 0$) in hope of getting target solutions at $(t = 1)$. A homotopy path, $x(t) \in \mathbb{C}^n$, can be tracked numerically using the so-called *predictor-corrector* technique. Suppose an approximation $\tilde{x}_0 \approx x(t_0)$ of a solution to H_{t_0} is available for some $t_0 \in [0,1]$ and suppose we are looking to find an approximation $\tilde{x}_1 \approx x(t_1)$ of a solution to H_{t_1} for some $t_1 > t_0$. Two steps are performed: the *predictor* makes a rough approximation of $x(t_1)$, then the *corrector* refines predictor's approximation.

2.1.2. Homotopy path tracking: predictor step. Differentiating $H_t(x(t)) = 0$ with respect to t we get

$$\left(\frac{\partial H_t}{\partial x}x'(t) + \frac{\partial H_t}{\partial t}\right)_{x=x(t)} = 0,$$

solving which for the derivative of the homotopy path $x(t)$ gives the following at the point $x_0 = x(t_0)$ of the path

$$x'(t_0) = c(x_0, t_0) = \left(\left(\frac{\partial H_t}{\partial x}\right)^{-1}\frac{\partial H_t}{\partial t}\right)_{x=x_0, t=t_0}. \tag{2.1.3}$$

This is a system of *ordinary differential equations (ODEs)* that can be *integrated numerically* using an arbitrary *numerical integration scheme*. We are, however, interested only in one step of a numerical integration procedure and list several popular choices here.

Order 0: This method makes the simplest prediction possible,

$$\tilde{x}_1 = \tilde{x}_0,$$

not using the ODEs (2.1.3) at all and, in particular, not depending on $\Delta t = t_1 - t_0$. As a global numerical integration scheme it is not very useful, but in our case the following corrector step makes even such a simple predictor step meaningful.

Order 1: The *tangent predictor* goes along the tangent line to the path:

$$\tilde{x}_1 = \tilde{x}_0 + c(\tilde{x}_0, t_0)\Delta t,$$

where the coefficient c is an approximation to $x'(t_0)$ derived by plugging in $x = \tilde{x}_0$ and $t = t_0$ in (2.1.3).

Order 2: The *Euler predictor* operates as follows:

$$\tilde{x}_1 = \tilde{x}_0 + \frac{c(t_0, \tilde{x}_0) + c\left(\tilde{x}_0 + c(t_0, \tilde{x}_0)\Delta t,\, t_0\right)}{2}\Delta t.$$

Note that this formula involves a part that is an exact copy of the tangent method.

Order m: In general, one can construct an integration scheme of an arbitrary order m. We skip the formal definition of the concept of *order*. Roughly speaking, a scheme is of order m if the error of approximatican be bounded from above as a constant multiple of $|\Delta t|^{m+1}$ as $\Delta t \to 0$.

The higher the order, the more accuracy is expected. On the other hand, the higher order methods typically are more involved in comparison to the lower order methods as demonstrated by the schemes of orders 0,1, and 2 above.

One of the most popular techniques is the *classical Runge-Kutta method*, which is or order 4.

2.1.3. Homotopy path tracking: correction step. One can extend *Newton's method* discussed in §1.1.2 to the multivariate setting. Let F be a square polynomial system of size n and assume one has an approximation $\tilde{x}_0$ of an exact solution

$x^* \in \mathbb{C}^n$. Then it can be refined by *Newton's operator*

$$N_F : \left\{ x \in \mathbb{C}^n \mid \det(\frac{\partial F}{\partial x})(x) \neq 0 \right\} \to \mathbb{C}^n,$$
$$N_F(x) = x - \left(\frac{\partial F}{\partial x}(x)\right)^{-1} F(x).$$

using Algorithm 1.1.1 almost word for word.

As long as the solution x^* is regular, which in case of a square system is equivalent to nonvanishing of the determinant of the Jacobian matrix $\frac{\partial F}{\partial x}(x^*)$, Newton's method *converges quadratically*. This, roughly speaking, means that the accuracy of the approximation (the number of correct digits) doubles at every step. We leave the exact definition of quadratic convergence to §2.3.

