Math 244 — Spring 2003

Nonlinear Systems

The key ideas of **nullcline**, **equilibrium point** (also called **critical point**), **separatrices** and **linearization** are introduced in Chapter 9, but definitions are spread over the chapter. This makes it difficult to identify material that will be needed for exams. These notes collect definitions in one place and emphasize techniques that should be mastered at the start of the study of nonlinear system. Useful exercises for these methods are **Exercises 5 – 16 of Section 9.3** and **Exercises 1 – 6 of Section 9.**.

These systems we consider are all **autonomous**. That is, the independent variable in the equation is t, (which usually represents time) and the derivatives of the dependent variables defining the system are independent of t. We fix the notation

$$\begin{aligned}\frac{dx}{dt} &= F(x, y)\\ \frac{dy}{dt} &= G(x, y)\end{aligned} \tag{1}$$

used in Equation (1) of Section 9.2 for the system of equations in the two-dimensional case.

For autonomous equations, the existence and uniqueness theorems imply that, if $\big(x(t), y(t)\big)$ is a solution of (1), so is $\big(x(t+c), y(t+c)\big)$. In other words, two functions that **draw the same curves in the same way with different starting points** satisfy the same systems (1). This allows the role of t to be pushed into the background with solutions represented by **trajectories** parameterized by solutions $(x(t), y(t))$ in the xy-plane, which is called the **phase plane**. The dependence of x and y on t is needed to tell **how** the trajectories is drawn by a solution to the equation, but the set of points visited by the solution still contains useful information and is worthy of study. Since any point on a trajectory can be chosen as $(x(0), y(0))$, the uniqueness theorem implies that **different trajectories are disjoint**.

We have previously claimed, both in lecture and in labs, that the **direction field** of a differential equation, or system of equations, gives a **picture** of the behavior of the solution. To get useful numerical accuracy, a numerical method is necessary, but we have seen many examples where the shape of the graphs of solutions could be seen in the direction field. Here is a secret: the equations in these examples were chosen very carefully to illustrate the desired conclusion. The direction field plotted by Maple usually has a cluttered appearance, with most of the action taking place in a small part of the picture. Consider the system

$$\begin{aligned}\frac{dx}{dt} &= 7x - 12y\\ \frac{dy}{dt} &= 4x - 7y\end{aligned} \tag{2}$$

The direction field of (2) is shown in Figure 1.

Since the system is **linear and homogeneous**, $\big(x(t), y(t)\big) = \big(0, 0\big)$ is a solution and all multiples of a solution are also solutions. In particular, the picture would be identical if both x and y were rescaled by the same factor. Thus, the confusion and clutter in this direction field is an **essential feature** of the equation. In spite of this, it is not too difficult to see that the origin is a **saddle point**, since some arrows seem to be moving toward the origin and others move away. Still, it would be desirable to have a less cluttered picture that retains some important features of the direction field.

Nullclines. First, recall that a direction field is generated by evaluating $F(x, y)$ and $G(x, y)$ at a grid of points and using these values at a point (x_i, y_i) in the grid to produce a **unit vector** in the direction of the

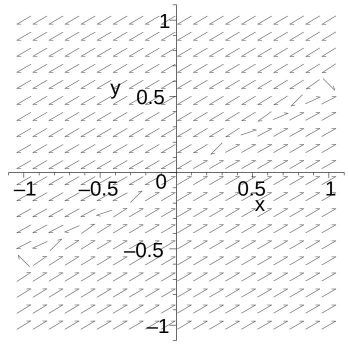

Figure 1

vector $\langle F(x_i, y_i).G(x_i, y_i)\rangle$ that will be drawn with its midpoint at (x_i, y_i). The clutter in a direction field is caused by drawing a vector at **every** point of the grid. In those regions where the direction is changing slowly, there are **too many** arrows, since they all have roughly the same direction; in those regions where the direction is changing rapidly, there are **too few** arrows, since the values at grid points don't provide a fair sample of all the directions at points of the region. The remedy is to decide on the directions to be drawn first and, then, to select points at which $\langle F(x, y).G(x, y)\rangle$ has those directions. Surprisingly, it suffices to use only **directions parallel to the axes** and only **a few** points for each direction. In particular, vertical directions are characterized by $dx/dt = 0$, so the tangent lines are vertical **precisely** at points where $F(x, y) = 0$. For many of the systems that we study, the solution to this equation is a curve that is easily drawn. Any curve on which $F(x, y) = 0$ is called a **nullcline** of the system; so is any curve on which $G(x, y) = 0$. If $G(x, y) = 0$, then the tangent lines are horizontal. In particular, this construction does not require any information about the solution to the system (1). Figure 2 shows both the direction field of (2) and the line on which the direction field is vertical.

