

3 Differential equations

3.1 Linear Differential Equations of First Order

A first order differential equation is an equation of the form

$$\frac{dx(t)}{dt} \equiv \dot{x}(t) = f(x, t). \quad (1)$$

As noted above, there will in general be a whole class of functions $x(t; c)$ (parameterized by c) that satisfies the differential equation (1). We need more information, like an initial condition for $x(t_0)$, to pin down the solution exactly.

Result 10 *Given that f is continuous and has continuous first derivatives, there is going to be a one function $x(t)$ that satisfies (1) and an initial condition.*

3.1.1 The simplest case

If f is independent of x , the solution is trivial. Then since

$$\dot{x}(t) = f(t), \quad (2)$$

the class of functions satisfying this is a primitive function of f plus any constant, i.e.,

$$x(t) = \int_{t_0}^t f(s) ds + \tilde{c} = F(t) - F(t_0) + \tilde{c} \quad (3)$$

where $F(t)$ is any primitive of f . Note that $F(t_0)$ is a constant. We can thus merge the two constants, defining $c \equiv \tilde{c} - F(t_0)$ and write

$$x(t) = F(t) + c \quad (4)$$

Certainly, for all c , (4) satisfies (1). For example, if

$$f(t) = e^{at}, \quad (5)$$

$$F(t) = \frac{e^{at}}{a} \quad (6)$$

is a primitive for $f(t)$. The arbitrary constant is pinned down with some other piece of information. So, if we want to find $x(t)$ and we know the value

of $x(0)$, we get

$$x(t) = F(t) + c \quad (7)$$

$$x(0) = F(0) + c \quad (8)$$

$$\rightarrow c = x(0) - F(0) \quad (9)$$

$$x(t) = F(t) + x(0) - F(0). \quad (10)$$

Note that there is only one degree of freedom in the constants c and t_0 . Choosing another t_0 simply means that the constant c has to be chosen in another way. Thus, one piece of information is sufficient to pin down the solution exactly.

3.1.2 A note on notation

Above, we have used the proper notation of an integral, where both the lower and upper limits and the dummy variable to integrate over have separate names and are all written out. Often, a more sloppy notation is used, for an arbitrary t_0 we write

$$\int_{t_0}^t f(s) ds = \int f(t) dt = F(t) \quad (11)$$

where it is understood that $\int f(t) dt$ is a (any) primitive of $f(t)$. This notation, called an indefinite integral, saves on variables, but can be confusing since t is used as both the variable to integrate over and as upper integration limit. Nevertheless, I will follow ordinary practice and use it below. Using this notation in (3), we rewrite (3)

$$x(t) = \int f(t) dt + c. \quad (12)$$

3.1.3 A little bit more complicated

Very often, a solution to a more complicated differential equation is derived by transforming the original equation into something that has the form of (2). Linear first order differential with constant coefficients equations can be solved directly using such a transformation. Consider

$$\dot{y}(t) + ry(t) = q. \quad (13)$$

In this case, we multiply both sides by e^{rt} (often called the integrating factor). After doing that, note that the LHS becomes

$$e^{rt}(\dot{y}(t) + ry(t)) = \frac{d(e^{rt}y(t))}{dt}. \quad (14)$$

Thus, thinking of $e^{rt}y(t)$ as simply a function of t , as $x(t)$ in (2), we get a LHS that is the time derivative of a known function of t and the RHS is also a function *only* of t . Let's set $t_0 = 0$, then the solution is found as in (3).

$$\frac{d(e^{rt}y(t))}{dt} = e^{rt}q \quad (15)$$

$$e^{rt}y(t) = \int_0^t e^{rs}q ds + c = \left[\frac{e^{rs}q}{r} \right]_0^t + c \quad (16)$$

$$= \frac{q}{r} (e^{rt} - 1) + c \quad (17)$$

$$\rightarrow y(t) = e^{-rt} \left(\frac{q}{r} (e^{rt} - 1) + c \right) = \frac{q}{r} + \left(c - \frac{q}{r} \right) e^{-rt}. \quad (18)$$

If we know $y(t)$ at some point in time, e.g. at $t = 0$.

$$y(0) = \frac{q}{r} + \left(c - \frac{q}{r} \right) = c \quad (19)$$

$$\rightarrow y(t) = \frac{q}{r} + \left(y(0) - \frac{q}{r} \right) e^{-rt} \quad (20)$$

Let's compare this to the case when we use the indefinite integral. Then, we have

$$\frac{d(e^{rt}y(t))}{dt} = e^{rt}q \quad (21)$$

$$e^{rt}y(t) = \int e^{rt}q dt + c = \frac{e^{rt}q}{r} + c \quad (22)$$

$$= \frac{q}{r} e^{rt} + c \quad (23)$$

$$\rightarrow y(t) = \frac{q}{r} + ce^{-rt}. \quad (24)$$

$$y(0) = \frac{q}{r} + c \rightarrow c = \left(y(0) - \frac{q}{r} \right) \quad (25)$$

$$\rightarrow y(t) = \frac{q}{r} + \left(y(0) - \frac{q}{r} \right) e^{-rt} \quad (26)$$

As you see, although the integration constants are different, the result is the same. The difference is that the constant submerges the primitive evaluated at the lower integration limit.

3.1.4 Variable exogenous parts

Let us generalize (13) by assuming that $q = q(t)$, a function of time. Consider an arbitrage condition where $y(t)$ is an asset price, with dividends $q(t)$ per

unit of time. Not to loose track of things, we use the notation with integration limits, rather than indefinite integrals.

Holding the asset should yield no excess return, i.e.,

$$\dot{y}(t) + q(t) = ry(t) \quad (27)$$

$$\dot{y}(t) - ry(t) = -q(t) \quad (28)$$

$$\Rightarrow e^{-rt}y(t) = - \int_{t_0}^t e^{-rs}q(s) ds + c \quad (29)$$

Often, economic theory let us use a so called "No-Ponzi" condition (typically requiring $r > 0$) . Here this gives a limiting end condition, implying that

$$\lim_{t \rightarrow \infty} e^{-rt}y(t) = 0. \quad (30)$$

Using this, we get

$$c = \lim_{t \rightarrow \infty} \int_{t_0}^t e^{-rs}q(s) ds \equiv \int_{t_0}^{\infty} e^{-rs}q(s) ds, \quad (31)$$

giving

$$e^{-rt}y(t) = \int_{t_0}^{\infty} e^{-rs}q(s) ds - \int_{t_0}^t e^{-rs}q(s) ds \quad (32)$$

$$= \int_t^{\infty} e^{-rs}q(s) ds \quad (33)$$

$$y(t) = e^{rt} \int_t^{\infty} e^{-rs}q(s) ds = \int_t^{\infty} e^{-r(s-t)}q(s) ds, \quad (34)$$

i.e., that no arbitrage and "No-Ponzi" implies that the price must equal the discounted sum of future dividends.