EXERCISE 2.1.3. *Design an algorithm*

(1) $\tilde{x} = \text{CORRECTOR}(H, x_0, t)$ *that takes a homotopy* H*, an approximation* x *to a solution of the system* $F = H_t$ *for a given* t *and outputs* $N_F(x_0)$*.*
(2) $\tilde{x} = \text{PREDICTOR}(H, \tilde{x}_{t-\Delta t}, t, \Delta t)$ *that takes*
   - *a homotopy* H*,*
   - *an approximation* $\tilde{x}_{t-\Delta t}$ *to a solution of* $H_{t-\Delta t}$
   - *for a given* t
   - *and the step size* Δt

   and outputs the prediction for a solution of H_t *using Euler's method.*

2.1.4. Heuristic homotopy tracking. The word *heuristic* suggests that the algorithms described below terminate and produce a correct output for a large number of inputs and parameter settings, however, they do not provide a guarantee of neither termination nor correctness.

The most naïve algorithm of tracking a homotopy path is the following.

Algorithm 2.1.1 $\tilde{x}_1 = \text{NAIVEHOMOTOPYTRACKING}(H, \tilde{x}_0, N)$

Require: $H = H_t$, a polynomial homotopy as in 2.1.1;
$\tilde{x}_0 \in k$, an approximation to a solution of H_0;
N, the number of steps taken on the homotopy path;
Ensure: $\tilde{x}_1$, an approximation of a solution of H_1.

```
Δt ← 1/N
for t = Δt to 1 with step Δt do
  x̂_t ← PREDICTOR(H, x̃_{t−Δt}, t, Δt)
  x̃_t ← CORRECTOR(H, x̂_t, t)
end for
```

The hope is that for a large enough value of N the naïve algorithm does not deviate from the homotopy path. A more sophisticated approach adjusts the size of the step in accordance with the "difficulty" of prediction and correction. Algorithm 2.1.2 presents one popular approach.

Note: The existing practical implementations of the heuristic homotopy tracking algorithms take even more parameters than TRACKHOMOTOPY in Algorithm 2.1.2. For instance, see the function `track` of `NumericalAlgebraicGeometry` package of `Macaulay2`.

Algorithm 2.1.2 $\tilde{x}_1 = \text{TRACKHOMOTOPY}(H, \tilde{x}_0, \Delta t_0, c, \delta, max_{corr})$

Require: $H = H_t$, a polynomial homotopy as in 2.1.1;
$\tilde{x}_0 \in k$, an approximation to a solution of H_0;
Δt_0, an initial step size;
c, the step increase factor;
δ, the backward error tolerance;
max_{corr}, the maximal number of corrections;
Ensure: $\tilde{x}_1$, an approximation of a solution of H_1.

$\tilde{x} \leftarrow \tilde{x}_0$, $t \leftarrow 0$
repeat
  repeat
    if $t + \Delta t > 1$ **then**
      $\Delta t = 1 - t$
    end if
    $\tilde{x}' \leftarrow \text{PREDICTOR}(H, \tilde{x}, t, \Delta t)$
    $success \leftarrow$ **false**, $i \leftarrow 0$
    while not $success$ and $i < max_{corr}$ **do**
      $\tilde{x}'' \leftarrow \text{CORRECTOR}(H, \tilde{x}', t)$
      if $|\tilde{x}'' - \tilde{x}'| < \delta$ **then**
        $success \leftarrow$ **true**
      end if
      $\tilde{x}' \leftarrow \tilde{x}''$, $i \leftarrow i + 1$
    end while
    if $success$ **then**
      $\Delta t \leftarrow \min(c\Delta t, 1 - t)$ -- increase the step size
    else
      $\Delta t \leftarrow c^{-1}\Delta t$ -- decrease the step size
    end if
  until $success$
  $t \leftarrow t + \Delta t$, $\tilde{x} \leftarrow \tilde{x}'$
until $t = 1$
$\tilde{x}_1 \leftarrow \tilde{x}$

Suppose the homotopy path $x(t)$ is *smooth* at every point perhaps with an exception of $t = 1$, i.e.,

$$\det\left(\frac{\partial H_t}{\partial x}(x(t))\right) \neq 0, \quad t \in [0, 1).$$

Above Algorithms 2.1.1 and 2.1.2, in fact, terminate and produce correct results for a sufficiently close approximation $\tilde{x}_0$ of $x(0)$, and sufficiently large N and Δt_0, respectively. However, it is practically impossible to determine the sufficient values of the parameters in the general case.

2.1.5. Randomization and γ-trick. For a path $x(t)$ containing a singular point, i.e., a singular solution $x^* = x(t^*)$ of the system H_{t^*} for some $t^* \in [0, 1)$, all numerical homotopy tracking algorithms are likely to fail: the corrector step becomes *ill-conditioned* close to x^*, since the Jacobian matrix is not invertible at x^*.