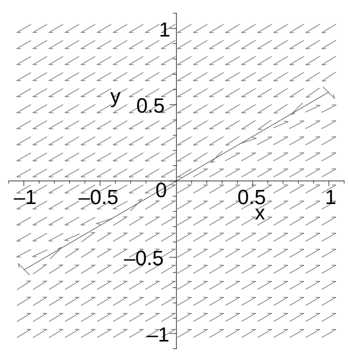

Figure 2

The direction is changing so rapidly near this line that the arrows at grid points near the line do not appear vertical. The only evidence of the validity of our claim is that the arrows above the line point to the left while those below point to the right. We do not show the nullcline designating horizontal tangents because it would be indistinguishable, at this resolution, from the one we have drawn. You can see this by noticing that (with few exceptions) arrows point down when they point to the left and up when they point to the right.

Since the words **horizontal** and **vertical** are used **in the inclusive sense**, it is possible for these nullclines to intersect. A point of intersection of the different nullclines is a point at which both $dx/dt = 0$ and $dy/dt = 0$, so that the **constant function** of t that takes this value is a solution of the system (1). Such points are called **equilibrium points** (or **critical points**) of (1). In many of our examples, $F(x, y)$ and $G(x, y)$ are polynomials, and it is possible that these polynomials factor. The curves where any of these factors are zero are the nullclines of the system. The equilibrium points are the intersections of curves determined by one factor of F and one factor of G. Before attempting to use a graph of nullclines to identify critical points, one should label the curves in some way to indicate whether they designate horizontal or vertical tangents. **Only points where curves of different designation meet are equilibrium points.**

The direction field in the phase plane consists of vectors pointing in the direction of increasing t on the a trajectory. In particular, on a nullcline where all tangent lines are horizontal, some of the arrows may point to the right while others point to the left. **Where can the direction reverse?** On such a curve where $G(x, y) = 0$, the sign of $F(x, y)$ provides the answer: the tangent points to the right if $F(x, y) > 0$ at a point and to the left if $F(x, y) < 0$. All of our studies are based on an assumption that $F(x, y)$ and $G(x, y)$ are continuous functions (extensions to equations with tame discontinuities have been mentioned, but these are dealt with by restricting to regions where the functions are continuous and splicing solutions together to create the illusion of a continuous solution to the original equation), so $F(x, y)$ can change sign on an arc only at a point where $F(x, y) = 0$. Since we are considering a curve where $G(x, y) = 0$, **these changes occur only at equilibrium points**. In a complete diagram showing the nullclines of an equation, it is only necessary to have a single arrow on each arc joining equilibrium point. Introducing these arrows and removing the direction field gives Figure 3.

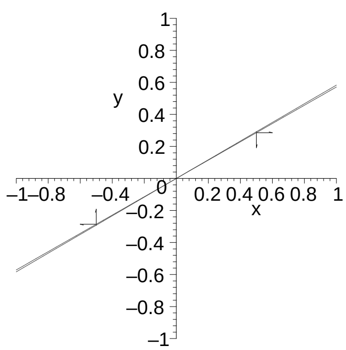

Figure 3

It is important to note that the nullclines are determined **directly** — and usually **easily** — from the equation. They are usually **not trajectories**. Indeed, the only examples of nullclines that are trajectories are of the form $x = c$ with $F(c, y) = 0$ for all y or $y = c$ with $G(x, c) = 0$ for all x.

The only evidence that both lines were drawn in Figure 3 is that there are horizontal and vertical arrows in the picture. In Figure 4, we zoom in to a small window around (0.5, 0.3).

The base of the arrow is at the point at which the direction was calculated, so we can see that the tangent points downward on the upper line and to the right on the lower line. The sign of the vertical component of the direction vector only changes on the lower line and the sign of the first component changes on the upper line. Thus, on the region **between** the lines, direction vectors point down **and** to the right. This is enough to assure that trajectories crossing the region enter along the upper line, pointing downward and exit along the lower line, pointing to the right.

