Lyapunov's **Direct Method for Flows**

Lyapunov's 2nd or Direct Method is a technique for investigating stability which generalises how energy behaves in a physical system to an arbitrary dynamical system. The "*Lyapunov* function" plays the role of this energy function: it is minimum when the system is in equilibrium, it increases in value when the equilibrium is perturbed since such a perturbation is equivalent to injecting energy into the system, and if the system is stable the energy will not increase, indeed it usually decreases as the system returns to equilibrium.
We start by looking at a number of functional properties that we'll use in describing *Lyapunov* functions.

1 Positive Definite Functions & Quadratic Functions

Given a neighbourhood $\mathcal{N}$ of $\mathbf{0} \in \mathbb{R}^n$ a function $W : \mathbb{R}^n \to \mathbb{R}$ is

- positive definite (pd) if

(pd 1) $W(\mathbf{0}) = 0$
(pd 2) $W(\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0} \in \mathcal{N}$

- positive semi-definite (psd) if

(psd 1) $W(\mathbf{0}) = 0$
(psd 2) $W(\mathbf{x}) \geq 0$ for $\mathbf{x} \in \mathcal{N}$

- negative definite (nd) if $-W(\mathbf{x})$ is pd
- negative semi-definite (nsd) if $-W(\mathbf{x})$ is psd.

A quadratic function $QF : \mathbb{R}^n \to \mathbb{R}$ is given by

$$\begin{aligned} QF(\mathbf{x}) &= \sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij}x_ix_j \\ &= \mathbf{x}^T P \mathbf{x} \end{aligned}$$

where P is a symmetric matrix $P = P^T = [P_{ij}]$.

It may be shown that a quadratic function QF is pd (respectively psd, nd, nsd) if all the eigenvalues of P are positive (respectively greater than or equal to zero, negative, less than or equal to zero). In all cases the matrix P inherits the corresponding name of the quadratic function. A computationally straightforward way of checking whether a symmetric matrix P is pd is given by *Sylvester*'s criterion: P is pd iff all its leading principal minors are positive.

Examples:

$$QF(\mathbf{x}) = x^2 - 2xy + 2y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

is pd since its leading principal minors are $P_{11} = 1$ and $det\ P = 1$ respectively.

$$QF(\mathbf{x}) = -2x_1^2 - 2x_1x_2 + 2x_1x_3 - 3x_2^2 - x_3^2 = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{bmatrix} -2 & -1 & 1 \\ -1 & -3 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

is not pd since its leading principal minors are $P_{11} = -2$, $(P_{11})(P_{22}) - (P_{12})(P_{21}) = 5$ and $det(P) = -2$. Can *Sylvester*'s criterion tell us anything else? We can use it to check whether P is nd by investigating $-P$. The leading principal minors of this latter matrix are 2, 5 and 2, and so we can conclude that P is nd.
We remark in passing that *Sylvester*'s criterion cannot be used in general to identify psd matrices.

2 *Lyapunov* Local Stability Theorems

Consider the flow

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \qquad \mathbf{x}_e = \mathbf{0} \tag{1}$$

Define the *Lyapunov* function $V : \mathcal{N} \to \mathbb{R}$ by

(L1) V is pd

(L2) $\frac{dV}{dt} = \left(\frac{\partial V}{\partial \mathbf{x}}\right)^T \dot{\mathbf{x}} = \left(\frac{\partial V}{\partial \mathbf{x}}\right)^T \mathbf{f}(\mathbf{x}) = ||\frac{\partial V}{\partial \mathbf{x}}||\ ||\dot{\mathbf{x}}||\ \cos\theta$ is nsd

$\frac{\partial V}{\partial \mathbf{x}}$ is the gradient of V, and it is orthogonal to the contours of V and pointing away from the origin; θ is the angle between $\frac{\partial V}{\partial \mathbf{x}}$ and $\dot{\mathbf{x}}$. Thus if $\frac{dV}{dt} < 0$ at a point $\mathbf{x} = \mathbf{x}_c$ in the phase plane, then the system's trajectory is crossing the contour $V(\mathbf{x}_c)$ from a higher value of V to a lower value of V, i.e. is heading towards $\mathbf{0}$. (See Fig. 1)