Note: A good indicator of how close to singularity in the condition number of the Jacobian of H_{t^*} at the current approximation of $x(t^*)$.

For a regular $n \times n$ matrix A the *condition number* is defined as

$$\kappa(A) = \|A\| \, \|A^{-1}\|, \quad \|A\| = \max_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|}.$$

The larger is $\kappa(A)$, the less reliable is a numerical approximate solution of a linear system $Ax = b$.

EXERCISE 2.1.4. *Consider a homotopy*

$$H_t(x) = (1-t)G(x) + \gamma t F(x). \tag{2.1.4}$$

Let $G = x^2 - 1$ and $F = x^2 + 2x - 3$. Find all $\gamma \in \mathbb{C}$ such that there exists $t \in \mathbb{R}$ with a singular solution x to H_t.

Exercise 2.1.4 suggests a way to avoid such singularities replacing a homotopy of (2.1.1) with that of (2.1.4) picking the value of $\gamma \in \mathbb{C}$ at random. This is known as the *γ-trick*.

Note: In practice, γ is picked on the unit circle in the complex plane with uniform distribution.

The following result allows us to compute *all* solutions of the target system via homotopy continuation.

THEOREM 2.1.5. *Let F be a 0-dimensional square polynomial system.*

Then for the homotopy (2.1.4) using the total-degree start system G and a generic $\gamma \in \mathbb{C}$

1. *for every start solution $x_0 \in \mathbb{V}(G)$ the homotopy path $x(t)$ starting at $x_0 = x(0)$ is regular for $t \in [0, 1)$;*
2. *every target solution $x_1 \in \mathbb{V}(F)$ is at the end of some homotopy path $x(t)$, i.e., $x_1 = x(1)$.*

PROOF. add a reference □

COROLLARY 2.1.6 (Bézout bound). *Let $F = (f_1, \ldots, f_n)$ with $f_i \in \mathbb{C}[x_1, \ldots, x_n]$ of degree d_i for $i = 1, \ldots, n$ be a 0-dimensional polynomial system.*

Then $|\mathbb{V}(F)| \leq d_1 \cdots d_n$.

2.1.6. Non-square systems. Consider a 0-dimensional *overdetermined* system

$$F = \begin{pmatrix} f_1 \\ \ldots \\ f_m \end{pmatrix} \in \mathbb{C}[x_1, \ldots, x_n]^m, \quad m > n.$$

We shall describe two ways of *squaring-up* an overdetermined system.

Take a matrix $A \in \mathbb{C}^{n \times m}$ and consider the square system $AF \in \mathbb{C}[x]^n$. Since the polynomials of AF are linear combinations of f_i,

$$\mathbb{V}(F) \subseteq \mathbb{V}(AF).$$

PROPOSITION 2.1.7. *For a 0-dimensional system F and a generic A as above, the system AF is 0-dimensional.*

PROOF. See Exercise 2.1.10 □

Note: If solution $x^* \in \mathbb{V}(F)$ is regular with respect to F, then it is regular with respect to AF in Proposition 2.1.7.

EXERCISE 2.1.8. *Construct an example of an overdetermined* 0*-dimensional system* F *in* n *variables such that every subset of* n *polynomials in* F *forms a positive-dimensional system.*

Proposition 2.1.7 suggests a simple method of finding $\mathbb{V}(F)$:

(1) pick a random A;
(2) find $\mathbb{V}(AF)$ for the square system AF;
(3) $\mathbb{V}(F) = \{\, x \in \mathbb{V}(AF) \,|\, F(x) = 0 \,\}$, which amounts to a finite number of checks for polynomial vanishing.

An alternative to the method above is to introduce the so-called *slack variables* $Y = (y_1, \ldots, y_{m-n})^T$. For a matrix $A \in \mathbb{C}^{m \times (m-n)}$, consider the square system

$$F + AY \in \mathbb{C}[x, y]^m.$$

We readily observe that solutions of F extend to solutions of $F + AY$ by appending zeros for y-coordinates.

$$\mathbb{V}(F + AY) \supseteq \{\, (x, 0) \in \mathbb{C}^m \,|\, x \in \mathbb{V}(F) \,\}.$$

PROPOSITION 2.1.9. *For a* 0*-dimensional system* F *and a generic* A *as above, the system* $F + AY$ *is* 0*-dimensional and*

$$\{\, (x, y) \in \mathbb{V}(F + AY) \,|\, y = 0 \,\} = \{\, (x, 0) \in \mathbb{C}^m \,|\, x \in \mathbb{V}(F) \,\}.$$

PROOF. add a reference □

EXERCISE 2.1.10. *Prove Proposition 2.1.7 by reducing to Proposition 2.1.9.*

2.1.7. Eigenvalue problem via homotopy continuation. The following problem is central in linear algebra; here we use numerical homotopy continuation technique to solve it approximately.