3.1.5 Variable coefficients

If also the coefficient on the endogenous variable ($y(t)$) is varying over time, for example, consider the no-arbitrage equation for an asset price $y(t)$ with dividends $q(t)$ and interest rate $r(t)$

$$\dot{y}(t) + q(t) = r(t)y(t) \Rightarrow \dot{y}(t) - r(t)y(t) = -q(t). \quad (35)$$

Then, we need a more general integrating factor to make the LHS into the time differential of a known function. Here the integrating factor is

$$e^{-\int_{t_0}^t r(s)ds}. \quad (36)$$

with

$$\frac{de^{-\int_{t_0}^t r(s)ds}}{dt} = -r(t) e^{-\int_{t_0}^t r(s)ds}, \quad (37)$$

alternatively expressed as

$$\frac{de^{-\int r(t)dt}}{dt} = -r(t) e^{-\int r(t)dt}.$$

Approximating the integral as a sum of rectangles with base Δt as in section 2.2, (which is exact if $r(t)$ were piecewise constant), and defining $r(t_0 + s\Delta t) \equiv r_{t_s}$, for the integers $s \in \{0, S\}$, $S = (t-t_0)/\Delta t$, the integrating factor, can be written

$$e^{-\int_{t_0}^t r(s)ds} \approx e^{-r_0\Delta t} e^{-r_1\Delta t} \dots e^{-r_S\Delta t} \approx \left(\frac{1}{1+r_1\Delta t}\right) \left(\frac{1}{1+r_1\Delta t}\right) \dots \left(\frac{1}{1+r_S\Delta t}\right), \quad (38)$$

i.e., it is product of all short run discount factors between t_0 and t . To save on notation, this product is denoted as

$$e^{-\int_{t_0}^t r(s)ds} \equiv R(t; t_0), \quad (39)$$

or if the starting point is suppressed as $R(t)$. This has a straightforward interpretation. Suppose the variable discount rate is given by $r(s)$, then, the the discounted value of a payment $y(t)$ at t seen from t_0 is $R(t; t_0) y(t)$. Using, the integrating factor, we get

$$R(t) (\dot{y}(t) - r(t) y(t)) = -R(t) q(t) \quad (40)$$

$$\frac{dR(t) y(t)}{dt} = -R(t) q(t), \rightarrow \quad (41)$$

$$R(t) y(t) = -\int_{t_0}^t R(s) q(s) ds + c \quad (42)$$

Suppose again, $\lim_{t \rightarrow \infty} R(t) y(t) = 0$, implying $c = \int_{t_0}^{\infty} R(s) q(s) ds$. Then, noting that $R(t)^{-1} R(s) = e^{\int_{t_0}^t r(v)dv - \int_{t_0}^s r(v)dv} = e^{-\int_t^s r(v)dv} = R(s; t)$ we have

$$R(t) y(t) = -\int_{t_0}^t R(s) q(s) ds + \int_{t_0}^{\infty} R(s) q(s) ds = \int_t^{\infty} R(s) q(s) ds \quad (43)$$

$$y(t) = \int_t^{\infty} R(t)^{-1} R(s) q(s) ds = \int_t^{\infty} e^{-\int_t^s r(v)dv} q(s) ds, \quad (44)$$

i.e., that the asset price equals the discounted value of future dividends.

Another example, consider money on a bank account with variable interest rate and deposits $q(t)$, then

$$\dot{y}(t) = r(t)y(t) + q(t) \quad (45)$$

$$\dot{y}(t) - r(t)y(t) = q(t) \quad (46)$$

$$\frac{dR(t, t_0)y(t)}{dt} = R(t, t_0)q(t) \quad (47)$$

$$\rightarrow R(t, t_0)y(t) = \int_{t_0}^t R(s, t_0)q(s) + c \quad (48)$$

$$= \int_{t_0}^t R(t, t_0)^{-1}R(s, t_0)q(s) + R(t, t_0)^{-1}c \quad (49)$$

Since $s \leq t$ here, the term $R(t, t_0)^{-1}R(s, t_0)$ is more conveniently¹ written $e^{\int_s^t r(v)dv}$. Using also an initial condition, $y(t_0)$, we have

$$y(t) = \int_{t_0}^t e^{\int_s^t r(v)dv}q(s) + e^{\int_{t_0}^t r(v)dv}y(t_0), \quad (50)$$

where the first term is the period t value of all deposits up until t and the second term is the period t value of the initial amount on the bank.

3.1.6 Separating variables

Sometimes, we can write a differential equation such that the LHS only contains functions of x and \dot{x} and the RHS only a function of t . For example

$$\dot{x}(t) = \frac{h(t)}{g(x)} \quad (51)$$

$$\dot{x}(t)g(x) = h(t). \quad (52)$$

In this case, we can use the following trick. Let $G(x)$ be any primitive of $g(x)$, i.e., $\frac{dG(x)}{dx} = g(x)$. Then,

$$\dot{x}(t)g(x(t)) = \frac{dG(x(t))}{dt} = h(t) \quad (53)$$

$$G(x) = \int h(s)ds + c \quad (54)$$

We can then recover x by inverting G . An example;

¹But remember $-\int_t^s r(v)dv = \int_s^t r(v)dv$.

$$\dot{x}(t) = (x(t)t)^2 \quad (55)$$

$$\dot{x}(t)x(t)^{-2} = t^2. \quad (56)$$

Now let $g(x) = x^{-2}$ implying $G(x) = -x^{-1}$. Then since $\frac{dG(x(t))}{dt} = \dot{x}(t)g(x(t)) = \dot{x}(t)x(t)^{-2}$, we have

$$\frac{dG(x(t))}{dt} = t^2 \quad (57)$$

$$G(x(t)) = -x(t)^{-1} = \int t^2 dt + c = \frac{t^3}{3} + c. \quad (58)$$

Using the fact that $z = G(x) = -x^{-1} \Rightarrow x = -z^{-1}$, so the inverse function is given by $G(z) = -z^{-1}$, we get

$$x(t) = G^{-1}\left(\frac{t^3}{3} + c\right) = -\left(\frac{t^3}{3} + c\right)^{-1}.$$

3.2 Linear differential equations of higher order

3.2.1 Linear second order differential equations

A linear second order differential equation has the form

$$\ddot{y}(t) + p(t)\dot{y}(t) + q(t)y(t) = R(t) \quad (59)$$

This cannot be solved directly by transformations in the simple way we did with first order equations. Instead we use a more general method (that would have worked above also). First some definitions

Definition 1 *The homogeneous part of differential equation is achieved by setting all exogenous variables, including constants, to zero.*

Definition 2 *Two functions $y_1(t)$ and $y_2(t)$ are linearly independent in a region Ω iff there is no $c_1, c_2 \neq \{0, 0\}$ s.t.*

$$c_1 y_1(t) = c_2 y_2(t) \quad \forall t \in \Omega. \quad (60)$$

Result 11 *The general solution (the complete set of solutions) to a differential equation is the general solution to the homogeneous part plus any particular solution to the complete equation.*

The general solution of the homogeneous part of a second order linear differential equation can be expressed as where $c_1 y_1(t) + c_2 y_2(t)$, where $y_1(t)$ and $y_2(t)$ are two linearly independent particular solutions to the homogeneous equations and c_1 and c_2 are some arbitrary integrations constants.

