Order and stability for general linear methods

John Butcher

The University of Auckland

New Zealand

NUMDIFF 11

Martin Luther Universität Halle–Wittenberg

4–8 September 2006

There is always a conflict between three basic aims in the design of algorithms to solve ordinary differential equations.

There is always a conflict between three basic aims in the design of algorithms to solve ordinary differential equations. These aims are

- Good accuracy using high order methods
- Good stability — A-stability in the case of stiff methods
- Modest computational costs

There is always a conflict between three basic aims in the design of algorithms to solve ordinary differential equations. These aims are

- Good accuracy using high order methods
- Good stability — A-stability in the case of stiff methods
- Modest computational costs

A fourth unwritten aim is to keep the method as simple as possible but this is related to the other aims.

There is always a conflict between three basic aims in the design of algorithms to solve ordinary differential equations. These aims are

- Good accuracy using high order methods
- Good stability — A-stability in the case of stiff methods
- Modest computational costs

A fourth unwritten aim is to keep the method as simple as possible but this is related to the other aims.

I will discuss some of the conflicts between order and stability using order arrows and order stars to illustrate how they are inter-connected.

Contents

- General linear methods and Obrechkov methods

Contents

- General linear methods and Obrechkov methods
- Order stars and order arrows

Contents

- General linear methods and Obrechkov methods
- Order stars and order arrows
- Order arrows and stability results

Contents

Contents

- General linear methods and Obrechkov methods
- Order stars and order arrows
- Order arrows and stability results
- The Daniel-Moore theorem
- Padé approximations to the exponential function

Contents

Contents

- General linear methods and Obrechkov methods
- Order stars and order arrows
- Order arrows and stability results
- The Daniel-Moore theorem
- Padé approximations to the exponential function
- The Ehle theorem
- Quadratic approximations to the exponential

Contents

General linear methods and Obrechkov methods

General linear methods are multivalue-multistage methods in which the input to a step $y^{[n-1]}$ and the output from the step $y^{[n]}$ are related to the stage values Y and the stage derivatives $F = f(Y)$ by the equations

$$\begin{bmatrix} Y \\ y^{[n]} \end{bmatrix} = \begin{bmatrix} A & U \\ B & V \end{bmatrix} \begin{bmatrix} hF \\ y^{[n-1]} \end{bmatrix}$$

General linear methods and Obrechkov methods

General linear methods are multivalue-multistage methods in which the input to a step $y^{[n-1]}$ and the output from the step $y^{[n]}$ are related to the stage values Y and the stage derivatives $F = f(Y)$ by the equations

$$\begin{bmatrix} Y \\ y^{[n]} \end{bmatrix} = \begin{bmatrix} A & U \\ B & V \end{bmatrix} \begin{bmatrix} hF \\ y^{[n-1]} \end{bmatrix}$$

For the "linear test problem" $y'(x) = qy(x)$, we obtain the solution in the form

$$y^{[n]} = M(z)^n y^{[0]}, \qquad z = hq,$$

where $$M(z) = V + zB(I - zA)^{-1}U.$$

Because we are interested in stable behaviour of powers of $M(z)$, we want to know properties of the stability function $\Phi(w,z)$ given by

$$\frac{\Phi(w,z)}{\det(I-zA)} = \det(wI - M(z)).$$

Because we are interested in stable behaviour of powers of $M(z)$, we want to know properties of the stability function $\Phi(w, z)$ given by

$$\frac{\Phi(w, z)}{\det(I - zA)} = \det(wI - M(z)).$$

The open stability region is the set of values of z in the complex plane for which any solution to the equation

$$\Phi(w, z) = 0$$

lies in the interior of the unit disc.

A necessary condition for order p is that $\exp(z) + O(z^{p+1})$ is an eigenvalue of $M(z)$ for small z.

A necessary condition for order p is that $\exp(z) + O(z^{p+1})$ is an eigenvalue of $M(z)$ for small z.

Because we will not consider details of the method, but only the stability function, we will regard this as the definition of the order of $\Phi(w, z)$.

If we have available not only a formula for the first derivative $y'(x) = f(y(x))$, but also higher derivatives $y''(x) = f_2(y(x)), \ldots,$ we can widen the type of method considerably.

If we have available not only a formula for the first derivative $y'(x) = f(y(x))$, but also higher derivatives $y''(x) = f_2(y(x)), \dots,$ we can widen the type of method considerably. In particular we can include Obrechkov methods or multiderivative linear multistep methods.

If we have available not only a formula for the first derivative $y'(x) = f(y(x))$, but also higher derivatives $y''(x) = f_2(y(x)), \ldots$, we can widen the type of method considerably. In particular we can include Obrechkov methods or multiderivative linear multistep methods.

We will look at two examples, each of which is a second-derivative generalization of a BDF method.

The first example has order 4

$$y_n = \frac{66}{85}hy'_n - \frac{18}{85}h^2y''_n + \frac{108}{85}y_{n-1} - \frac{27}{85}y_{n-2} + \frac{4}{85}y_{n-3},$$

The first example has order 4

$$y_n = \frac{66}{85}hy'_n - \frac{18}{85}h^2y''_n + \frac{108}{85}y_{n-1} - \frac{27}{85}y_{n-2} + \frac{4}{85}y_{n-3},$$

leading to the stability function

$$(1 - \tfrac{66}{85}z + \tfrac{18}{85}z^2)w^3 - \tfrac{108}{85}w^2 + \tfrac{27}{85}w - \tfrac{4}{85}.$$

The first example has order 4

$$y_n = \frac{66}{85}hy'_n - \frac{18}{85}h^2y''_n + \frac{108}{85}y_{n-1} - \frac{27}{85}y_{n-2} + \frac{4}{85}y_{n-3},$$

leading to the stability function

$$(1 - \tfrac{66}{85}z + \tfrac{18}{85}z^2)w^3 - \tfrac{108}{85}w^2 + \tfrac{27}{85}w - \tfrac{4}{85}.$$

We can verify the order by substituting $w = \exp(z)$ and obtaining the result $O(z^5)$.

The first example has order 4

$$y_n = \frac{66}{85}hy'_n - \frac{18}{85}h^2y''_n + \frac{108}{85}y_{n-1} - \frac{27}{85}y_{n-2} + \frac{4}{85}y_{n-3},$$

leading to the stability function

$$(1 - \tfrac{66}{85}z + \tfrac{18}{85}z^2)w^3 - \tfrac{108}{85}w^2 + \tfrac{27}{85}w - \tfrac{4}{85}.$$

We can verify the order by substituting $w = \exp(z)$ and obtaining the result $O(z^5)$.

We will refer to this method as $[2, 0, 0, 0]$, where the name gives the degrees of the polynomials in z which appear as the coefficients of w^3, w^2, w^1, w^0.

The first example has order 4

$$y_n = \frac{66}{85} h y_n' - \frac{18}{85} h^2 y_n'' + \frac{108}{85} y_{n-1} - \frac{27}{85} y_{n-2} + \frac{4}{85} y_{n-3},$$

leading to the stability function

$$(1 - \tfrac{66}{85} z + \tfrac{18}{85} z^2) w^3 - \tfrac{108}{85} w^2 + \tfrac{27}{85} w - \tfrac{4}{85}.$$

We can verify the order by substituting $w = \exp(z)$ and obtaining the result $O(z^5)$.

We will refer to this method as $[2, 0, 0, 0]$, where the name gives the degrees of the polynomials in z which appear as the coefficients of w^3, w^2, w^1, w^0. The approximation has maximal order for these degrees.