PROBLEM 2.1.11. *Given* $A \in \mathbb{C}^{n \times n}$ *find the eigenvalues and eigenvectors of* A*.*

Consider the case $n = 2$; the approach outlined here works for all n. Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad v = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

There are three unknowns in the eigenproblem: the eigenvalue λ and the coordinates of the vector v. We seek solutions of the system

$$(Av - \lambda v) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 - \lambda x_1 \\ a_{21}x_1 + a_{22}x_2 - \lambda x_2 \end{pmatrix}.$$

This is a system of two quadratic polynomials in $\mathbb{C}[\lambda, x_1, x_2]$. The system is positive-dimensional: there are fewer equations than unknowns. Indeed, if v is an eigenvector, then there is the whole line of solutions corresponding to Span (v).

EXERCISE 2.1.12. *For the* 2×2 *matrix* A *find a condition (involving the entries* a_{ij}*) that ensures that the characteristic polynomial of* A *has a double root.*

Augment the system with one linear equation:

$$F = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 - \lambda x_1 \\ a_{21}x_1 + a_{22}x_2 - \lambda x_2 \\ b_1x_1 + b_2x_2 + 1 \end{pmatrix}.$$

Assume that the eigenspaces of A are one-dimensional (this is the case, in particular, when the eigenvalues are not repeated; cf. Exercise 2.1.12). If $b_1, b_2 \in \mathbb{C}$ are generic, the last equation in the system picks out one nonzero vector in each eigenspace. We conclude that F is a 0-dimensional system generically, since a randomly picked matrix has distinct eigenvalues.

Now if we solve the system F, consisting of two quadratic and one linear polynomials, using the total-degree homotopy, we would have to track $4(= 2 \cdot 2 \cdot 1)$ paths. However, $|\mathbb{V}(F)| = 2$, meaning that this homotopy is not *optimal*.

Instead, let us take a start system that arises from the eigenproblem that we know a solution to: take a diagonal matrix D with distinct entries on the diagonal,

$$D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}, \quad d_1 \neq d_2$$

In a similar way as above we translate D into the 0-dimensional system

$$G = (Dv - \lambda v) = \begin{pmatrix} d_1 x_1 - \lambda x_1 \\ d_2 x_2 - \lambda x_2 \\ b_1 x_1 + b_2 x_2 + 1 \end{pmatrix}. \tag{2.1.5}$$

The two start points $(\lambda, x_1, x_2) \in \mathbb{V}(G)$ are $(d_1, \frac{-1}{b_1}, 0)$ and $(d_2, 0, \frac{-1}{b_2})$.

EXERCISE 2.1.13. *Show that the solutions of the system G in (2.1.5) are regular (for generic b_1 and b_2).*

The homotopy connecting G to F is

$$\begin{aligned} H_t = (1-t)G + tF &= \begin{pmatrix} (tD + (1-t)A)\, v - \lambda v \\ b_1 x_1 + b_2 x_2 + 1 \end{pmatrix} \\ &= \begin{pmatrix} ((1-t)d_1 + ta_{11}) & x_1 & + & ta_{12} & x_2 & - & \lambda x_1 \\ ta_{21} & x_1 & + & ((1-t)d_2 + ta_{22}) & x_2 & - & \lambda x_2 \\ b_1 & x_1 & + & b_2 & x_2 & + & 1 \end{pmatrix}. \end{aligned} \tag{2.1.6}$$

For all $t \in [0,1]$ the system H_t represents the eigenproblem for the matrix

$$M_t = (1-t)D + tA. \tag{2.1.7}$$

For almost all choices of the entries d_1 and d_2 of D, the eigenvalues of M_t are distinct and the solutions $\mathbb{V}(H_t)$ are regular for all $t \in [0,1)$.

EXERCISE 2.1.14. *Consider the homotopy M_t in (2.1.7) for*

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \textit{and} \quad D = \begin{bmatrix} \gamma & 0 \\ 0 & 0 \end{bmatrix}, \quad \gamma \in \mathbb{C}$$

Find the locus of values of γ that produce M_t with a repeated eigenvalue (see Exercise 2.1.12) for some $t \in \mathbb{R}$.