3.2.2 Homogeneous Equations with Constant Coefficients

Consider the homogeneous part of a differential equation, given by

$$\ddot{y}(t) + p\dot{y}(t) + qy(t) = 0. \quad (61)$$

To solve this equation, we first define the *characteristic equation*, for this equation;

$$r^2 + pr + q = 0. \quad (62)$$

Since this is a second-order polynomial, it has two roots

$$r_{1,2} = -\frac{1}{2}p \pm \frac{1}{2}\sqrt{(p^2 - 4q)}. \quad (63)$$

Now, it is straightforward to see that $e^{r_1 t}$ and $e^{r_2 t}$ both are solutions to the homogeneous equation, by noting that for $i \in \{1, 2\}$

$$\frac{de^{r_i t}}{dt} = r_i e^{r_i t} \quad (64)$$

$$\frac{d^2 e^{r_i t}}{dt^2} = r_i^2 e^{r_i t} \rightarrow \quad (65)$$

$$\frac{d^2 e^{r_i t}}{dt^2} + p \frac{de^{r_i t}}{dt} + q e^{r_i t} = (r_i^2 + pr_i + q) e^{r_i t} = 0. \quad (66)$$

So then, result 11 tells us that the general solution to the homogeneous equation is

$$y_h(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}, \quad (67)$$

provided $e^{r_1 t}$ and $e^{r_2 t}$ are linearly independent. It is easy to verify that this is the case, if and only if $r_1 \neq r_2$.

3.2.3 Complex roots

Complex roots pose no particular difficulty, we simply have to recall that for any real number b

$$e^{bi} = \cos(b) + i \sin(b), \quad (68)$$

$$e^{-bi} = \cos(b) - i \sin(b), \quad (69)$$

yielding for the complex roots $r_{1,2} = a \pm bi$,

$$c_1 e^{(a+bi)t} + c_2 e^{(a-bi)t} = \quad (70)$$

$$c_1 e^{at} (\cos(bt) + i \sin(bt)) + c_2 e^{at} (\cos(bt) - i \sin(bt)) \quad (71)$$

$$= e^{at} ((c_1 + c_2) \cos(bt) + (c_1 - c_2) i \sin(bt)) \quad (72)$$

Note here, that the constants c_1 and c_2 may be chosen from full set of complex number. Defining

$$c_1 + c_2 \equiv \bar{c}_1, \quad (73)$$

$$(c_1 - c_2) i \equiv \bar{c}_2, \quad (74)$$

It turns out that for any \bar{c}_1 and \bar{c}_2 on the real line, we can find c_1 and c_2 satisfying this definition. Since we, at least in economics, are only interested in solutions in the real space, we can use the restricted set of constants satisfying \bar{c}_1 and \bar{c}_2 being on the real line. We can then write the general solution *in the real space* as

$$y_h(t) = e^{at} (\bar{c}_1 \cos(bt) + \bar{c}_2 \sin(bt)). \quad (75)$$

3.2.4 Repeated roots

The general solution to the homogenous equation in the case when the roots are repeated, i.e., $r_1 = r_2 \equiv r$ is

$$y_h(t) = c_1 e^{rt} + c_2 t e^{rt}. \quad (76)$$

Convince yourself that they are linearly independent and check that they are both solutions if the roots are repeated but not otherwise!

3.2.5 Non-Homogeneous Equations with Constant Coefficients

Relying on result 11, the only added problem when we have a non-homogeneous equation is that we have to find *one* particular solution to the complete equation. Consider

$$\ddot{y}(t) + p\dot{y}(t) + qy(t) = R(t). \quad (77)$$

Typically we guess a form of this solution and then use the *method of undetermined coefficients*. Often a good guess is a solution of a form similar to $R(t)$, e.g., if it is a polynomial of degree n we guess on a general n 'th degree polynomial with unknown coefficients. The simplest example is if $R(t)$ equals a constant R , we then guess on a constant, $y_p(t) = y^{ss}$, i.e., a *steady state*, in which case $\ddot{y}(t)$ and $\dot{y}(t)$ both are zero. To satisfy the differential equation,

$$qy^{ss} = R \quad (78)$$

$$y^{ss} = \frac{R}{q}. \quad (79)$$

Another example is

$$\ddot{y}(t) - 2\dot{y}(t) + y(t) = 3t^2 + t, \quad (80)$$

in which case we guess

$$y_p(t) = At^2 + Bt + C, \quad (81)$$

for some, yet *undetermined coefficients* A, B and C . We then solve for these constants by substituting into the differential equation

$$2A - 2(2At + B) + At^2 + Bt + C = 3t^2 + t. \quad (82)$$

For this to hold for each t , we need

$$A = 3 \quad (83)$$

$$-4A + B = 1 \quad (84)$$

$$2A - 2B + C = 0 \quad (85)$$

yielding $A = 3, B = 13$, and $C = 20$. So a particular solution is

$$y_p(t) = 3t^2 + 13t + 20. \quad (86)$$

The characteristic equation is

$$r^2 - 2r + 1 = (r - 1)^2 \rightarrow r_{1,2} = 1. \quad (87)$$

So the general solution is

$$y(t) = c_1 e^t + c_2 t e^t + 3t^2 + 13t + 20. \quad (88)$$

3.2.6 Linear nth Order Differential Equations

Consider the nth order differential equation

$$y^n(t) + P_1 y^{n-1}(t) + \dots + P_n y(t) = R(t), \quad (89)$$

where

$$y^n(t) \equiv \frac{d^n y(t)}{dt^n}. \quad (90)$$

The solution technique is here exactly analogous to the second order case. First, some definitions

Definition 3 For any differentiable function $y(t)$, the differential operator D is defined as

$$Dy(t) \equiv \frac{dy(t)}{dt}, \quad (91)$$

satisfying $D^n y(t) \equiv \frac{d^n y(t)}{dt^n}$.

Using this definition, and defining the following polynomial, letting then write

$$P(r) \equiv r^n + P_1 r^{n-1} + \dots + P_n, \quad (92)$$

we have

$$P(D)y(t) = (D^n + P_1 D^{n-1} + \dots + P_n)y(t) \quad (93)$$

$$= D^n y(t) + P_1 D^{n-1} y(t) + \dots + P_n y(t) \quad (94)$$

so we can write (89) in the condensed form,

$$P(D)y(t) = R(t) \quad (95)$$

and the characteristic equation is

$$P(r) = 0, \quad (96)$$

with roots $r_{1,\dots,n}$

The general solution to the homogenous part is then the sum of the n solutions corresponding to each of n roots. The only thing to note is that if one root, r , is repeated $k \geq 2$ times, the solutions corresponding to this root is given by

$$c_1 e^{rt} + c_2 t e^{rt} + \dots + c_k t^{k-1} e^{rt}. \quad (97)$$

Repeated complex roots are handled the same way. Say the root $a \pm bi$ is repeated in k pairs. Their contribution to the general solution is given by

$$e^{at} (c_1 \cos(bt) + c_2 \sin(bt) + t(c_3 \cos(bt) + c_4 \sin(bt))) + \dots \quad (98)$$

$$+ e^{at} t^{k-1} (c_{2k-1} \cos(bt) + c_{2k} \sin(bt)). \quad (99)$$

A particular solution to the complete equations can often be solved by the method of undetermined coefficients.