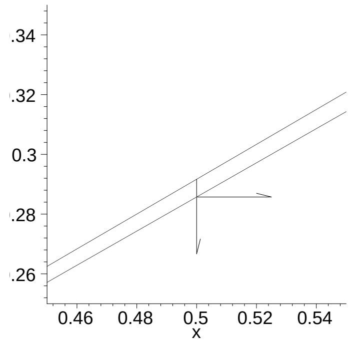

Figure 4

The competing species example Consider Example 1 from Section 9.4 of the textbook

$$\begin{aligned} \frac{dx}{dt} &= x(1 - x - y) \\ \frac{dy}{dt} &= y\left(\frac{3}{4} - y - \frac{1}{2}x\right) \end{aligned} \tag{3}$$

These equation describe two variable which, in the absence of the other, satisfy a **logistic growth** equation. In addition, there is an xy term that has a negative coefficient in both equations. Thus, as soon as both variables are positive, the growth of both is retarded. The model suggests that attention should be confined to the region where both variables are positive. Since both lines $x = 0$ and $y = 0$ satisfy (3) (a closer look shows that each is broken by an equilibrium point at the origin and one at the stable logistic value into distinct trajectories), a trajectory that contain any point in the first quadrant remains in that quadrant for all t. Although, it is possible that the equation may be of **mathematical** interest outside the first quadrant, only that part of the phase space will be considered here.

The lines $x = 0$ (which is the y-axis) and $x + y = 1$ are **nullclines** designating **vertical** tangents. Their intersection at (0, 1) is of no particular significance.

The lines $y = 0$ (which is the x-axis) and $2x + 4y = 3$ are **nullclines** designating **horizontal** tangents. Their intersection at $(3/2, 0)$ is of no particular significance. Figure 5 shows the nullclines for (3) with a segment on the lines other than the axes showing designation of that nullcline.

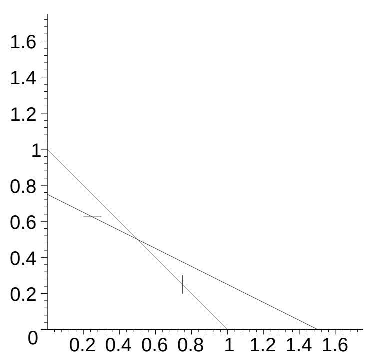

Figure 5

The equilibrium points are found by intersecting a line of the first type with one of the second. (A purely algebraic approach to finding equilibrium points, without considering nullclines would exploit factorization in exactly the same way.) This gives the following **complete list of equilibrium points**: $(0, 0)$, $(0, 3/4)$, $(1, 0)$, $(1/2, 1/2)$. Since the restriction to the x-axis is a logistic equation, both trajectories on the positive part of the x-axis point toward $(1, 0)$. This shows that the arrow at $(3/2, 0)$ points to the left. The same must be true of the segment of the **other** nullcline through that point between the x-axis and the critical point $(1, 1)$. This allows us to add arrows on each piece of a nullcline between equilibrium points. The result is shown in Figure 6.

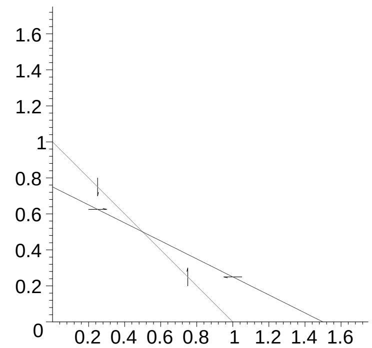

Figure 6

From Figure 6, one sees, by considering the direction of arrows on the sides of the region, that all trajectories in the region near the origin move **up and to the right**, causing them to exit the region along one of the boundaries inside the quadrant. There is **no place that a trajectory can enter this region**, so extrapolating back to $t = -\infty$ must lead to the equilibrium point $(0, 0)$. In the region where both x and y are large, each trajectory moves **down and to the left**. As in the first case, this leads to trajectories coming from infinity to one of the interior lines, where it exits in the direction of the arrow on that segment. We shall call both of these regions **quadrilaterals** when we discuss this behavior in more detail later. A similar analysis in the triangular regions shows trajectories **entering along one of the sides** and **never leaving**. The allows only the possibility that a trajectory approaches $(1/2, 1/2)$ as $t \to +\infty$.