2.2. Singular solutions

The γ-trick of §2.1.5 makes sure the homotopy H_t is such that the points of the variety $\mathbb{V}(H_t)$ are regular for all $t \in [0, 1)$. This guarantees fast convergence of the corrector step at every point of the path with the possible exception at the very end.

This section discusses ways to detect singular solutions and regularize the system $F = H_1$, therefore restoring the quadratic convergence of Newton's method.

2.2.1. Path clustering. We say that a solution $x^* \in \mathbb{V}(F)$ is *isolated* if there is a neighborhood of x^* containing no other solution. If F is 0-dimensional then all its solutions are isolated.

Define *path multiplicity* of $x^* \in \mathbb{V}(F)$ relative to homotopy H_t with $H_1 = F$ to be the number of homotopy paths $x(t)$ with $x(1) = x^*$.

THEOREM 2.2.1. *Let F be a 0-dimensional square polynomial system.*

Then the path multiplicity of $x^* \in \mathbb{V}(F)$ relative to a total-degree homotopy (2.1.4) for a generic $\gamma \in \mathbb{C}$ is the same.

We will denote this generic multiplicity by $\mu = \mu(x^*)$.

Note: If $x^{(1)}(t), \ldots, x^{(\mu)}(t)$ are the homotopy paths converging to x^* as $t \to 1$, their *centroid*

$$x_c(t) = \frac{x^{(1)}(t) + \cdots + x^{(\mu)}(t)}{\mu}$$

converges to x^* asymptotically faster.

This fact can be used in practical computation.

2.2.2. Deflation. First, let us consider the univariate case: suppose $f \in \mathbb{C}[x]$ has a multiple root x^*. Then $f = (x - x^*)^\mu g$ for some $\mu > 1$ where g is not divisible by $x - x^*$.

Differentiating f we get

$$f' = \mu(x - x^*)^{\mu-1} g + (x - x^*)^\mu g' = (x - x^*)^{\mu-1}(\mu g + (x - x^*)g').$$

Since $\mu g + (x - x^*)g'$ does not have x^* as a root, we conclude that f' has x^* as a root with multiplicity $\mu - 1$.

As was mentioned in Chapter 1, Newton's method converges much slower around a multiple root. One way to restore the fast quadratic convergence is to approximate x^* as a root of $f^{(\mu-1)}$.

However, in practice, the multiplicity μ may be unknown. Suppose we have a modified NEWTON subroutine that gives up if the convergence is (heuristically) deemed to be slow.

Algorithm 2.2.1 $(\tilde{x}, fast) = \text{NEWTONFAST}(f, x_0, \delta)$

Require: $f \in k[x]$, a polynomial;
$x_0 \in k$, an initial approximation;
δ, the desired absolute error tolerance;

Ensure: either $fast =$ **true** (convergence is quadratic) and $\tilde{x}$ is an approximate root with backward error estimated to be at most δ
or $fast =$ **false** (convergence is not quadratic) and $\tilde{x}$ is a possibly finer approximation with no estimate on the error.

Then it is possible to design a heuristic regularization algorithm that takes one derivative at a time.

Algorithm 2.2.2 $\tilde{x} = \text{UNIVARIATEDEFLATION}(f, x_0, \delta)$

Require: $f \in k[x]$, a polynomial;
$x_0 \in k$, an initial approximation to a root of f;
δ, the absolute error tolerance;
Ensure: $\tilde{x}$ is an approximate root with backward error estimated to be at most δ.

```
x̃ ← x_0
repeat
  (x̃, fast) = NewtonFast(f, x̃, δ)
  if not fast then
    f ← f′
  end if
until fast
```

Now suppose we have a system of equations $F = (f_1, \ldots, f_m)$ of multivariate polynomials $f_i \in \mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_n]$ with an isolated solution x^*. This implies $m \geq n$.