The second example has order 5

$$y_n = \frac{60}{83}hy'_n - \frac{72}{415}h^2y''_n + \frac{576}{415}y_{n-1} - \frac{216}{415}y_{n-2} + \frac{64}{415}y_{n-3} - \frac{9}{415}y_{n-4},$$

The second example has order 5

$$y_n = \frac{60}{83}hy'_n - \frac{72}{415}h^2y''_n + \frac{576}{415}y_{n-1} - \frac{216}{415}y_{n-2} + \frac{64}{415}y_{n-3} - \frac{9}{415}y_{n-4},$$

leading to the stability function

$$(1 - \tfrac{60}{83}z + \tfrac{72}{415}z^2)w^4 - \tfrac{576}{415}w^3 + \tfrac{216}{415}w^2 - \tfrac{64}{415}w + \tfrac{9}{415}.$$

The second example has order 5

$$y_n = \frac{60}{83} h y_n' - \frac{72}{415} h^2 y_n'' + \frac{576}{415} y_{n-1} - \frac{216}{415} y_{n-2} + \frac{64}{415} y_{n-3} - \frac{9}{415} y_{n-4},$$

leading to the stability function

$$(1 - \tfrac{60}{83} z + \tfrac{72}{415} z^2) w^4 - \tfrac{576}{415} w^3 + \tfrac{216}{415} w^2 - \tfrac{64}{415} w + \tfrac{9}{415}.$$

Again we can verify the order by substituting $w = \exp(z)$, this time obtaining the result $O(z^6)$.

The second example has order 5

$$y_n = \frac{60}{83}hy_n' - \frac{72}{415}h^2y_n'' + \frac{576}{415}y_{n-1} - \frac{216}{415}y_{n-2} + \frac{64}{415}y_{n-3} - \frac{9}{415}y_{n-4},$$

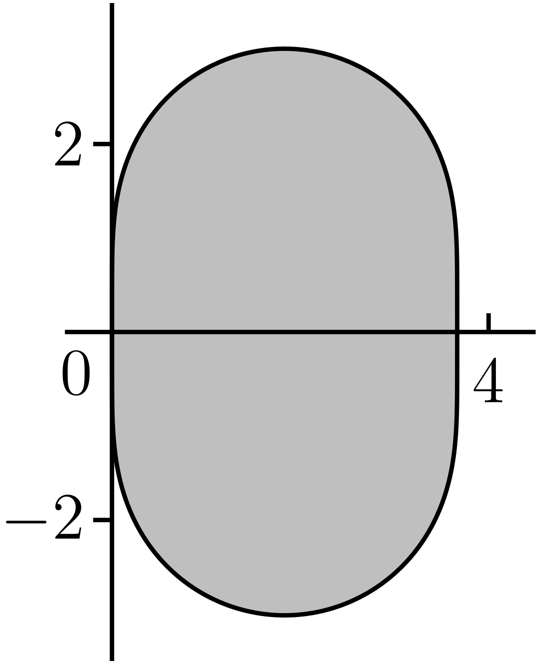
leading to the stability function

$$(1 - \tfrac{60}{83}z + \tfrac{72}{415}z^2)w^4 - \tfrac{576}{415}w^3 + \tfrac{216}{415}w^2 - \tfrac{64}{415}w + \tfrac{9}{415}.$$

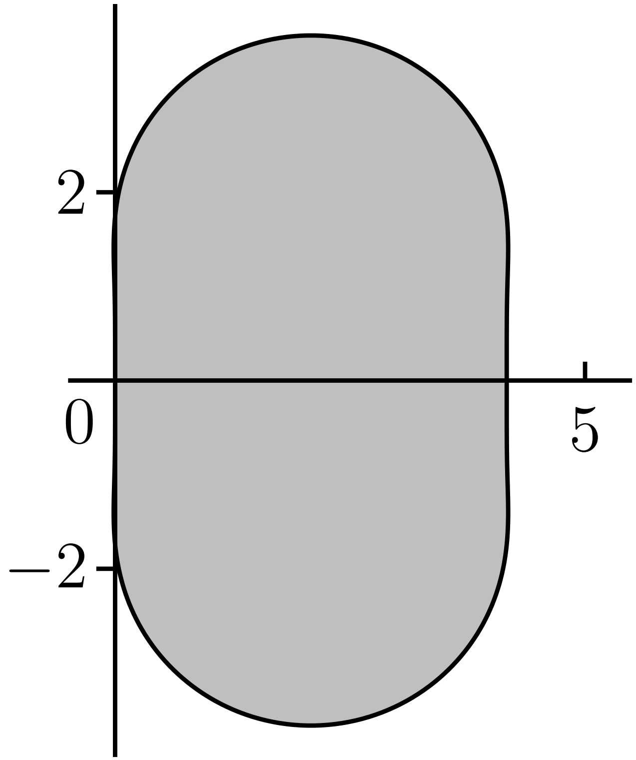
Again we can verify the order by substituting $w = \exp(z)$, this time obtaining the result $O(z^6)$.

This is the $[2, 0, 0, 0, 0]$ approximation.

The stability regions of these two methods are the unshaded regions in the diagrams:

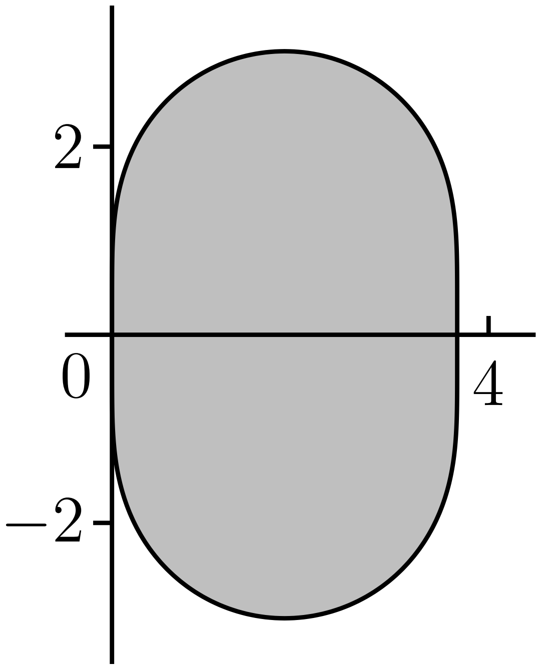


$[2, 0, 0, 0]$

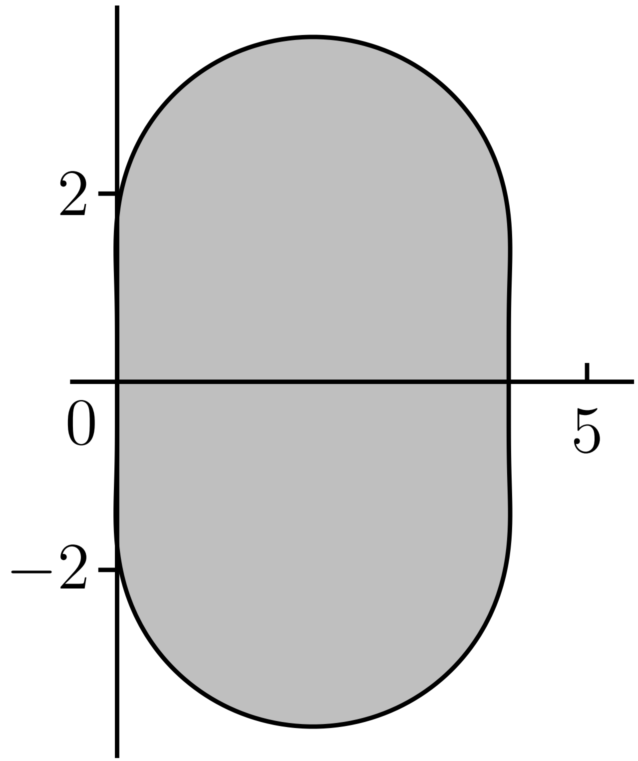


$[2, 0, 0, 0, 0]$

The stability regions of these two methods are the unshaded regions in the diagrams:



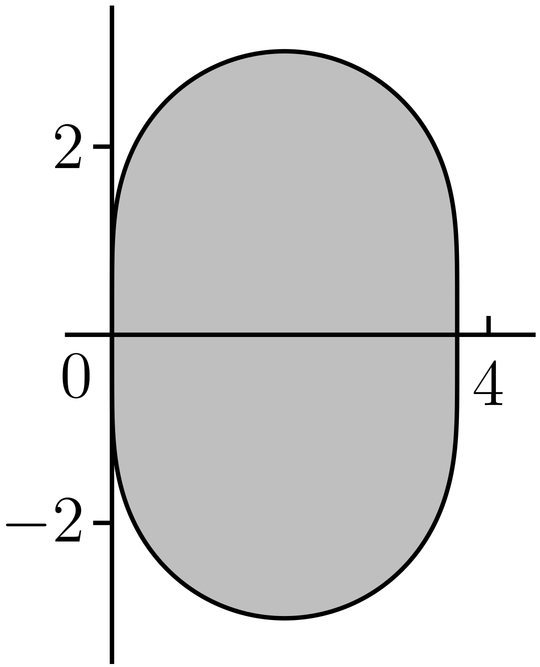
$[2, 0, 0, 0]$



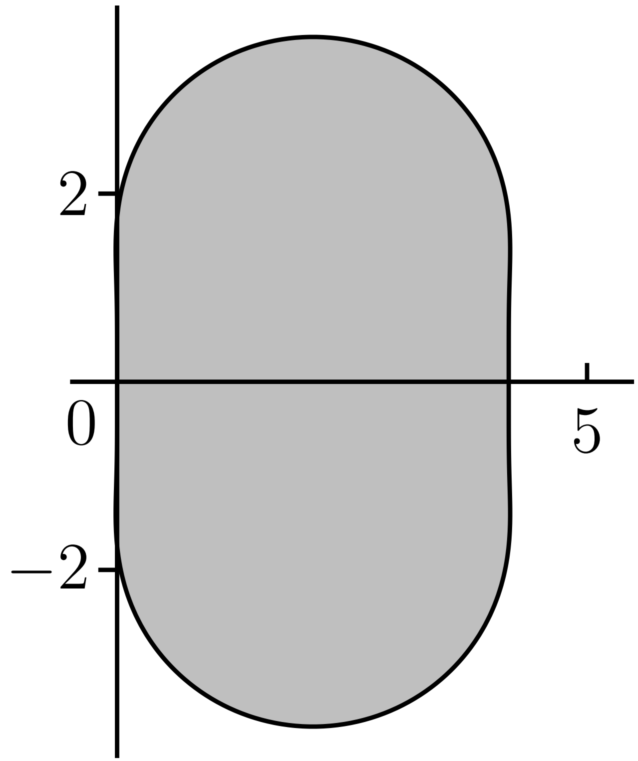
$[2, 0, 0, 0, 0]$

The methods are A-stable

The stability regions of these two methods are the unshaded regions in the diagrams:



$[2, 0, 0, 0]$



$[2, 0, 0, 0, 0]$

The methods are A-stable and A(89.365°)-stable respectively.

Order stars and order arrows

The use of order stars in settling stability questions is well-known.

Order stars and order arrows

The use of order stars in settling stability questions is well-known.

An alternative to order stars is “order arrows” and this is the approach we will emphasise.

Order stars and order arrows

The use of order stars in settling stability questions is well-known.

An alternative to order stars is “order arrows” and this is the approach we will emphasise.

In order stars we consider the sets of (w, z) pairs such that

$$\Phi(w \exp(z), z) = 0, \qquad (\star)$$

and such that $|w| > 1$ (or such that $|w| < 1$).

Order stars and order arrows

The use of order stars in settling stability questions is well-known.

An alternative to order stars is "order arrows" and this is the approach we will emphasise.

In order stars we consider the sets of (w, z) pairs such that

$$\Phi(w \exp(z), z) = 0, \qquad (\star)$$

and such that $|w| > 1$ (or such that $|w| < 1$).

For order arrows we consider the set of (w, z) pairs satisfying $(\star)$ such that w is real and positive.

Before considering complicated examples like the $[2, 0, 0, 0]$ and $[2, 0, 0, 0, 0]$ approximations we will look at standard Padé approximations to the exponential function.

Before considering complicated examples like the $[2, 0, 0, 0]$ and $[2, 0, 0, 0, 0]$ approximations we will look at standard Padé approximations to the exponential function.

We consider the example of the $[2, 1]$ Padé approximation for which

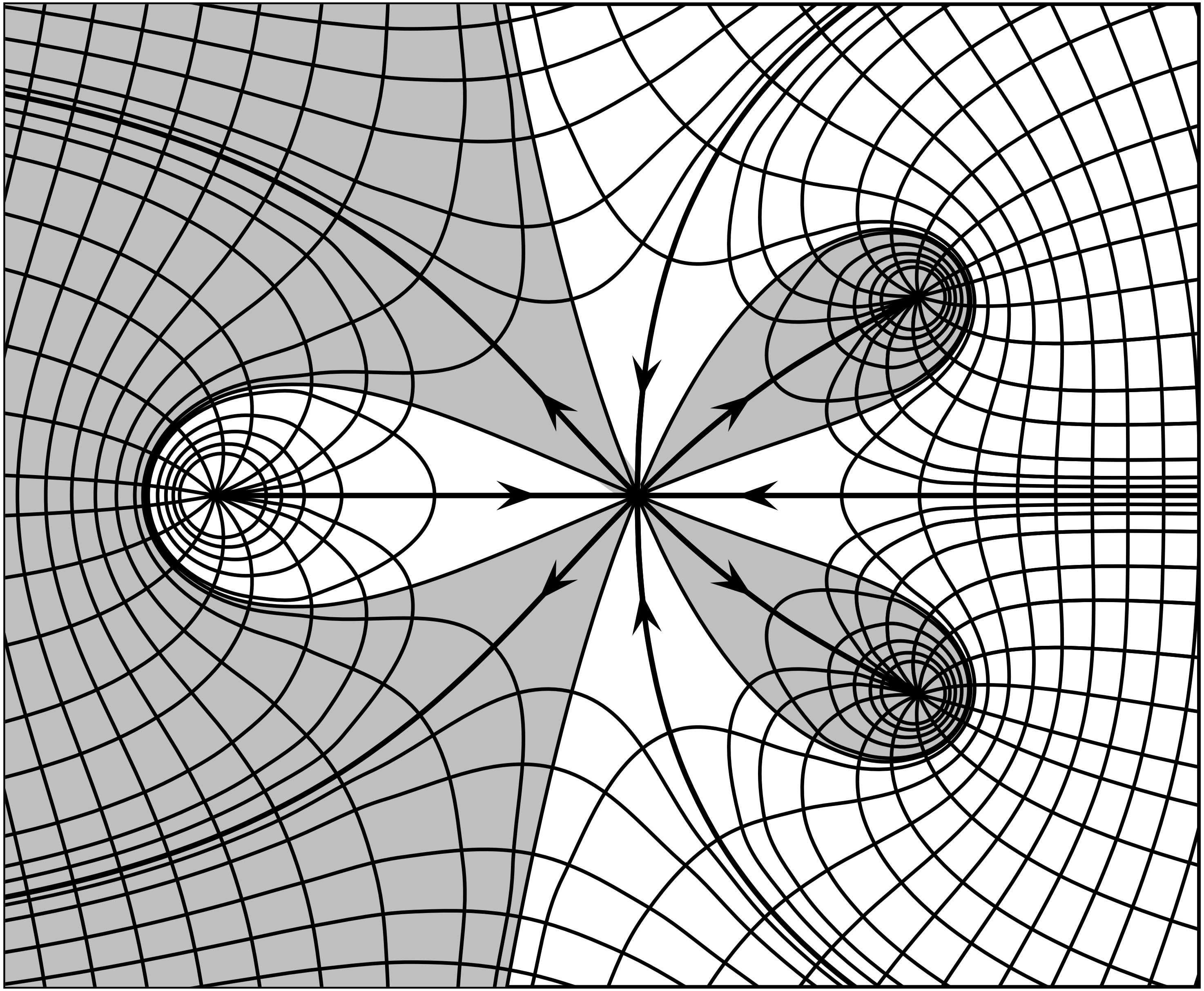
$$R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}.$$

Before considering complicated examples like the $[2, 0, 0, 0]$ and $[2, 0, 0, 0, 0]$ approximations we will look at standard Padé approximations to the exponential function.