Separatrices. If we take a point inside one of the quadrilaterals as the **initial data** $\big(x(0), y(0)\big)$ for a solution to the competing species system (3), the solution is expected to **leave the quadrilateral in finite time**. The quadrilateral is thus **cut into** a piece consisting of points where the trajectory starting at that point leaves along one side of the boundary, and a piece where the trajectory leaves along the other side (the axes can be excluded from this discussion since trajectories originating inside the region are disjoint from the trajectories along the axes). There will be some **exceptional points** for which we cannot decide which edge they lead to. A closer look at the existence and uniqueness theorems will show that **solutions depend continuously on initial conditions** (under suitable conditions that will be satisfied for all of our examples). This allows us to conclude that the this classification by trajectory-exit divides the quadrilateral into two large **connected** sets separated by another connected set. As we allow the **exit time** to go to $+\infty$, the separating set **becomes a curve**. Such a curve is called a **separatrix** (the plural is **separatrices**).

In the **triangular regions**, points are classified by where the trajectory through the point **entered** the region. Here, the axis is a limit of curves that enter along the outer edge, and all such trajectories can be traced back to the boundary in finite time. This forces the separatrix to approach the equilibrium point on the axis as $t \to -\infty$. In Figure 7, the arrows of Figure 6 are replaced by **computed solutions** with initial conditions (in the **middle** of the computed solution) on the nullclines to illustrate the division between solutions crossing different boundaries.

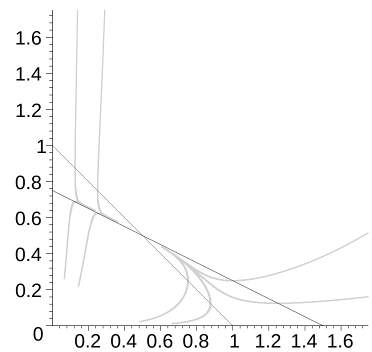

Figure 7

To summarize, equation (3) has **four special trajectories**. Each starts near a stationary point (or at infintiy) for $t \to -\infty$ and approaches the stationary point $(\frac{1}{2}, \frac{1}{2})$ as $t \to +\infty$.

Linearization. Suppose that we expand $F(x, y)$ and $G(x, y)$ in a **Taylor series** about an equilibrium point. In a case like (3), where F and G are polynomials, this can be done by substituting $x = x_0 + u$ and $y = y_0 + v$ to get polynomials in u and v. Since this substitution gives $dx/dt = du/dt$ and $dy/dt = dv/dt$, the effect of the substitution on the differential equation is to create a new system in which the equilibrium point (x_0, y_0) has become the origin in (u, v) coordinates. This process is known as **linearization**.

Because an equilibrium point was selected, the series around that point **has no constant term**. Furthermore, if u and v are small, then **all terms of degree greater than 1 are very small**. This means that the solutions of the system near the equilibrium point will resemble the solutions of the linear equation obtained by selecting only the linear terms in the series expansions of F and G. For equation (3), it is easy enough to do this at all four equilibrium points. For example, at $(0, 0)$, removing the second degree terms leaves $dx/dt = x$ and $dy/dt = 3y/4$. Our study of linear systems tells us that, even when the equation is as easy to solve as this, it is useful to write the system in matrix form as

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{3}{4} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

The matrix in this system is **immediately** seen to have positive eigenvalues. This marks the point as an **unstable node**. Since solutions of the nonlinear equation near $(0, 0)$ are close to those of this linearization, the same classification holds for the nonlinear equation. This agrees with the analysis based on nullclines, providing a check on our earlier work. However, the eigenvalues and eigenvectors of the linearization give finer information about solutions near the equilibrium point.

Even if F and G are polynomials, the **Taylor series** approach to linearization is usually preferred. Only the **linear terms** of $F(x_0 + u, y_0 + v)$ and $G(x_0 + u, y_0 + v)$ need to be retained, so any computation of terms of higher degree will be discarded (and the computation of the constant term must give zero if **evaluated correctly at a stationary point**). Furthermore, there must be a separate computation at each stationary point. It is usually better to compute the matrix

$$J(x,y) = \begin{bmatrix} \frac{\partial F(x,y)}{\partial x} & \frac{\partial F(x,y)}{\partial y} \\ \frac{\partial G(x,y)}{\partial x} & \frac{\partial G(x,y)}{\partial y} \end{bmatrix}$$

and evaluate it at each equilibrium point to get the linearization at that point. This is in **the best spirit of the calculus**: a derivative is found **as a function** and evaluated at special places to get the quantities need in an application.

For (3), we have

$$J(x,y) = \begin{bmatrix} 1 - 2x - y & -x \\ -y/2 & 3/4 - 2y - x/2 \end{bmatrix}$$

Substituting the coordinates of each critical point int this expression gives the matrices found in the textbook.