If x^* is singular, then the Jacobian $J = \frac{\partial F}{\partial x}(x^*)$ is rank deficient and has a nonzero kernel. That means that

$$r = \operatorname{rank} J < n \text{ and}$$
$$c = \operatorname{corank} J = n - r = \dim \ker J > 0.$$

Pick a generic constant matrix $A = (a_{ij}) \in \mathbb{C}^{c \times n}$ and consider the following augmented system of polynomials in $\mathbb{C}[x, \lambda] = \mathbb{C}[x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_n]$:

$$\begin{pmatrix} F \\ \frac{\partial F}{\partial x}\,\lambda \\ A\lambda + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \end{pmatrix} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \\ \frac{\partial f_1}{\partial x_1}\lambda_1 + \cdots + \frac{\partial f_1}{\partial x_n}\lambda_n \\ \vdots \\ \frac{\partial f_m}{\partial x_1}\lambda_1 + \cdots + \frac{\partial f_m}{\partial x_n}\lambda_n \\ a_{11}\lambda_1 + \cdots + a_{1n}\lambda_n + 1 \\ \vdots \\ a_{c1}\lambda_1 + \cdots + a_{cn}\lambda_n + 1 \end{pmatrix} \tag{2.2.1}$$

The three blocks of polynomials in the system correspond to

(1) the original equations;
(2) equations that place the column vector of indeterminates λ in the kernel of the Jacobian;
(3) equations that describe a random r-plane (of dimension $r = n - c$) in the space of parameters λ (of dimension n).

PROPOSITION 2.2.2. *For a generic choice of A the system (2.2.1) has an isolated solution $(x^*, \lambda^*) \in \mathbb{C}^{2n}$ for some $\lambda^* \in \mathbb{C}^n$.*

PROOF. Set $x = x^*$, then the second block of equations in the augmented system describes $K = \ker J$, where $J = \frac{\partial F}{\partial x}(x^*)$. Intersecting K, which is of dimension $c = n - r$, with a random r-plane cut out by the third block we get one point λ^*.

This leads to the conclusion that the (x^*, λ^*) solves the augmented system and is isolated. □

Taking a closer look at the system (2.2.1) one may notice that c parameters can be easily eliminated using the third block of equations. Alternatively, one can reformulate the augmentation procedure: take a generic constant matrix $B \in \mathbb{C}^{n\times(r+1)}$ and consider the system

$$D_B F = \begin{pmatrix} F \\ \frac{\partial F}{\partial x}\, B \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_r \\ 1 \end{bmatrix} \end{pmatrix} \in \mathbb{C}[x, \lambda]^{2m}, \tag{2.2.2}$$

where $\mathbb{C}[x, \lambda] = [x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_r]$. Note that the *deflation* procedure, denoted by D (we write D_B if the choice of a generic matrix B needs to be emphasized), produces DF, a system of m equations in $n + r$ unknowns, while the system (2.2.1) has $2m + c$ equations in $2n$ unknowns.

EXERCISE 2.2.3. *Consider the system*

$$F = \begin{pmatrix} x_1^2 - x_2^4 \\ x_1^2 - x_2^6 \end{pmatrix}.$$

1. *Show that the origin $x^* = (0, 0)$ is a singular isolated solution of F.*
2. *Find the rank r of the Jacobian at x^*.*
3. *Construct the deflated system $DF = D_B F$ picking entries of B to be nonzero integers.*
4. *Is the lifted solution $(x^*, \lambda^*) \in \mathbb{V}(DF)$ regular?*

PROPOSITION 2.2.4. *For a generic choice of B the system (2.2.2) has an isolated solution $(x^*, \lambda^*) \in \mathbb{C}^{n+r}$ for some $\lambda^* \in \mathbb{C}^r$.*

PROOF. The argument is similar to that of Proposition 2.2.2.

Note that $B(\lambda, 1)^T$ parametrizes an r-plane and the the second block of equations in 2.2.2 implies that $(x^*, \lambda) \in \mathbb{V}(DF)$ iff

$$B \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \in \ker J, \quad J = \frac{\partial F}{\partial x}(x^*).$$

There is a unique point λ^* in the intersection of the $(n - r)$-dimensional kernel and a generic r-plane, which leads to the conclusion. □

The following algorithm can be seen as a generalization of Algorithm 2.2.2 to the multivariate setting. It relies on two subroutines:

- a multivariate generalization of NEWTONFAST(F, x_0, δ), which implements the multivariate Newton's method as in §2.1.3 (if given an overdetermined

system F, Newton's method is applied to a square system formed using one of the approaches in §2.1.6);
- a heuristic function NUMERICALRANK($\tilde{A}$), which given an approximation $\tilde{A}$ of a matrix A attempts to recover the (exact) rank of A.

Algorithm 2.2.3 $\tilde{x} = \text{DEFLATION}(F, x_0, \delta)$

Require: $F \in \mathbb{C}[x]^m$, a polynomial system;
$x_0 \in \mathbb{C}^n$, an initial approximation to an isolated solution of F;
δ, the absolute error tolerance;
Ensure: $\tilde{x}$ is an approximate solution with the estimated backward error at most δ.