Theorem 1 (Stability) *If there exists a Lyapunov function for the system of Eq(1), then* $\mathbf{x}_e = \mathbf{0}$ *is stable.*

Theorem 2 (Asymptotic stability) *If there exists a Lyapunov function for the system of Eq(1), with the additional property that*
(L3) $\frac{dV}{dt}$ is nd
then $\mathbf{x}_e = \mathbf{0}$ *is asymptotically stable.*

Theorem 3 (Instability) *If there exists a pd function V for which $\frac{dV}{dt}$ is also pd for the system of Eq(1), then* $\mathbf{x}_e = \mathbf{0}$ *is unstable.*

3 *Lyapunov* Global Stability Theorems

These mirror the local stability theorems, but with some differences

- $\mathcal{N} = \mathbb{R}^n$ (of course),
- $V(\mathbf{x}) \to \infty$ as $||\mathbf{x}|| \to \infty$. A function with this property is said to be radially unbounded and it is needed to ensure that the contours of V define closed curves.

With these provisos, the global stability theorems read the same as the local stability ones.

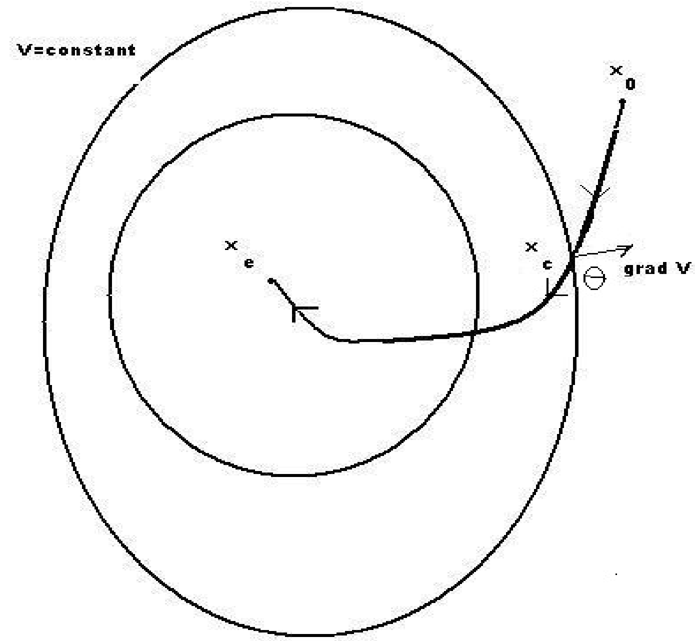

Figure 1: Trajectory crossing a contour

4 La Salle's Invariance Principle

This principle can be of use in the situation where the Direct Method enables us to infer stability but not asymptotic stability, but we suspect that the latter is indeed the case. Recall that a (forward) invariant set S for system of Eq (1) has the property that

$$\mathbf{x}_0 \in S \qquad \Rightarrow \qquad \mathbf{x}(t) = \Phi(t, \mathbf{x}_0) \in S, \;\; t \geq 0$$

Assume that we have a function V: Define the neighbourhood $\mathcal{N}_b = \{\mathbf{x} : V(\mathbf{x}) < b\}$. Furthermore assume that $\frac{dV}{dt} \leq 0$ for $\mathbf{x} \in \mathcal{N}_b$. Define $R = \{\mathbf{x} \in \mathcal{N}_b : \frac{dV}{dt} = 0\}$ and let M be the union of all invariant sets in R. The Invariance Principle says that if $\mathbf{x}_0 \in \mathcal{N}_b$ then $\mathbf{x}(t) \to M$ as $t \to \infty$. [See Fig. 2].

Corollary 4 *If (i) V is pd, (ii) $\frac{dV}{dt}$ is nsd and (iii) $M = \{\mathbf{0}\}$, then $\mathbf{0}$ is locally asymptotically stable.*

Corollary 5 *If (i) V is pd everywhere and radially unbounded, (ii) $\frac{dV}{dt}$ is nsd everywhere and (iii) $M = \{\mathbf{0}\}$, then $\mathbf{0}$ is globally asymptotically stable.*