3.3 Stability

From the solutions to the differential equations we have seen we find that the terms corresponding to roots that have positive real parts (unstable, or explosive roots) tend to explode as t goes to infinity. This means that also the solution explodes unless the corresponding integration constants are zero. Terms with roots that have negative real parts (stable roots), on the other hand, always converge to zero.

If all roots are *strictly* negative for a differential equation, $P(D)y(t) = r$, the system converges to a unique point as time goes to infinity, wherever it starts. This point, is often called a *sink* and the system is defined as globally stable.

If all roots are strictly positive the system is globally unstable, but there is still a steady-state, this is sometimes called the *origin*. The reason for this is that if the system is unstable in all dimensions when time goes forward, it will be *stable* in all dimensions if time goes in reverse. Starting from *any* point and going *backward*, one reaches, in the limit the steady state, i.e., it is the origin of all paths.

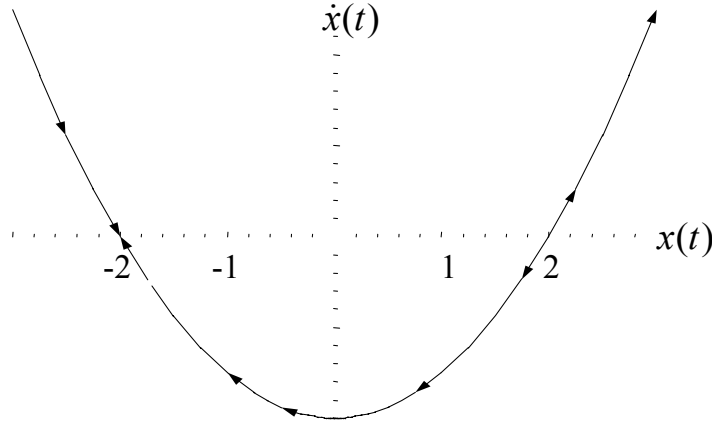
If a system has both stable and unstable roots, it is called saddle-path stable. Then, in some sub-dimensions it is stable.

3.3.1 Non-linear equations and local stability

Look at the nonlinear differential equation

$$\dot{x}(t) = x(t)^2 - 4. \tag{100}$$

Although we have not learned how to solve such an equation, we can say something about it. Let us plot the relation $x(t) \rightarrow \dot{x}(t)$



$$\dot{x}(t) = x(t)^2 - 4$$

We see that $x(t) = 2$ and $x(t) = -2$ are stationary points. We also see that $x(t) = -2$ is locally stable. In the region $[-\infty, 2)$ x converge to $x = -2$. 2 is an unstable stationary point and in the region $(2, \infty]$, x explodes over time. In a plot $x(t) \rightarrow \dot{x}(t)$, local stability is equivalent to a negative slope at the stationary point.

3.4 Systems of Linear First Order Differential Equations

Consider the following system of two first order differential equations

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} p_1(t) \\ p_2(t) \end{bmatrix} \quad (101)$$

$$\dot{\mathbf{y}}(t) = \mathbf{A}\mathbf{y}(t) + \mathbf{p}(t) \quad (102)$$

As in the one equations case we start by finding the general solutions to the homogeneous part. This plus some particular solution is the general solution to the complete system.

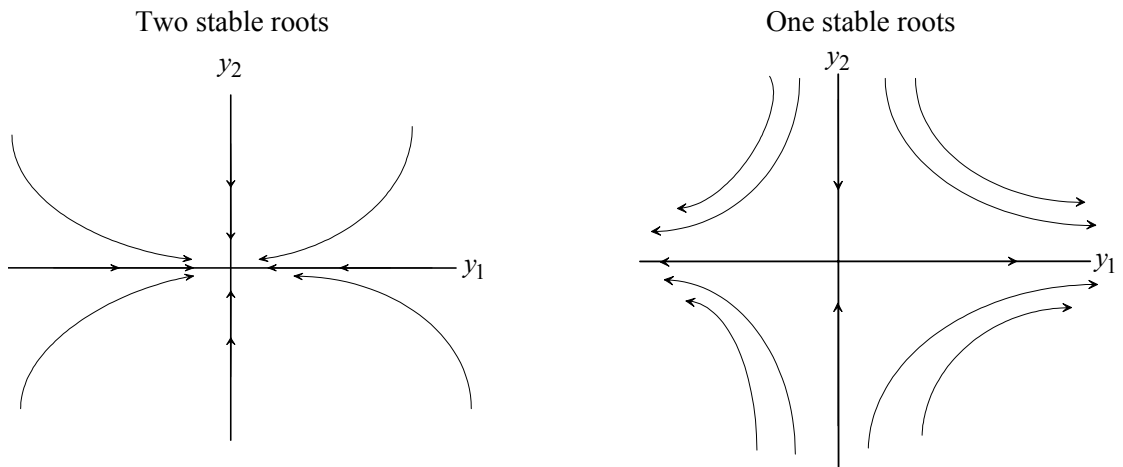
If the off diagonal terms are zero the solution to the homogeneous part is trivial, since there is no interdependency between the equations

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \quad (103)$$

$$\rightarrow \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} e^{a_{11}t} & 0 \\ 0 & e^{a_{22}t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \quad (104)$$

The system in (103) has an important property, time has no direct effect on the law-of-motion. Given knowledge of $y_1(t)$ and $y_2(t)$, $\dot{y}_1(t)$ and $\dot{y}_2(t)$ are fully determined, *regardless of t* . A system that has this property is called *autonomous*. The system (101) is not an autonomous system, unless $\mathbf{p}(t)$ is constant. Note that the homogeneous solution is always autonomous, given, of course, that the parameters are constant.

The behavior of an autonomous system can be depicted in a graph, a phase diagram. We see in the phase diagram that if the roots are stable, i.e., negative, the homogeneous part always goes to zero as t goes to infinity. With only one root stable, there is just one stable path.