$(\tilde{x}, fast) \leftarrow \text{NEWTONFAST}(F, x_0, \delta)$
if not $fast$ **then**
  $J \leftarrow \dfrac{\partial F}{\partial x}(\tilde{x})$
  $r \leftarrow \text{NUMERICALRANK}(J)$
  $B \leftarrow$ a random $n \times (r+1)$ matrix
  $\tilde{\lambda} \leftarrow$ the least-squares approximate solution of $JB \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = 0$
  $\tilde{x} \leftarrow$ the first n coordinates of $\text{DEFLATION}(D_B F, (\tilde{x}, \tilde{\lambda}), \delta)$
end if

Unlike in the univariate case it is not clear whether the recursion in DEFLATION terminates. In fact, one can define an intrinsic notion of *multiplicity* of an isolated solution $\nu(x^*)$, such that $\nu(x^*) = 1$ iff x^* is regular, and prove that it decreases when deflation is applied.

THEOREM 2.2.5. *Let $x^* \in \mathbb{V}(F)$ be an isolated solution of a polynomial system F and $(x^*, \lambda^*) \in \mathbb{V}(DF)$ be the corresponding solution of the deflated system DF. Then $\nu(x^*, \lambda^*) < \nu(x^*)$.*

PROOF. add a reference □

COROLLARY 2.2.6. *Algorithm 2.2.3 terminates.*

EXAMPLE 2.2.7. *Consider*

$$F = \begin{pmatrix} x_1^3 - x_2^2 \\ x_1 x_2^2 \\ x_2^3 \end{pmatrix}.$$

This system has one solution, $x^* = (0, 0)$. However, x^* is singular:

$$\frac{\partial F}{\partial x} = \begin{bmatrix} 3x_1^2 & -2x_2 \\ x_2^2 & 2x_1 x_2 \\ 0 & 3x_2^2 \end{bmatrix}, \quad \frac{\partial F}{\partial x}(x^*) = 0, \quad r = 0, \quad c = 2.$$

The first deflation step is not using any new variables: take a generic $B_1 \in \mathbb{C}^{2\times 1}$, *for this example* $B_1 = \begin{bmatrix}1\\1\end{bmatrix}$ *works, then*

$$D_{B_1}F = \begin{pmatrix} F \\ \frac{\partial F}{\partial x}\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}1\end{bmatrix} \end{pmatrix} = \begin{pmatrix} x_1^3 - x_2^2 \\ x_1x_2^2 \\ x_2^3 \\ 3x_1^2 - 2x_2 \\ x_2^2 + 2x_1x_2 \\ 3x_2^2 \end{pmatrix} \in \mathbb{C}[x_1, x_2]^6.$$

However, x^* *is singular for this system as well:*

$$\frac{\partial F}{\partial x} = \begin{bmatrix} 3x_1^2 & -2x_2 \\ x_2^2 & 2x_1x_2 \\ 0 & 3x_2^2 \\ 6x_1 & -2 \\ 2x_2 & 2(x_2+x_1) \\ 0 & 6x_2 \end{bmatrix}, \quad \frac{\partial F}{\partial x}(x^*) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad r = 1, \quad c = 1.$$

In the second deflation step we take a generic $B_2 \in \mathbb{C}^{2\times 2}$, *here* $B_2 = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$ *works, to construct*

$$D_{B_2}D_{B_1}F = \begin{pmatrix} F \\ \frac{\partial (D_{B_1}F)}{\partial x}\begin{bmatrix}1\\\lambda_1\end{bmatrix} \end{pmatrix} = \begin{pmatrix} x_1^3 - x_2^2 \\ x_1x_2^2 \\ x_2^3 \\ 3x_1^2 - 2x_2 \\ x_2^2 + 2x_1x_2 \\ 3x_2^2 \\ 3x_1^2 - 2x_2\lambda_1 \\ x_2^2 + 2x_1x_2\lambda_1 \\ 3x_2^2\lambda_1 \\ 6x_1 - 2\lambda_1 \\ 2x_2 + 2(x_2+x_1)\lambda_1 \\ 6x_2\lambda_1 \end{pmatrix} \in \mathbb{C}[x_1, x_2, \lambda_1]^{12}.$$

The above system has one solution $(x^*, \lambda^*) = (0, 0, 0)$ *that is regular, checking which we leave as an exercise.*

Note: In practice, to avoid doubling the number of equations at each deflation step, one may want to square-up the deflated after each deflation: use the first approach of §2.1.6.