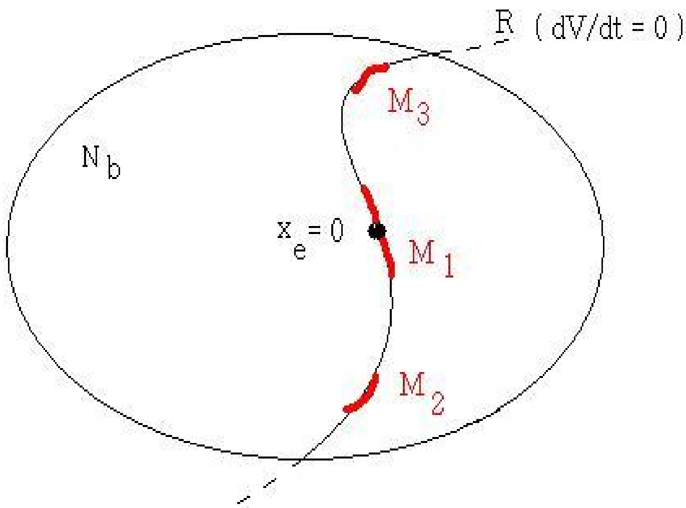

Figure 2: La Salle's Invariance Principle: Invariant set $M = M_1 \cup M_2 \cup M_3$ for which $\frac{dV}{dt} = 0$

5 Two canonical examples

- **1-d flow**

$$\boxed{\dot{x} = -c(x)} \tag{2}$$

(In the context of *Lyapunov*-like approaches , this is frequently written as $\dot{x} + c(x) = 0$) where $c(x)$ is continuous and has the same sign as x i.e.

$$xc(x) > 0, \qquad x \neq 0$$

Continuity requires that $c(0) = 0$ [See Fig. 3] and thus $x_e = 0$.

Consider the radially unbounded pd function

$$\begin{aligned} V(x) &= \frac{1}{2}x^2 \\ \Rightarrow \frac{dV}{dt} &= x\dot{x} \\ &= -xc(x) \\ &< 0, \qquad \text{if } x \neq 0 \end{aligned}$$

Thus $\frac{dV}{dt}$ is nd, and we can conclude that $x_e = 0$ is globally asymptotically stable. Specific examples are:

$$\dot{x} + 2x = 0, \qquad \dot{x} + x|x| = 0, \qquad \dot{x} + x^3 = 0, \qquad \dot{x} + x - \sin x = 0$$

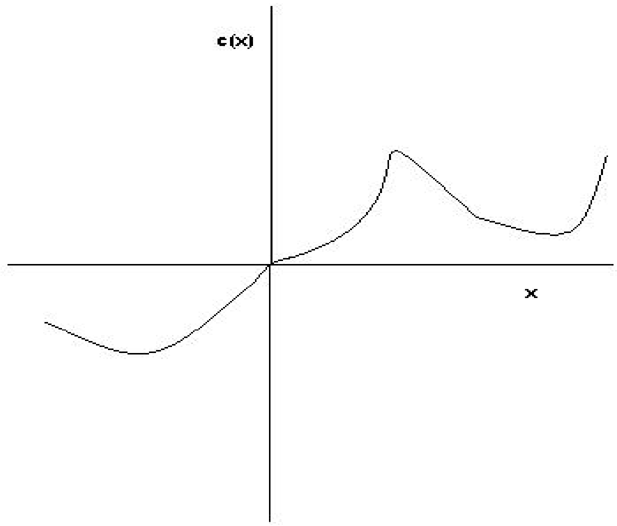

Figure 3: the function $c(x)$

Can anything be be said about

$$\dot{x} + x^2 = 0?$$

What can be said if c is continuous but that $xc(x) > 0$ on a finite interval containing $x = 0$?

- **2-d flow** Consider the specific class of 2-d systems described by

$$\ddot{x} + b(\dot{x}) + c(x) = 0$$

where b and c are continuous functions and have the same sign as their arguments i.e.

$$\begin{aligned} \dot{x}b(\dot{x}) &> 0, & \dot{x} \neq 0 \\ xc(x) &> 0, & x \neq 0 \end{aligned}$$

Continuity requires that $b(0) = 0$ and $c(0) = 0$. Both b and c as functions of their respective arguments have the same "shape" as the function in Fig. 3.
The example is a generalisation of the spring-mass-damper system with nonlinear spring ($c(x)$) and nonlinear damping ($b(\dot{x})$). With the choice of state variables $x = x,\ y = \dot{x}$, the state equation becomes

$$\boxed{\dot{x} = y, \qquad \dot{y} = -c(x) - b(y)} \tag{3}$$

and with the given conditions on b and c, $\mathbf{x}_e = \mathbf{0}$.