A linear inhomogeneous example. Consider this modification of (2)

$$\begin{aligned} \frac{dx}{dt} &= 7x - 12y - 9 \\ \frac{dy}{dt} &= 4x - 7y - 5 \end{aligned} \tag{4}$$

We have met one method to solve this system: use **undetermined coefficients** to find a **particular solution** in which x and y are **constant functions**, and add the general solution of the homogeneous equation (2) to obtain a **general solution**. Specialization of the approach used for nonlinear equations leads to similar computations, although the interpretation is different. This approach begins by finding the **equilibrium point**, which is the simultaneous solution of $7x - 12y - 9 = 0$ and $4x - 7y - 5 = 0$. The solution is $(x, y) = (3, 1)$. Since there are no terms of degree greater than one in $F(x, y)$ or $G(x, y)$, the matrix J is a constant function of x and y, so the expression of (4) in local coordinates at (3, 1) looks exactly like (2).

Polar nullclines. In our analysis of Figure 4, we were able to use the direction of the vector field on the nullclines to give detailed information about a **very small** part of the plane, but there were large sectors where we only knew that trajectories entered across the nullclines and could not reach the origin. This led to the conclusion that they must escape to infinity somehow, but it appeared that it would be necessary to solve the equation to learn more. The equation is easy to solve in this case since (2) is a linear system, but it would be useful to have some methods that could be applied in the nonlinear case.

If **polar coordinates** (r, θ) are introduced by the usual equations $x = r\cos\theta$ and $y = r\sin\theta$, then

$$\begin{aligned} r\frac{dr}{dt} &= x\frac{dx}{dt} + y\frac{dy}{dt} \\ r^2\frac{d\theta}{dt} &= x\frac{dy}{dt} - y\frac{dx}{dt} \end{aligned} \tag{5}$$

and the right sides of these equations can be found — as functions of x and y — from the original system (1). In the **important special case** in which $F(x, y)$ and $G(x, y)$ are **polynomials**, these new expressions are also polynomials. The curves where the right sides of two equations in (5) are zero ape used to extend the usefulness of nullclines, but do not seem to have a standard name. In order to talk about their role, we propose the use of the name **polar nullcline** for curves on which either dr/dt or $d\theta/dt$ is zero. Their geometric significance should be clear: $dr/dt = 0$ shows where trajectories are closest to (or farthest from) the origin, and $d\theta/dt = 0$ separates regions of counterclockwise motion about the origin, where $d\theta/dt > 0$, from regions of clockwise motion.

If the origin is an equilibrium point, the series expansion of the expressions defining the polar nullclines will have **no constant term** and **no linear terms**. Usually, the quadratic terms will completely determine the behavior of the polar nullclines near the origin. The directions in which the quadratic part of $d\theta/dt$ is zero give the **tangential directions** of that polar nullcline, which are **the directions of the eigenvectors** of the linearization at the origin. This characterization of eigenvectors may also be derived by equating the ratio of components in $\mathbf{v}$ and $M\mathbf{v}$.

If $F(x, y)$ and $G(x, y)$ are polynomials, it is useful to consider the terms of **largest** degree. These are the terms that dominate when x and y are large, so they can be used to identify the behavior of trajectories "at infinity".

If there are several equilibrium points, it might be useful to consider all translations coordinates that make one of them the origin and set up polar coordinates in all of those systems. There will be different polar nullclines for each of these choices. The behavior near the origins of those systems must reflect the different linearizations at the different equilibrium points. However, the behavior at infinity will be the same for all of these choices, since translations of coordinates only change the terms of degree less than the maximal degree of a polynomial.

There are still some cases in which this approach doesn't quite work. A simple example in which these methods work well enough to get close to uncovering all qualitative features of the system is

$$\begin{aligned} \frac{dx}{dt} &= y - x \\ \frac{dy}{dt} &= x^2 - y \end{aligned} \tag{6}$$

Figure 8 shows nullclines and some trajectories for this system.

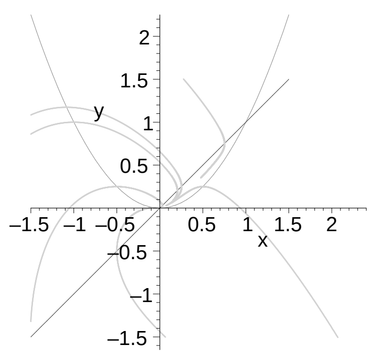

Figure 8