The fact that it is trivial to solve a diagonal system, suggests a way of finding the solution to the homogeneous part of (101). Suppose we can make a transformation of the variables so that the transformed system is diagonal. Start by defining the new set of variables

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \equiv \mathbf{B} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \rightarrow \quad (105)$$

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} \equiv \mathbf{B} \begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{bmatrix} = \mathbf{BA} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \mathbf{BAB}^{-1} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad (106)$$

If we can find a \mathbf{B} such that \mathbf{BAB}^{-1} is diagonal we are half way. The solutions for \mathbf{x} is then

$$(107)$$

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} e^{r_1 t} & 0 \\ 0 & e^{r_2 t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad (108)$$

where $r_{1,2}$ are the diagonal terms of the matrix \mathbf{BAB}^{-1} . The solution for \mathbf{y} then follows from the definition of \mathbf{x}

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \mathbf{B}^{-1} \begin{bmatrix} e^{r_1 t} & 0 \\ 0 & e^{r_2 t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \quad (109)$$

From linear algebra we know that \mathbf{B}^{-1} is the matrix of eigenvectors of \mathbf{A} and that the diagonal terms of \mathbf{BAB}^{-1} are the corresponding eigenvalues. The eigenvalues are given by the characteristic equation of \mathbf{A}

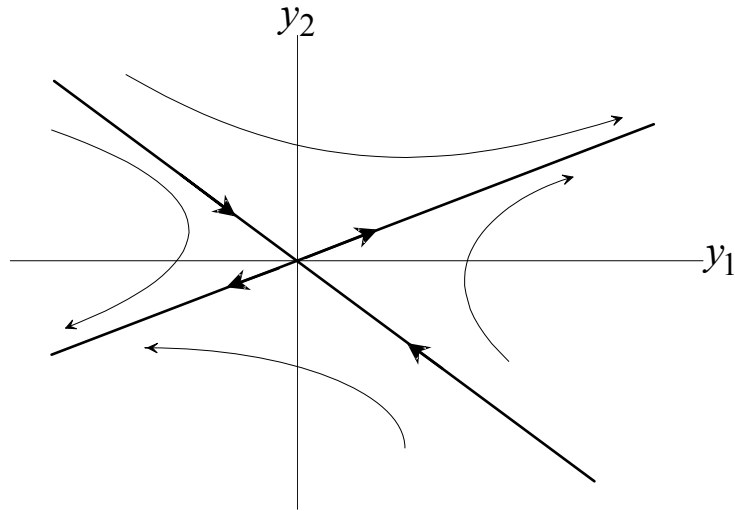
$$\left| \begin{bmatrix} a_{11} - r & a_{12} \\ a_{21} & a_{22} - r \end{bmatrix} \right| = 0 \quad (110)$$

$$\rightarrow (a_{11} - r)(a_{22} - r) - a_{12}a_{21} = 0 \quad (111)$$

$$r^2 - rTr(\mathbf{A}) + |\mathbf{A}| = 0, \quad (112)$$

where $Tr(\mathbf{A})$ is the trace of \mathbf{A} . The only crux is that we need the roots to be distinct, otherwise \mathbf{B}^{-1} is not always invertible. Distinct root implies that \mathbf{B}^{-1} is invertible. (If \mathbf{A} is symmetric \mathbf{B}^{-1} is also invertible.)

Let us draw a phase diagram with the eigenvectors of \mathbf{A} . The dynamic system behaves as the diagonal one in but the eigenvectors have replaced the standard, orthogonal axes.



One stable root

What is remaining is to find a particular solution of the complete system. One way is here to use the method of undetermined coefficients. Sometimes, we can find a steady state of the system, i.e., a point where the time derivatives are all zero. This is easy if the exogenous part is constant. We then set the differential equal to zero so

$$0 = \mathbf{A}\mathbf{y}(t) + \mathbf{p} \quad (113)$$

$$\rightarrow \mathbf{y}_p(t) = \mathbf{y}^{ss} = -\mathbf{A}^{-1}\mathbf{p}. \quad (114)$$

Given that we know the value of $\mathbf{y}(0)$ we can now give the general solution. The formula is given in matrix form and is valid for any dimension of the system. First define

$$\mathbf{r}(t) \equiv \begin{bmatrix} e^{r_1 t} & 0 & \dots & 0 \\ 0 & e^{r_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{r_n t} \end{bmatrix}, \quad (115)$$

then we have

$$\mathbf{y}(t) = \mathbf{B}^{-1}\mathbf{x}(t) + \mathbf{y}^{ss} = \mathbf{B}^{-1}\mathbf{r}(t)\mathbf{c} + \mathbf{y}^{ss} \quad (116)$$

$$\mathbf{y}(0) = \mathbf{B}^{-1}\mathbf{x}(0) + \mathbf{y}^{ss} = \mathbf{B}^{-1}\mathbf{c} + \mathbf{y}^{ss} \quad (117)$$

$$\mathbf{c} = \mathbf{B}(\mathbf{y}(0) - \mathbf{y}^{ss}) \quad (118)$$

$$\rightarrow \mathbf{y}(t) = \mathbf{B}^{-1}\mathbf{r}(t)\mathbf{B}(\mathbf{y}(0) - \mathbf{y}^{ss}) + \mathbf{y}^{ss}. \quad (119)$$

The method outlined above works also in the case of complex roots of the characteristic equation. If the roots are $a \pm bi$ we have

$$\mathbf{y}(t) = \mathbf{B}^{-1} \begin{bmatrix} e^{(a+bi)t} & 0 \\ 0 & e^{(a-bi)t} \end{bmatrix} \mathbf{c} + \mathbf{y}^{ss} \quad (120)$$

Example;

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (121)$$

$$r_{1,2} = -1 \pm i \quad (122)$$

$$\mathbf{B}^{-1} = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}. \quad (123)$$

So,

$$\rightarrow \mathbf{y}(t) = \quad (124)$$

$$\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{-t}(\cos t + i \sin t) & 0 \\ 0 & e^{-t}(\cos t - i \sin t) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \mathbf{y}^{ss} \quad (125)$$

$$= e^{-t} \begin{bmatrix} i(c_1 - c_2) \cos t - (c_1 + c_2) \sin t \\ (c_1 + c_2) \cos t + i(c_1 - c_2) \sin t \end{bmatrix} + \mathbf{y}^{ss} \quad (126)$$

$$= e^{-t} \begin{bmatrix} \tilde{c}_1 \cos t + \tilde{c}_2 \sin t \\ -\tilde{c}_2 \cos t + \tilde{c}_1 \sin t \end{bmatrix} + \mathbf{y}^{ss} \quad (127)$$

The steady-state is found from,

$$\mathbf{0} = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (128)$$

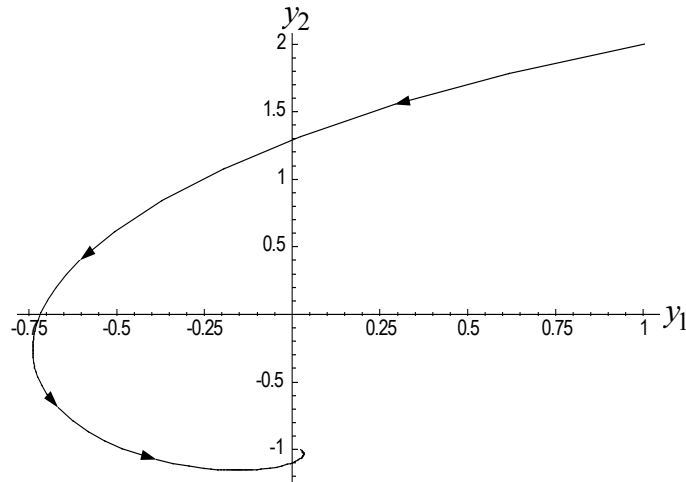
$$\begin{bmatrix} y_1^{ss} \\ y_2^{ss} \end{bmatrix} = - \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (129)$$