EXERCISE 2.2.8. *For the system* F *of Exercise 2.2.3 construct a sequence of* s *deflations picking matrices* $B_1, \ldots, B_s$ *(take* s *to be as large as needed) such that the resulting system* $D_{B_s}\cdots D_{B_1}F$ *has a regular solution projecting to the origin.*

2.3. Certification

In this section we go back to the analysis of Newton's method and outline the cornerstone results of Smale's alpha theory. These can be used to show that heuristically obtained approximate solutions are certifiably correct.

2.3.1. Approximate zeros. Let $F \in \mathbb{C}[x]^n$ be a square system of polynomials. For $m \in \mathbb{N}$, let

$$N_F^m(x) \;:=\; \underbrace{N_F \circ \cdots \circ N_F(x)}_{m \text{ times}}$$

be the m^{th} Newton iteration of F starting at x. Let $\|\cdot\|$ be the *hermitian norm* on $\mathbb{C}^n$:

$$\|(x_1, \ldots, x_n)\| = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}.$$

A point x is a *approximate zero* of F with the *associated zero* $x^* \in \mathbb{V}(F)$ if

$$\|N_F^m(x) - x^*\| \;\leq\; \left(\frac{1}{2}\right)^{2^m-1} \|x - x^*\|, \tag{2.3.1}$$

for every $m \in \mathbb{N}$. In other words, the sequence $\{N_F^m(x) \mid m \in \mathbb{N}\}$ *converges quadratically* to x^*.

2.3.2. Smale's α-theorem. Smale's α-theory provides sufficient conditions for a given point x to be a approximate zero of F. It operates with the numbers $\alpha(F,x)$, $\beta(F,x)$, and $\gamma(F,x)$ that are defined if the Jacobian $J(x) = \frac{\partial F}{\partial x}(x)$ is invertible:

$$\beta(F,x) = \|x - N_F(x)\| \;=\; \|J(x)^{-1}F(x)\|$$

was used before as the absolute backward error estimator,

$$\gamma(F,x) = \sup_{m \geq 2} \left\| \frac{J(x)^{-1} \frac{\partial^m F}{\partial x^m}(x)}{m!} \right\|^{\frac{1}{m-1}},$$

and

$$\alpha(F,x) = \beta(F,x)\gamma(F,x).$$

Note: Beyond the univariate case, the higher-order derivatives of F and the norm used in the definition of γ have highly nontrivial descriptions. add a reference

Nevertheless, these are computable. So is γ, since the supremum is taken over a finite number of values m: the derivatives of order higher than the order of polynomials vanish.

THEOREM 2.3.1. *The point $x \in \mathbb{C}^n$ with*

$$\alpha(F,x) \;<\; \frac{13 - 3\sqrt{17}}{4} \;\approx\; 0.157671 \tag{2.3.2}$$

is a approximate zero of F. Moreover, $\|x - x^*\| \leq 2\beta(F,x)$ where $x^* \in \mathbb{V}(F)$ is the associated zero for x.

PROOF. add a reference □

EXERCISE 2.3.2. *For a polynomial $f = x^2 - 2x + 3$ determine whether the point x passes the α-test (2.3.2) for*

(1) $x = 1$;
(2) $x = 1 + \boldsymbol{i}$;
(3) $x = 1 + \frac{3}{2}\boldsymbol{i}$.

If so, what is the associated zero of the point?

THEOREM 2.3.3. *Let $x \in \mathbb{C}^n$ with $\alpha(F,x) < 0.03$ and $x^* \in \mathbb{V}(f)$ the associated zero for x. If $y \in \mathbb{C}^n$ satisfies*

$$\|x - y\| \ < \ \frac{1}{20\gamma(F,x)}\,, \tag{2.3.3}$$

then y is a approximate zero of F with associated solution x^*.

PROOF. add a reference □

EXERCISE 2.3.4. *For a polynomial $f = x^2 - 1$ find an upper bound on $\varepsilon > 0$ such that $x = 1+\varepsilon$ and $y = 1-\varepsilon$ pass the* robust α-test*, i.e., satisfy the hypotheses of Theorem 2.3.3.*

EXERCISE 2.3.5. *Let $f \in \mathbb{R}[x]$, assume that all points in $\mathbb{V}(f) \subset \mathbb{C}$ are regular, and let $Z \subset \mathbb{C}$ be a set of $\deg f$ approximate zeros associated to distinct zeros of f.*

Design a procedure that selects approximate zeros in Z that are associated to real zeros of f.