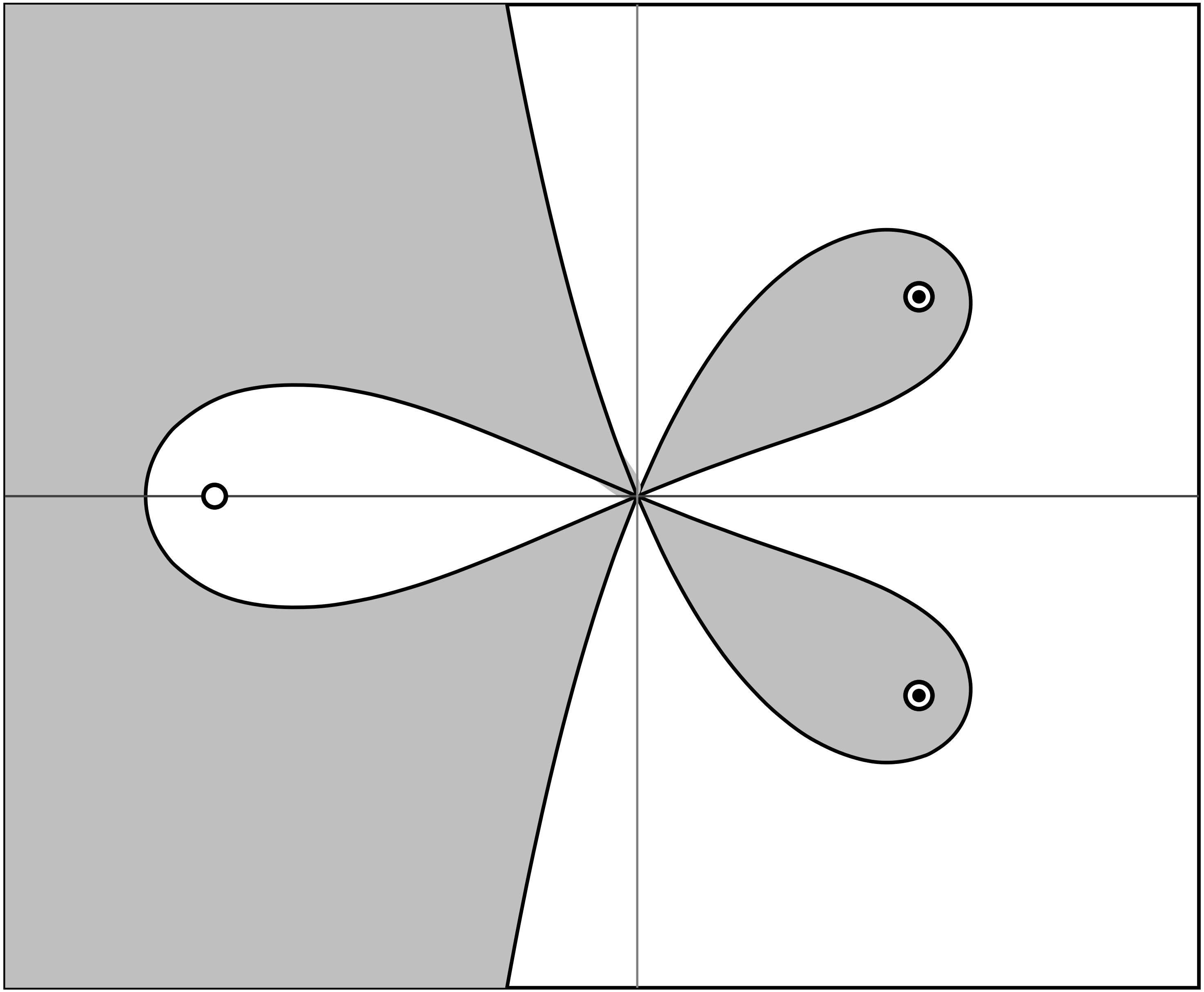
We consider the example of the $[2, 1]$ Padé approximation for which

$$R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}.$$

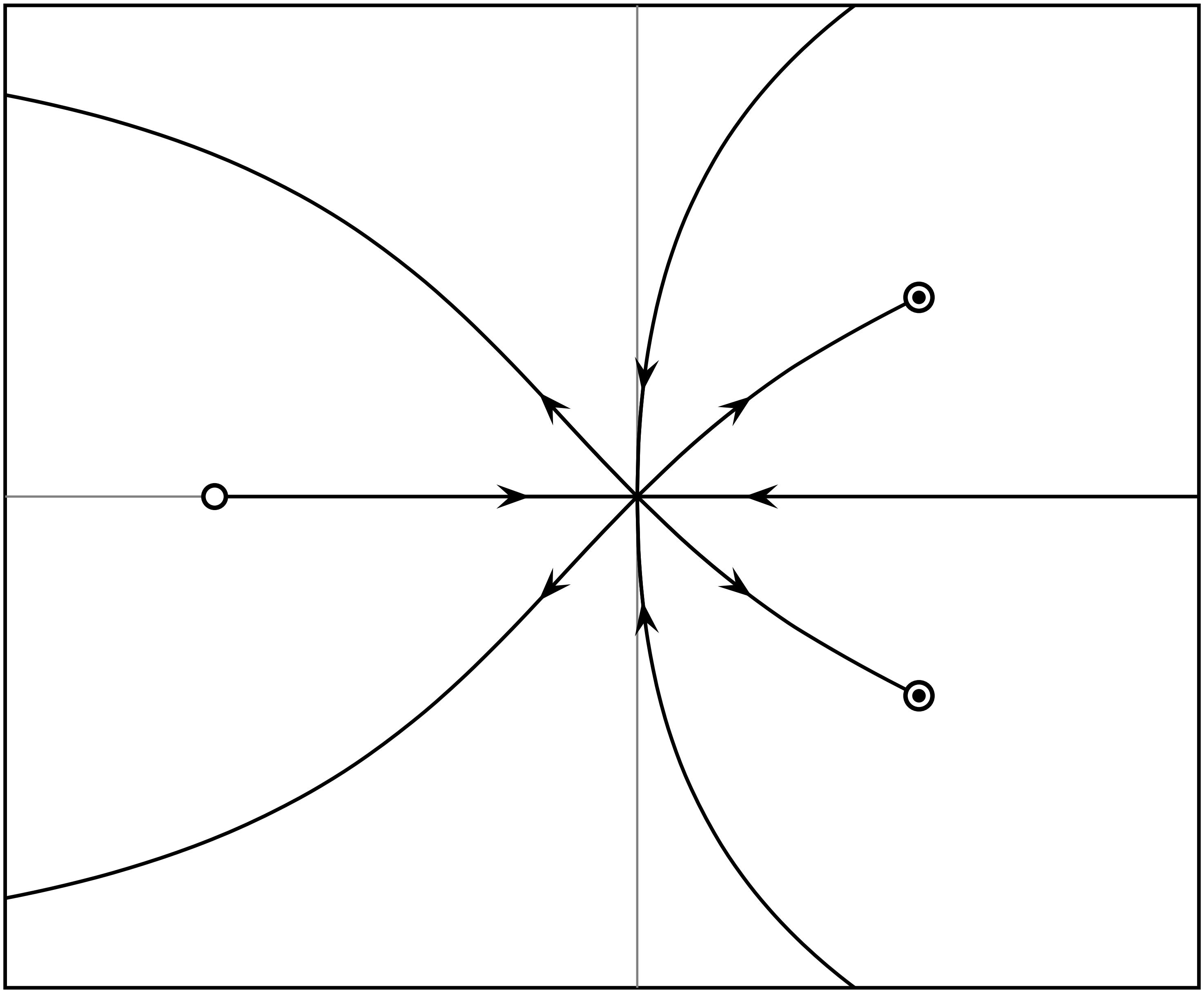
The figure on the next slide gives information on both the order star and the order arrows:



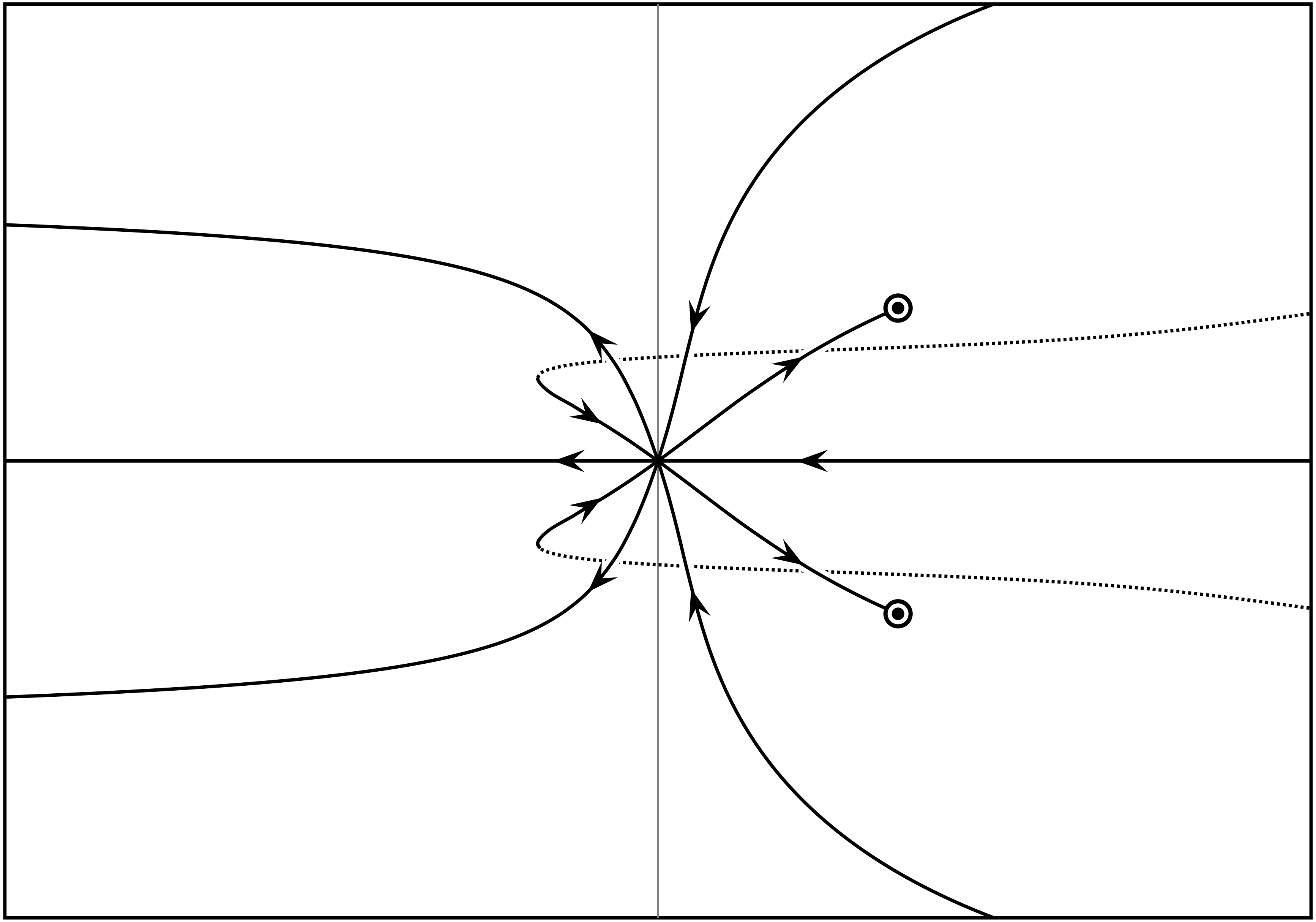
We can separate out the order star picture



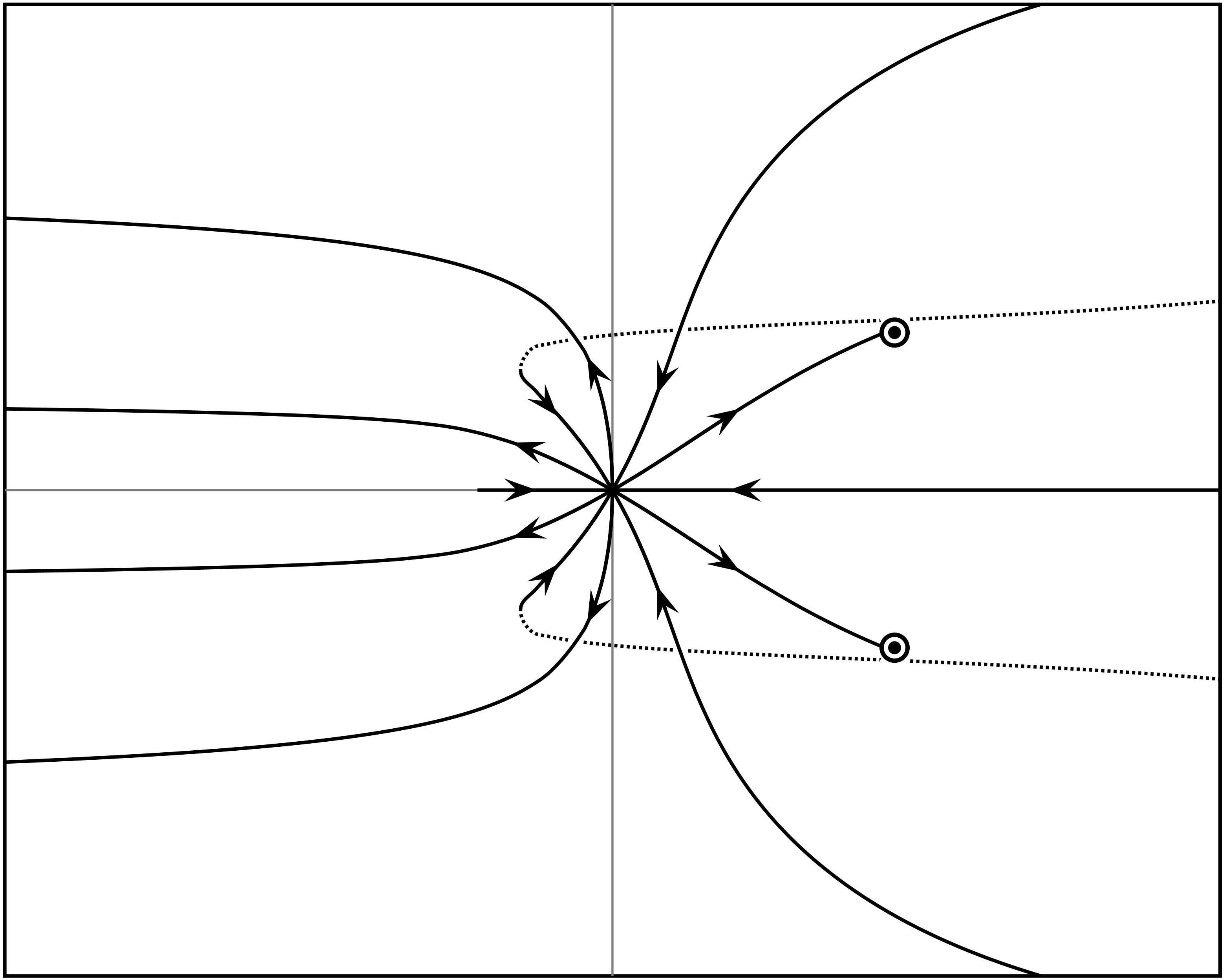
And the order arrow picture



Now consider the $[2, 0, 0, 0]$ ***approximation***



And the $[2, 0, 0, 0, 0]$ *approximation*



Order arrows and stability results

For an A-stable approximation, an upward arrow from 0 cannot cross or be tangential to the imaginary axis.

Order arrows and stability results

For an A-stable approximation, an upward arrow from 0 cannot cross or be tangential to the imaginary axis.

This is similar to the observation that, in the order star analysis, a finger cannot overlap the imaginary axis if the method is to be A-stable.