Consider the pd function

$$\begin{aligned} V(x,y) &= \frac{1}{2}y^2 + \int_0^x c(s)ds \\ \Rightarrow \frac{dV}{dt} &= y\dot{y} + c(x)\dot{x} \\ &= -c(x)y - b(y)y + c(x)y \\ &= -b(y)y \\ &\le 0 \end{aligned}$$

Thus $\frac{dV}{dt}$ is nsd, implying that $\mathbf{x}_e = \mathbf{0}$ is stable. But is it asymptotically stable? Now $\frac{dV}{dt} = 0$ only if $y = 0$ giving us that $R = \{(x,0)^T\}$ - the x - axis. What is M? If $\mathbf{x} \in R$ but $x \neq 0$ then $\dot{y} = -c(x) \neq 0$ leading to the conclusion that y changes value i.e. is no longer 0 and so the trajectory leaves R. Hence if $\mathbf{x} \in M$ then x must be zero and so $M = \{(0,0)^T\} = \mathbf{0}$. The Invariance Principle then tells us that $\mathbf{0}$ is locally asymptotically stable.
In addition, if the integral $\int_0^x c(s)ds$ is unbounded as $|x| \to \infty$, then V is radially unbounded and we get that $\mathbf{0}$ is globally asymptotically stable.
Specific examples are:

$$\ddot{x} + 2\dot{x}|\dot{x}| + x - \frac{1}{3}x^3 = 0, \qquad \ddot{x} + \dot{x}^3 + x^5 - x^4 \sin^2 x = 0$$

Again what can be said if b or c are continuous such that $yb(y) > 0$ or $xc(x) > 0$ on a finite interval containing 0?

6 Linear Systems

Even though *Lyapunov*'s first or indirect method (linearization) is usually adequate for linear systems, the use of the direct method gives us another tool for stability analysis. Recall that for linear systems, local stability and global stability are synonymous. For the linear system

$$\dot{\mathbf{x}} = A\mathbf{x}, \qquad \mathbf{x}_e = \mathbf{0} \tag{4}$$

consider the quadratic function

$$\begin{aligned} V(\mathbf{x}) &= \mathbf{x}^T P \mathbf{x}, \qquad P^T = P \\ \Rightarrow \frac{dV}{dt} &= \dot{\mathbf{x}}^T P \mathbf{x} + \mathbf{x}^T P \dot{\mathbf{x}} \\ &= (A\mathbf{x})^T P \mathbf{x} + \mathbf{x}^T P (A\mathbf{x}) \\ &= \mathbf{x}^T A^T P \mathbf{x} + \mathbf{x}^T P A \mathbf{x} \\ &= \mathbf{x}^T (A^T P + PA)\mathbf{x} \\ &= -\mathbf{x}^T Q \mathbf{x} \end{aligned}$$

where we have set

$$\boxed{A^T P + PA = -Q} \tag{5}$$

This equation is known as the *Lyapunov* Equation. Note that Q is forced to be symmetric. Let's consider the case where we want to establish asymptotic stability of the fixed point

$\mathbf{x}_e = \mathbf{0}$. We require that (i) $V = \mathbf{x}^T P \mathbf{x}$ be pd or equivalently that the matrix P be pd and (ii) that $\frac{dV}{dt} = -\mathbf{x}^T Q \mathbf{x}$ be nd or equivalently that the matrix $-Q$ be nd, i.e. Q be pd. this leads to the following theorem

Theorem 6 (Linear Asymptotic Stability) *The (fixed point of the) system of Eq (4) is globally asymptotically stable if and only if, given any pd matrix Q, the solution P of Eq (5) is also pd.*

Example:

$$\dot{\mathbf{x}} = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} \mathbf{x}$$