If we know $\mathbf{y}(0)$, and using $\cos 0 = 1$ and $\sin 0 = 0$, we get

$$\begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} \tilde{c}_1 \\ -\tilde{c}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (130)$$

$$\rightarrow \mathbf{y}(t) = e^{-t} \begin{bmatrix} y_1(0) \cos t - (y_2(0) - 1) \sin t \\ (y_2(0) - 1) \cos t + y_1(0) \sin t \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (131)$$

A phase-diagram of this system is an inward spiral.



(132)

How would it look like if the real part of the root was -1 or $+1$?

3.4.1 Equivalent Systems

In the repeated root case the matrix of the eigenvectors may be singular, so that we cannot find \mathbf{B}^{-1} . Then we use the method of equivalent systems.

A linear n th order differential equation is equivalent to a system of n first order differential equations. Consider,

$$\ddot{y}(t) + a_1\dot{y}(t) + a_2y(t) + a_3y(t) = R(t) \quad (133)$$

We can transform this into a system of first order differential equation by defining

$$\dot{y}(t) \equiv x_1(t), \quad (134)$$

$$\ddot{y}(t) \equiv x_2(t) = \dot{x}_1(t), \quad (135)$$

$$\ddot{\dot{y}}(t) = \dot{x}_2(t), \quad (136)$$

giving

$$\begin{bmatrix} \dot{y}(t) \\ \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} y(t) \\ x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ R(t) \end{bmatrix} \quad (137)$$

Since the equations are equivalent they consequently have the same solutions. Sometimes one of the transformations is more convenient to solve. Let us also transform a two dimensional system into a second order equation

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}. \quad (138)$$

First use the first equation to express x_2 and then take the time derivative of the first. Then we can eliminate x_2 and its time derivative

$$\dot{x}_1(t) = a_{11}x_1(t) + a_{12}x_2(t) + k_1 \quad (139)$$

$$x_2(t) = \frac{\dot{x}_1(t) - a_{11}x_1(t) - k_1}{a_{12}} \quad (140)$$

$$\dot{x}_2(t) = \frac{\ddot{x}_1(t) - a_{11}\dot{x}_1(t)}{a_{12}} \quad (141)$$

$$\frac{\ddot{x}_1(t) - a_{11}\dot{x}_1(t)}{a_{12}} = a_{21}x_1(t) + a_{22}\frac{\dot{x}_1(t) - a_{11}x_1(t) - k_1}{a_{12}} + k_2 \quad (142)$$

$$\ddot{x}_1(t) - (a_{11} + a_{22})\dot{x}_1(t) + (a_{22}a_{11} - a_{12}a_{21})x_1(t) = -a_{22}k_1 + a_{12}k_2 \quad (143)$$

Note that the characteristic equation of this second order equation is the same as the one for the system in (138). Consequently, the roots and thus the dynamics are identical.

3.5 Non-linear systems and Linearization

Phase diagrams are convenient to analyze the behavior of a 2 dimensional system qualitatively that we cannot or refer no to solve explicitly. E.g., if

$$\dot{c}(t) = g_1(c(t), k(t)) \quad (144)$$

$$\dot{k}(t) = g_2(c(t), k(t)) \quad (145)$$

The first step here is to find the two curves in the c, k -space where c and k , respectively are constant. Setting the time derivatives equal to zero defines two relations between c and k , which we denote by G_1 and G_2 .

$$g_1(c(t), k(t)) = 0 \rightarrow c = G_1(k) \quad (146)$$

$$g_2(c(t), k(t)) = 0 \rightarrow c = G_2(k) \quad (147)$$

We then draw these curves in the c, k -space. For example, you will in the macro course analyze the Ramsey optimal consumption problem, where output is $f(k)$, interest by $f'(k)$, the subjective discount rate by θ and utility is $u(c)$ and u and f are assumed to be concave functions. The model will deliver the following system of differential equations

$$\dot{c}(t) = -\frac{u'(c(t))}{u''(c(t))} (f'(k(t)) - \theta) \quad (148)$$

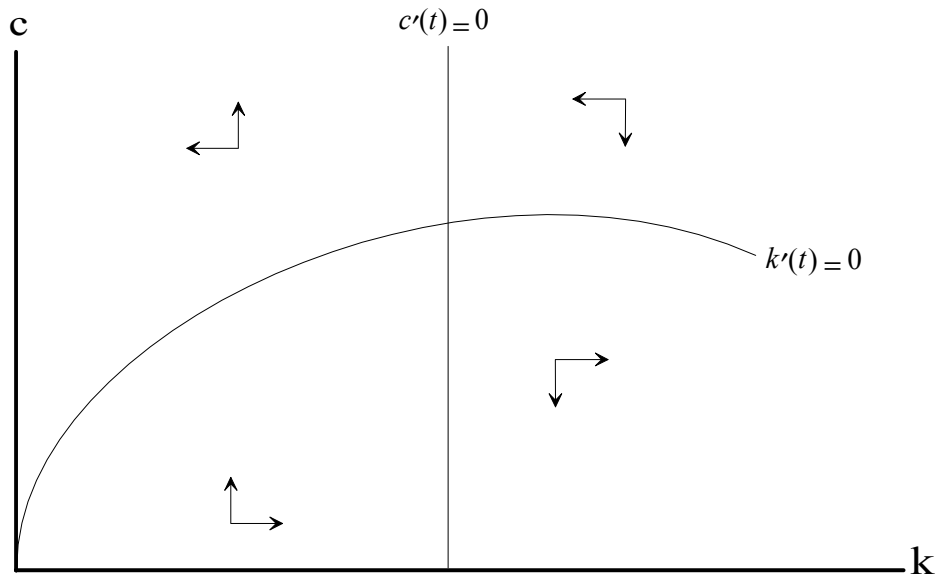
$$\dot{k}(t) = f(k(t)) - c(t). \quad (149)$$