Order arrows and stability results

For an A-stable approximation, an upward arrow from 0 cannot cross or be tangential to the imaginary axis.

This is similar to the observation that, in the order star analysis, a finger cannot overlap the imaginary axis if the method is to be A-stable.

In each case we also use the behaviour near zero of the locally defined function $w(z) = 1 + Cz^{p+1}$.

The Daniel-Moore theorem

Theorem. *For an A-stable method with* n_0 *poles, the order cannot exceed* $2n_0$*.*

The Daniel-Moore theorem

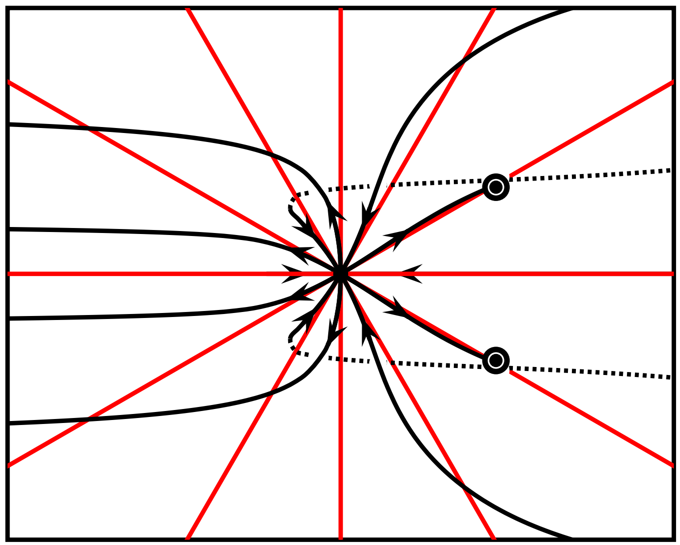
Theorem. *For an A-stable method with* n_0 *poles, the order cannot exceed* $2n_0$*.*

We illustrate how this theorem is proved by returning to the $[2, 0, 0, 0, 0]$ approximation.

The Daniel-Moore theorem

Theorem. *For an A-stable method with* n_0 *poles, the order cannot exceed* $2n_0$*.*

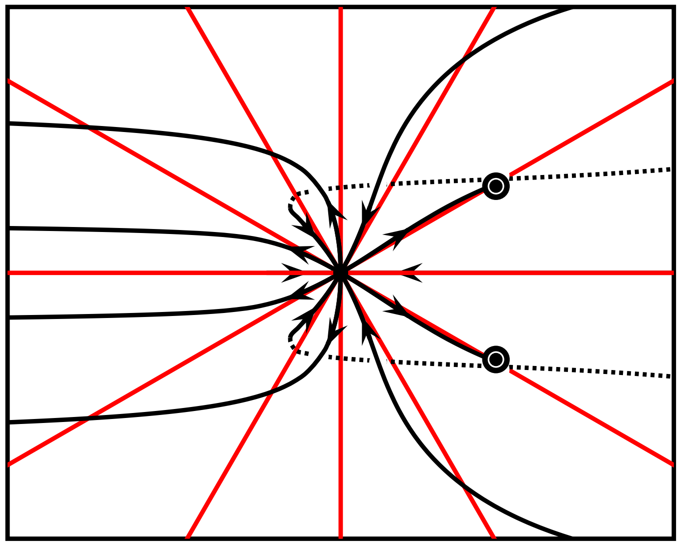
We illustrate how this theorem is proved by returning to the $[2, 0, 0, 0, 0]$ approximation.



The Daniel-Moore theorem

Theorem. *For an A-stable method with* n_0 *poles, the order cannot exceed* $2n_0$*.*

We illustrate how this theorem is proved by returning to the $[2,0,0,0,0]$ approximation.

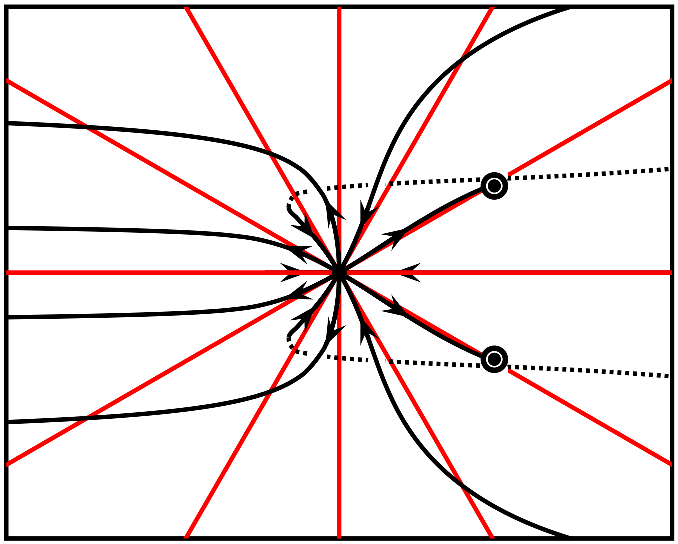


The red lines are tangent to the arrows and are spaced at angles of $\pi/(p+1) = \pi/6$.

The Daniel-Moore theorem

Theorem. *For an A-stable method with* n_0 *poles, the order cannot exceed* $2n_0$*.*

We illustrate how this theorem is proved by returning to the $[2, 0, 0, 0, 0]$ approximation.



The red lines are tangent to the arrows and are spaced at angles of $\pi/(p+1) = \pi/6$.

Hence there exist up-arrows tangent to the imaginary axis.

Padé approximations to the exponential function

We consider approximations of the form

$$w = \frac{N(z)}{D(z)} = \exp(z) + O(z^{p+1}),$$

where

$$\deg(D) = n_0, \quad \deg(N) = n_1, \quad p = n_0 + n_1.$$

Padé approximations to the exponential function

We consider approximations of the form

$$w = \frac{N(z)}{D(z)} = \exp(z) + O(z^{p+1}),$$

where

$$\deg(D) = n_0, \quad \deg(N) = n_1, \quad p = n_0 + n_1.$$

It is known that this approximation is A-stable iff

$$2n_0 - p \in \{0, 1, 2\}.$$

Padé approximations to the exponential function

We consider approximations of the form

$$w = \frac{N(z)}{D(z)} = \exp(z) + O(z^{p+1}),$$

where

$$\deg(D) = n_0, \quad \deg(N) = n_1, \quad p = n_0 + n_1.$$

It is known that this approximation is A-stable iff

$$2n_0 - p \in \{0, 1, 2\}.$$

The final step of this result, that $2n_0 - p \leq 2$ is necessary for A-stability, was proved in the famous Order Star paper of Hairer, Nørsett and Wanner.

Padé approximations to the exponential function

We consider approximations of the form

$$w = \frac{N(z)}{D(z)} = \exp(z) + O(z^{p+1}),$$

where

$$\deg(D) = n_0, \quad \deg(N) = n_1, \quad p = n_0 + n_1.$$

It is known that this approximation is A-stable iff

$$2n_0 - p \in \{0, 1, 2\}.$$

The final step of this result, that $2n_0 - p \leq 2$ is necessary for A-stability, was proved in the famous Order Star paper of Hairer, Nørsett and Wanner.
An alternative proof will be outlined using order arrows.

The Ehle theorem

Theorem. *A Padé approximation* $[n_0, n_1]$ *with order* $p = n_0 + n_1$, *is A-stable only if*

$$2n_0 - p \leq 2.$$

The Ehle theorem

Theorem. *A Padé approximation* $[n_0, n_1]$ *with order* $p = n_0 + n_1$, *is A-stable only if*

$$2n_0 - p \leq 2.$$

Some of the up-arrows from zero terminate at poles and some terminate at $-\infty$ in the sense that the real part has this limit and the imaginary part has a finite limit.

The Ehle theorem

Theorem. *A Padé approximation* $[n_0, n_1]$ *with order* $p = n_0 + n_1$, *is A-stable only if*

$$2n_0 - p \leq 2.$$

Some of the up-arrows from zero terminate at poles and some terminate at $-\infty$ in the sense that the real part has this limit and the imaginary part has a finite limit.