Since Q can be any pd matrix, it is quite common to choose $Q = I$. Building in symmetry in P, the *Lyapunov* Equation for this system becomes

$$\begin{pmatrix} 0 & -2 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} p_1 & p_2 \\ p_2 & p_3 \end{pmatrix} + \begin{pmatrix} p_1 & p_2 \\ p_2 & p_3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

or writing this as a system of linear equations

$$\begin{aligned} -4p_2 &= -1 && \text{position } (1,1) \text{ in matrix equation} \\ -2p_3 + p_1 - 3p_2 &= 0 && \text{position } (1,2) \text{ in matrix equation} \\ p_1 - 3p_2 - 2p_3 &= 0 && \text{position } (2,1) \text{ in matrix equation} \\ 2p_2 - 6p_3 &= -1 && \text{position } (2,2) \text{ in matrix equation} \end{aligned}$$

(Notice that the equation in position $(2,1)$ is the same as that in position $(1,2)$, due to the built-in symmetry, and so there are only 3 independent equations in the 3 unknowns) The system has solution

$$p_1 = \frac{5}{4}, \qquad p_2 = \frac{1}{4}, \qquad p_3 = \frac{1}{4}$$

So we have

$$P = \begin{pmatrix} 5/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$$

which has leading principal minors $P_{11} = \frac{5}{4}$ and $detP = \frac{1}{4}$. Thus $\mathbf{x}_e = \mathbf{0}$ is globally asymptotically stable.

7 Estimating the Basin of Attraction for an asymptotically stable fixed point

Consider the system of Eq (1), where we assume that $\mathbf{x}_e = \mathbf{0}$ is asymptotically stable. Then there exists a neighbourhood ($\mathcal{N}$) of $\mathbf{0}$ with a pd $V(\mathbf{x})$ and corresponding nd $\frac{dV}{dt}$. The following approach can be used to estimate the basin of attraction for $\mathbf{0}$ (it is illustrated in terms of a 2-d problem):

- Solve

$$\frac{dV}{dt} = 0$$

Typically, one solution of this is $\mathbf{x} = \mathbf{0}$. We are interested in the other solution(s) which are generically curves. These curves partition the plane into regions where $\frac{dV}{dt}$ alternates from positive to negative and back again. We choose the largest region (Z) that contains $\mathbf{0}$ and where $\frac{dV}{dt} \leq 0$ throughout, and find the largest V-contour that lies entirely in Z.

- Solve the optimisation problem

$$\min_{\mathbf{x}\in Z} V(\mathbf{x}), \qquad \text{subject to } \frac{dV}{dt} = 0, \;\; \mathbf{x} \neq \mathbf{0}$$

  Let $m \triangleq \min_{\mathbf{x}\in Z} V(\mathbf{x})$.

- The estimate of the basin of attraction is then

$$\mathcal{B} = \{\mathbf{x} : V(\mathbf{x}) < m\}$$

  [See Fig.(4)]. Using a different function for V will result in a different estimate of the basin of attraction.

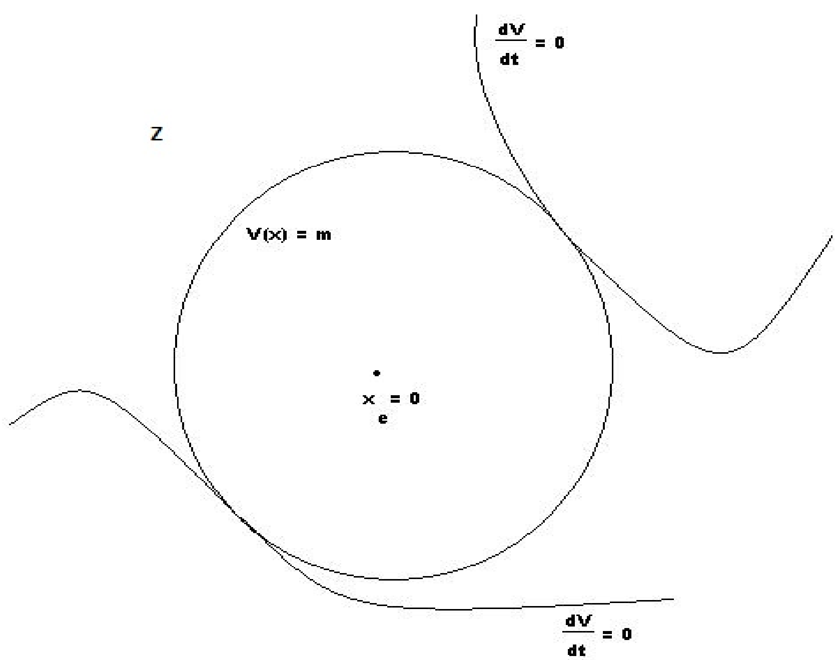

Figure 4: Basin of Attraction of stable fixed point

Example:

$$\dot{x} = -x + y, \qquad \dot{y} = -y + \frac{1}{2}x^2 \tag{6}$$

has a fixed point at $\mathbf{x}_e = \mathbf{0}$. The pd function

$$\begin{aligned} V(x,y) &= \frac{1}{2}x^2 + \frac{1}{2}xy + \frac{1}{2}y^2 \qquad (7)\\ \Rightarrow \frac{dV}{dt} &= (x + \frac{1}{2}y)\dot{x} + (\frac{1}{2}x + y)\dot{y} \\ &= -x^2 - \frac{1}{2}y^2 + \frac{1}{4}x^3 + \frac{1}{2}x^2 y \end{aligned}$$

We can argue that for $\mathbf{x}$ near $\mathbf{0}$ the third order terms in the expression for $\frac{dV}{dt}$ are negligible, and so $\frac{dV}{dt}$ is nd giving local asymptotic stability. We use the method discussed above to estimate the basin of attraction.

$$\begin{aligned} \frac{dV}{dt} &= 0 \\ \Rightarrow y^2 - x^2 y - \frac{1}{2}x^3 + 2x^2 &= 0 \end{aligned}$$

This defines two real branches for y as a function of x. The active branch - the one which is tangential to the relevant contour (see Fig. (5))- is

$$y = \frac{x}{2}\left(x - \sqrt{x^2 + 2x - 8}\right)$$

which function we will denote by $Y(x)$. Now define $G(x) \triangleq V(x, Y(x))$. Minimising G

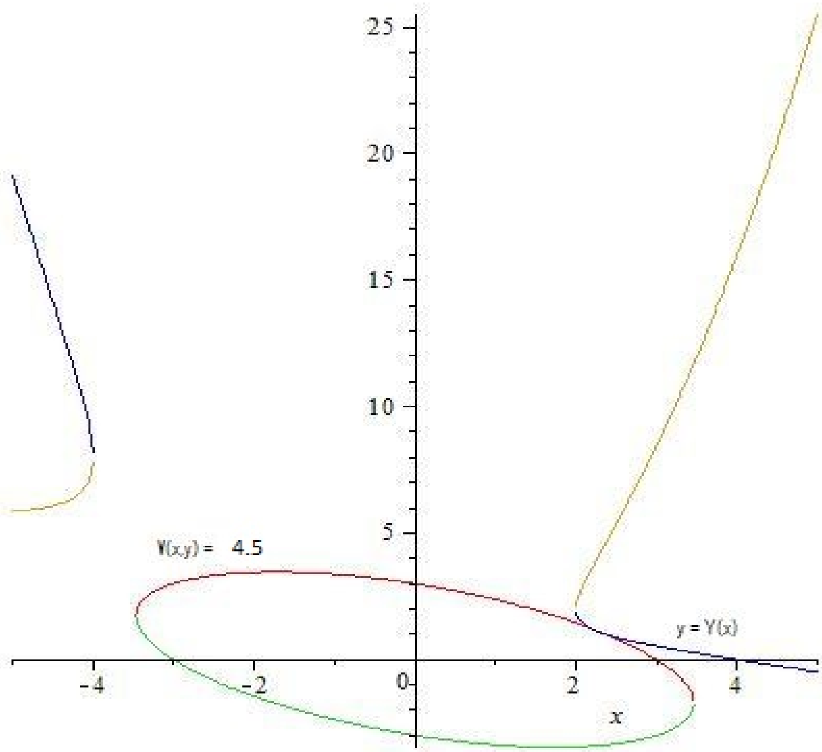

Figure 5: Basin of Attraction of stable fixed point of Eq(6) estimated using Eq(7).

as a function of x yields $m = \min_x G(x) = \frac{9}{2}$ at $x \approx 2.1523$ (found numerically). Hence the estimate of the basin of attraction is

$$\mathcal{B} = \left\{ (x,y)^T : \frac{1}{2}x^2 + \frac{1}{2}xy + \frac{1}{2}y^2 < \frac{9}{2} \right\}.$$