Setting the time derivatives to zero we get

$$f'(k(t)) = \theta, \quad (150)$$

$$c = f(k(t)). \quad (151)$$

Draw these curves in the c, k space

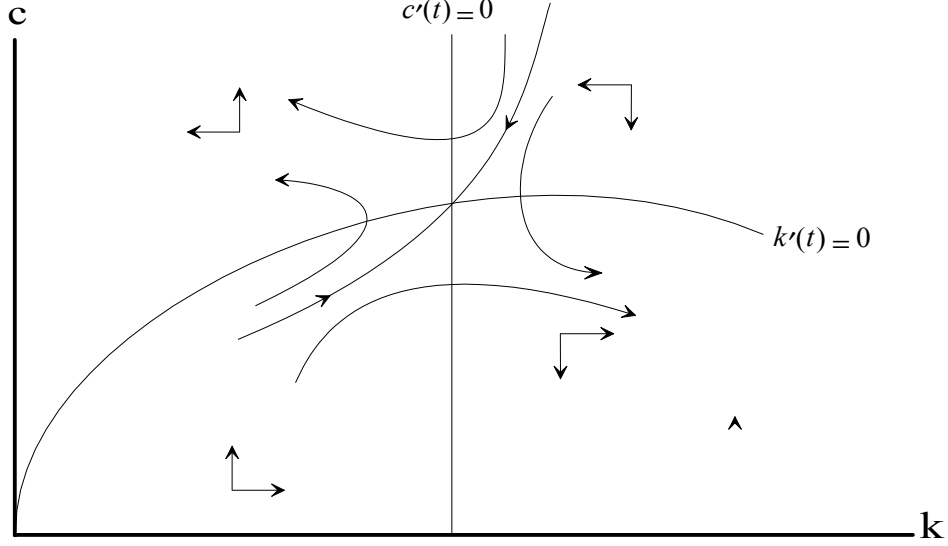


We then have to find the signs of \dot{c} and \dot{k} above and below their respective zero motion curves. From (148), we see that

$$\frac{\partial \dot{c}(t)}{\partial k} = -\frac{u'(c)}{u''(c)} f''(k) < 0 \quad (152)$$

$$\frac{\partial \dot{k}(t)}{\partial c} = -1. \quad (153)$$

This means that \dot{c} is positive to the left of and negative to the right of the curve $\dot{c} = 0$. For \dot{k} , we find that it is positive below and negative above $\dot{k} = 0$. Then draw these motions as arrows in the phase diagram. Note that no paths ever can cross.



We conclude that this system has saddle point characteristics and thus have only one stable trajectory towards the steady state.

The behavior close to the steady state should also be evaluated by means of linearization around the steady state. We do that by approximating in the following way

$$\begin{bmatrix} \dot{c}(t) \\ \dot{k}(t) \end{bmatrix} \approx \begin{bmatrix} \frac{\partial g_1(c^{ss}, k^{ss})}{\partial c} & \frac{\partial g_1(c^{ss}, k^{ss})}{\partial k} \\ \frac{\partial g_2(c^{ss}, k^{ss})}{\partial c} & \frac{\partial g_2(c^{ss}, k^{ss})}{\partial k} \end{bmatrix} \begin{bmatrix} c(t) - c^{ss} \\ k(t) - k^{ss} \end{bmatrix}. \quad (154)$$

with an obvious generalization to higher dimensions.

We now evaluate the roots of the matrix of derivatives. In the example we find that the coefficient matrix is

$$\begin{bmatrix} \frac{\partial g_1(c^{ss}, k^{ss})}{\partial c} & \frac{\partial g_1(c^{ss}, k^{ss})}{\partial k} \\ \frac{\partial g_2(c^{ss}, k^{ss})}{\partial c} & \frac{\partial g_2(c^{ss}, k^{ss})}{\partial k} \end{bmatrix} \quad (155)$$

$$= \begin{bmatrix} -(f' - \theta) \frac{\partial}{\partial c} \left(\frac{u'}{u''} \right) & -\frac{u'}{u''} f'' \\ -1 & f' \end{bmatrix}, \quad (156)$$

where all functions are evaluated at the steady state. There, $f' = \theta$, implying that the matrix simplifies to

$$\begin{bmatrix} 0 & -\frac{u'}{u''} f'' \\ -1 & f' \end{bmatrix} \quad (157)$$

with eigenvalues

$$\frac{1}{2} \left(f' \pm \sqrt{(f')^2 + 4 \frac{u'}{u''} f''} \right) \quad (158)$$

which clearly are of opposite signs.

3.6 Example: Steady-state asset distributions

In this example, we use our derived skills to find the steady-state wealth distribution in a simple model. As we will see, here the model gives us a system of differential equations for the wealth distribution that we can solve easily. We will see that the methods we have learned work as well when the differential equations are in wealth, rather than time.

In JPub (1999), we analyze preferences over unemployment insurance in a very simple economy using continuous time and where agents can borrow and save at a rate r . In the model, employed individuals earn a wage w per unit of time and become unemployed with an instantaneous probability q . This means that over a small (infinitesimal) interval of time dt , the probability of becoming unemployed is qdt . Unemployed individuals receive a flow of unemployment benefits b and become rehired with instantaneous probability hdt . In addition, we assume that there is a instantaneous death probability of δ and that there is an inflow of newborn unemployed with zero assets of δ so that the total population is constant.

In the paper, we show that individuals have CARA utility ($U = -e^{\gamma c}$) and wages and benefits are constant, individuals choose constant savings amount for each of the two employment states. Denoting these s_e and s_u , where $s_e > 0$ and $s_u < 0$, we have that individual asset accumulation for the two types, conditional on surviving is

$$\dot{A}_t = s_e, \quad (159)$$

for employed and

$$\dot{A}_t = s_u \quad (160)$$

for unemployed.

In the paper, we don't calculate the steady state wealth distribution of assets. That's the purpose of this exercise. First, we calculate the steady state share of unemployed, μ_u . For this purpose, we note that in steady state, the inflow and the outflow to the stock of unemployed must be constant. Thus,

$$(1 - \mu_u)q + \delta = \mu_u(h + \delta) \quad (161)$$

$$\rightarrow \mu_u = \frac{q + \delta}{q + h + \delta} \quad (162)$$

Now, we want to calculate the steady state distribution of assets among employed and unemployed in this economy. Let us denote these densities, by $f_e(A)$ and $f_u(A)$. We will derive these by solving a system of differential equations. Consider first the number (density) of employed individuals with assets $A \neq 0$, given by $f_e(A)$. Over a small period dt , a number $f_e(A)(1 - (\delta + q)dt)$ of them remain employed and alive. Over the same time, a number $f_u(A)hdt$ of unemployed with assets A become hired. Finally, over the period dt these individuals add to their assets an amount $s_e dt$. Writing this down yields,