We will assume that each of the poles is a termination point for an up-arrow from zero.

The Ehle theorem

Theorem. *A Padé approximation* $[n_0, n_1]$ *with order* $p = n_0 + n_1$, *is A-stable only if*

$$2n_0 - p \leq 2.$$

Some of the up-arrows from zero terminate at poles and some terminate at $-\infty$ in the sense that the real part has this limit and the imaginary part has a finite limit.

We will assume that each of the poles is a termination point for an up-arrow from zero.

This question will be discussed later.

Because adjacent up-arrows subtend an angle

$$\frac{2\pi}{p+1}$$

and n_0 of them terminate at poles, the total angle subtended is at least

$$\frac{2(n_0-1)}{p+1}\pi \geq \pi \quad \text{if} \quad 2n_0 - p > 2.$$

Because adjacent up-arrows subtend an angle

$$\frac{2\pi}{p+1}$$

and n_0 of them terminate at poles, the total angle subtended is at least

$$\frac{2(n_0-1)}{p+1}\pi \geq \pi \quad \text{if} \quad 2n_0 - p > 2.$$

Hence, either up-arrows terminating at poles are tangential to the imaginary axis or protrude into the left half-plane.

Because adjacent up-arrows subtend an angle

$$\frac{2\pi}{p+1}$$

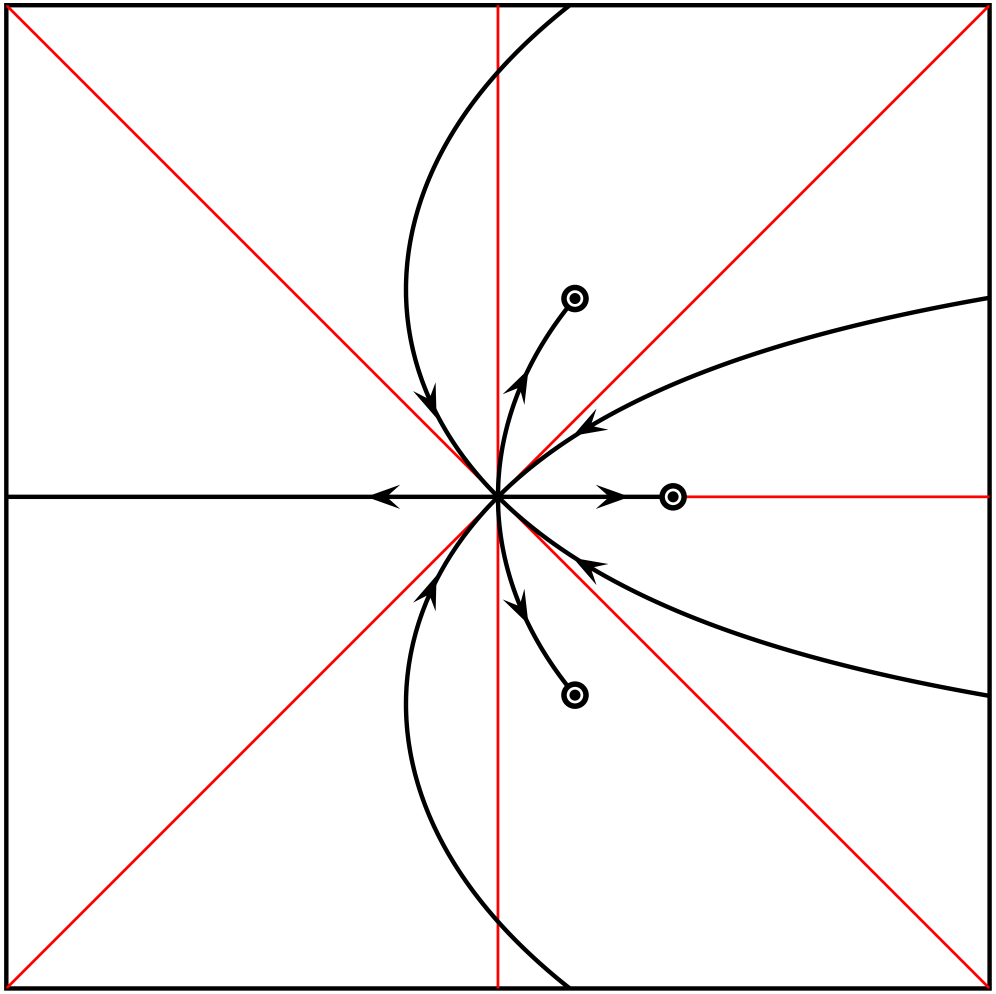
and n_0 of them terminate at poles, the total angle subtended is at least

$$\frac{2(n_0-1)}{p+1}\pi \geq \pi \quad \text{if} \quad 2n_0 - p > 2.$$

Hence, either up-arrows terminating at poles are tangential to the imaginary axis or protrude into the left half-plane.

In the latter case, there are poles in the left half-plane or an up-arrow crosses back across the imaginary axis.

We will illustrate this result in the $[3, 0]$ case.



Now return to a crucial part of the proof:

Why should every pole be at the end of an up-arrow from zero?

Now return to a crucial part of the proof:

Why should every pole be at the end of an up-arrow from zero?

For Padé approximations this follows simply from the fact that up-arrows from zero and down-arrows from zero cannot cross.

Now return to a crucial part of the proof:

Why should every pole be at the end of an up-arrow from zero?

For Padé approximations this follows simply from the fact that up-arrows from zero and down-arrows from zero cannot cross.

But in the general case, where everything happens on a Riemann surface, we cannot use this argument in a simple way.

Now return to a crucial part of the proof:

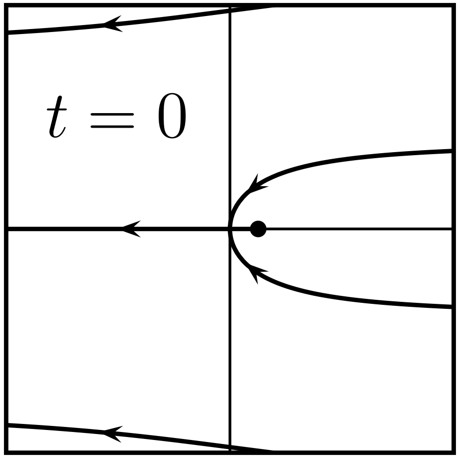
Why should every pole be at the end of an up-arrow from zero?

For Padé approximations this follows simply from the fact that up-arrows from zero and down-arrows from zero cannot cross.

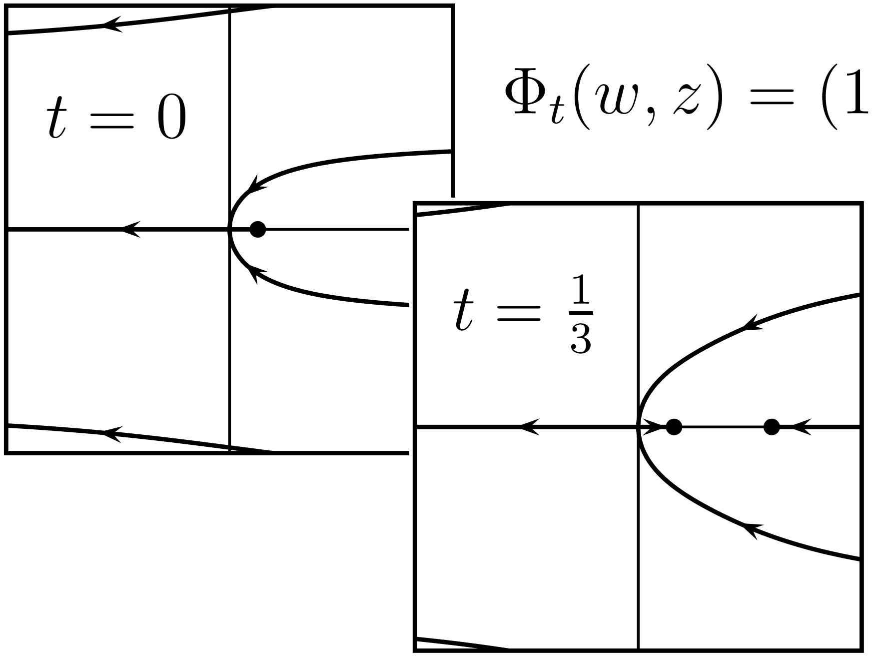
But in the general case, where everything happens on a Riemann surface, we cannot use this argument in a simple way.