$$f_e(A + s_e dt) = f_e(A)(1 - (\delta + q)dt) + f_u(A)hdt, \quad (163)$$

$$f_e(A) + f'_e(A)s_e dt = f_e(A)(1 - (\delta + q)dt) + f_u(A)hdt \quad (164)$$

$$f'_e(A) = -f_e(A)\frac{\delta + q}{s_e} + f_u(A)\frac{h}{s_e}. \quad (165)$$

By the same reasoning,

$$f_u(A + s_e dt) = f_u(A)(1 - (\delta + h)dt) + f_e(A)qdt, \quad (166)$$

$$f_u(A) + f'_u(A)s_e dt = f_u(A)(1 - (\delta + h)dt) + f_e(A)qdt \quad (167)$$

$$f'_u(A) = -f_u(A)\frac{\delta + h}{s_u} + f_e(A)\frac{q}{s_u}, \quad (168)$$

yielding the system

$$\begin{bmatrix} f'_e(A) \\ f'_u(A) \end{bmatrix} = \begin{bmatrix} -\frac{\delta+q}{s_e} & \frac{h}{s_e} \\ \frac{q}{s_u} & -\frac{\delta+h}{s_u} \end{bmatrix} \begin{bmatrix} f_e(A) \\ f_u(A) \end{bmatrix}. \quad (169)$$

$$\equiv \mathbf{A} \begin{bmatrix} f_e(A) \\ f_u(A) \end{bmatrix} \quad (170)$$

Let us now, assigns some numbers to the parameters. Say $\delta = 1/40$, $h = 2$, $q = 1/5$, $s_e = 1$ and $s_u = -2$, all measured as probabilities per year. Then, we can calculate the roots and the eigenvalues

$$r_1 = \frac{63}{160} + \frac{1}{160}\sqrt{4681} \approx .821 \quad (171)$$

$$r_2 = \frac{63}{160} - \frac{1}{160}\sqrt{4681} \approx -0.0339 \quad (172)$$

$$\mathbf{B}^{-1} = \begin{bmatrix} \frac{99}{320} + \frac{1}{320}\sqrt{4681} & \frac{99}{320} - \frac{1}{320}\sqrt{4681} \end{bmatrix} \approx \begin{bmatrix} 1.0 & 1.0 \\ .523 & 0.0956 \end{bmatrix} \quad (173)$$

Thus, except at $A = 0$, where there is an inflow of newborn, that we have not considered, we have

$$\begin{bmatrix} f_e(A) \\ f_u(A) \end{bmatrix} = \mathbf{B}^{-1} \mathbf{r}(t) \mathbf{c} \quad (174)$$

$$= \begin{bmatrix} 1.0 & 1.0 \\ .523 & 0.0956 \end{bmatrix} \begin{bmatrix} e^{.821A} & 0 \\ 0 & e^{-.0339A} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad (175)$$

$$= \begin{bmatrix} c_1 e^{.821A} + c_2 e^{-.0339A} \\ .523c_1 e^{.821A} + .0956c_2 e^{-.0339A} \end{bmatrix} \quad (176)$$

Now, we first note that $f_e(A)$ and $f_u(A)$ cannot be explosive in any of the directions. That would violated that these functions are densities, i.e., the sum of their respective integrals over the real line must be unity. Thus, for $A < 0$, c_2 must be zero and for $A > 0$, $c_1 = 0$. Furthermore, we know

$$\int_{-\infty}^{\infty} f_e(A) dA = 1 - \mu_u = \frac{h}{q + h + \delta} \approx 0.899 \quad (177)$$

$$\int_{-\infty}^{\infty} f_u(A) dA = \mu_u = \frac{q + \delta}{q + h + \delta} \approx 0.101. \quad (178)$$

Using this, we can calculate the integrations constants.

$$\int_{-\infty}^{\infty} f_e(A) dA = \int_{-\infty}^0 c_1 e^{.821A} dA + \int_0^{\infty} c_2 e^{-.0339A} dA \quad (179)$$

$$= \frac{c_1}{0.821} + \frac{c_2}{.0339} = 0.899 \quad (180)$$

$$\int_{-\infty}^{\infty} f_u(A) dA = \int_{-\infty}^0 .523c_1 e^{.821A} dA + \int_0^{\infty} .0956c_2 e^{-.0339A} dA \quad (181)$$

$$= \frac{.523c_1}{0.821} + \frac{0.0956c_2}{.0339} = 0.101 \quad (182)$$

Yielding,

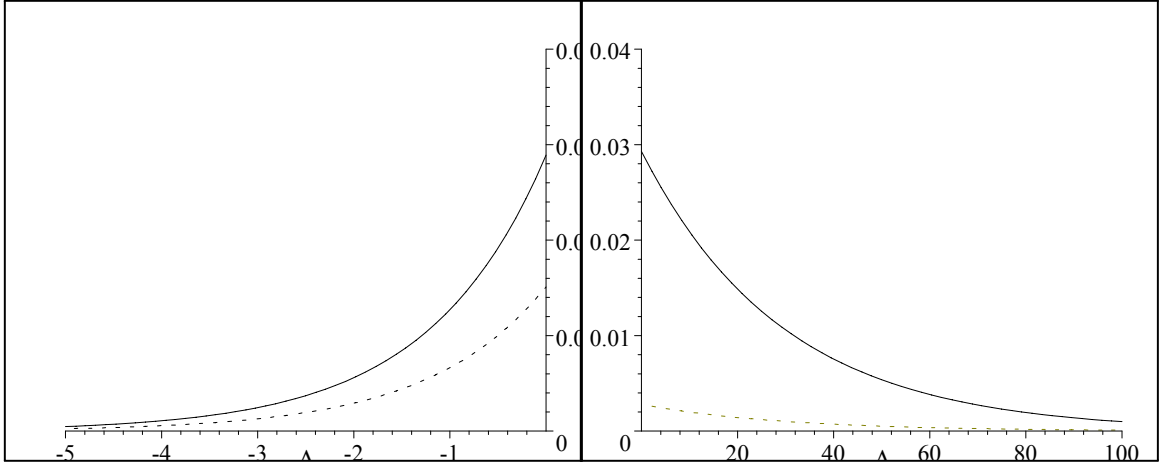
$$c_1 = 0.0289, \quad (183)$$

$$c_2 = 0.0293. \quad (184)$$

This concludes our calculations,

$$f_e(A) = \begin{cases} 0.0289e^{.821A} & \text{for } A < 0 \\ 0.0293e^{-.0339A} & \text{for } A > 0 \end{cases}, \quad (185)$$

$$f_u(A) = \begin{cases} 0.0151e^{.821A} & \text{for } A < 0 \\ 0.00280e^{-.0339A} & \text{for } A > 0 \end{cases}. \quad (186)$$



Plotting the densities, using a solid (dotted) line to denote wealth of employed (unemployed), we see that unemployed are more concentrated to the left. We can also calculate average assets among the two types from

$$A_e = \int_{-\infty}^{\infty} A f_e(A) dA = \quad (187)$$

$$\int_{-\infty}^0 A \cdot 0.0289 e^{821A} dA + \int_0^{\infty} A \cdot 0.0293 e^{-0.339A} dA \approx 25.45 \quad (188)$$

$$A_u = \int_{-\infty}^{\infty} A f_u(A) dA = \quad (189)$$

$$\int_{-\infty}^0 A \cdot 0.0151 e^{821A} dA + \int_0^{\infty} A \cdot 0.00280 e^{-0.339A} dA \approx 2.41 \quad (190)$$