We will look to see if homotopy might be a useful approach.

$$\begin{aligned}\Phi_t(w,z) &= (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)\\ &= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)\end{aligned}$$

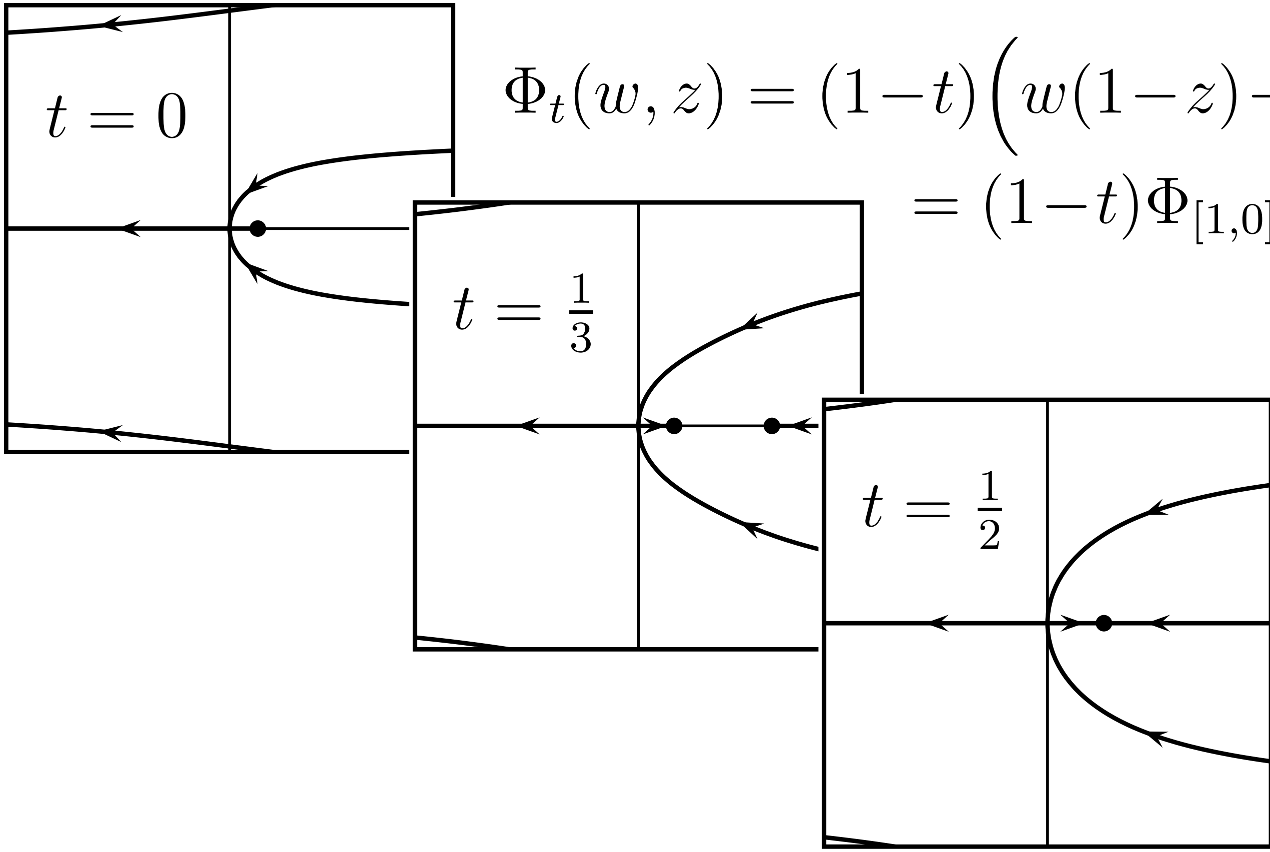


$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$

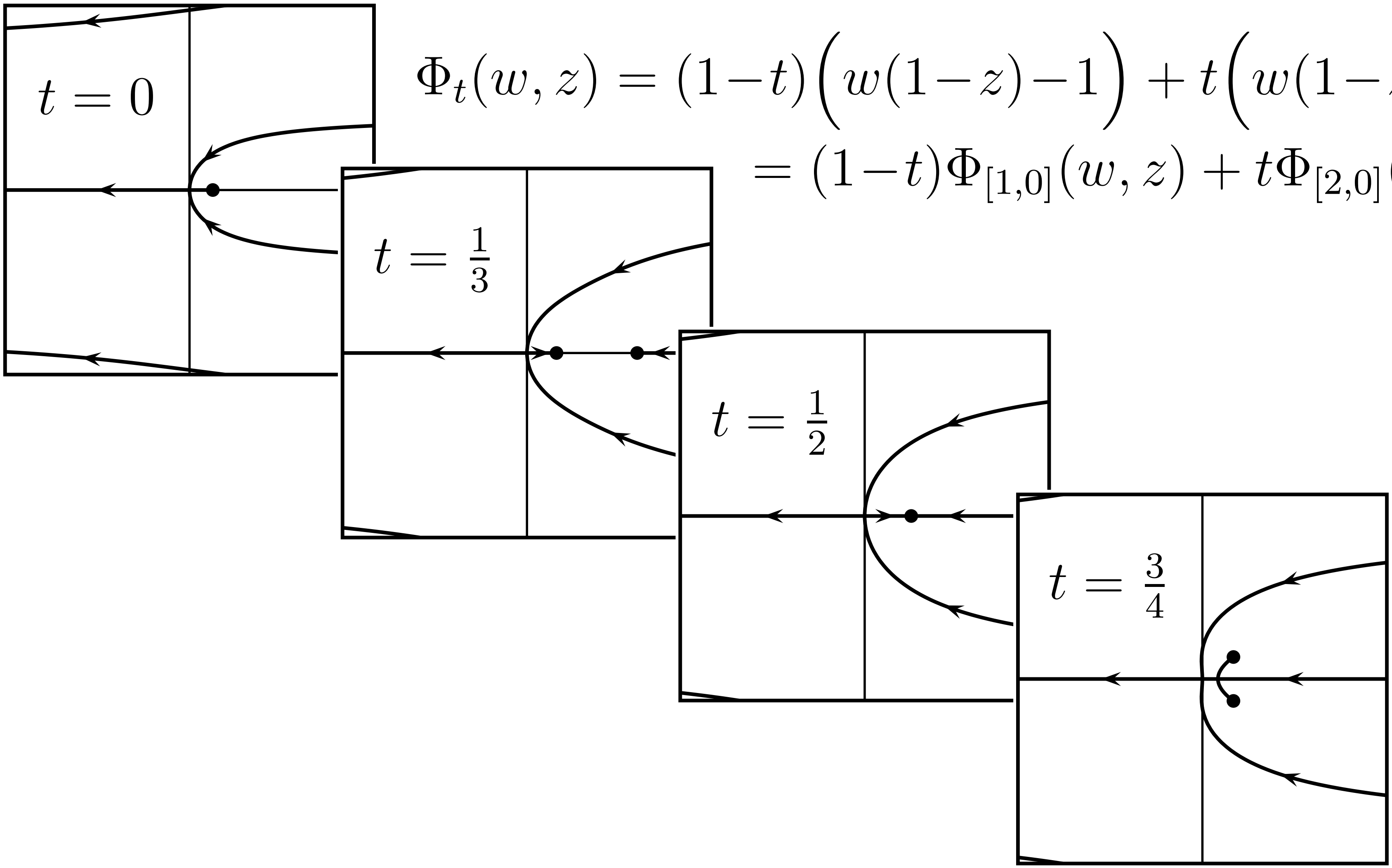


$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$

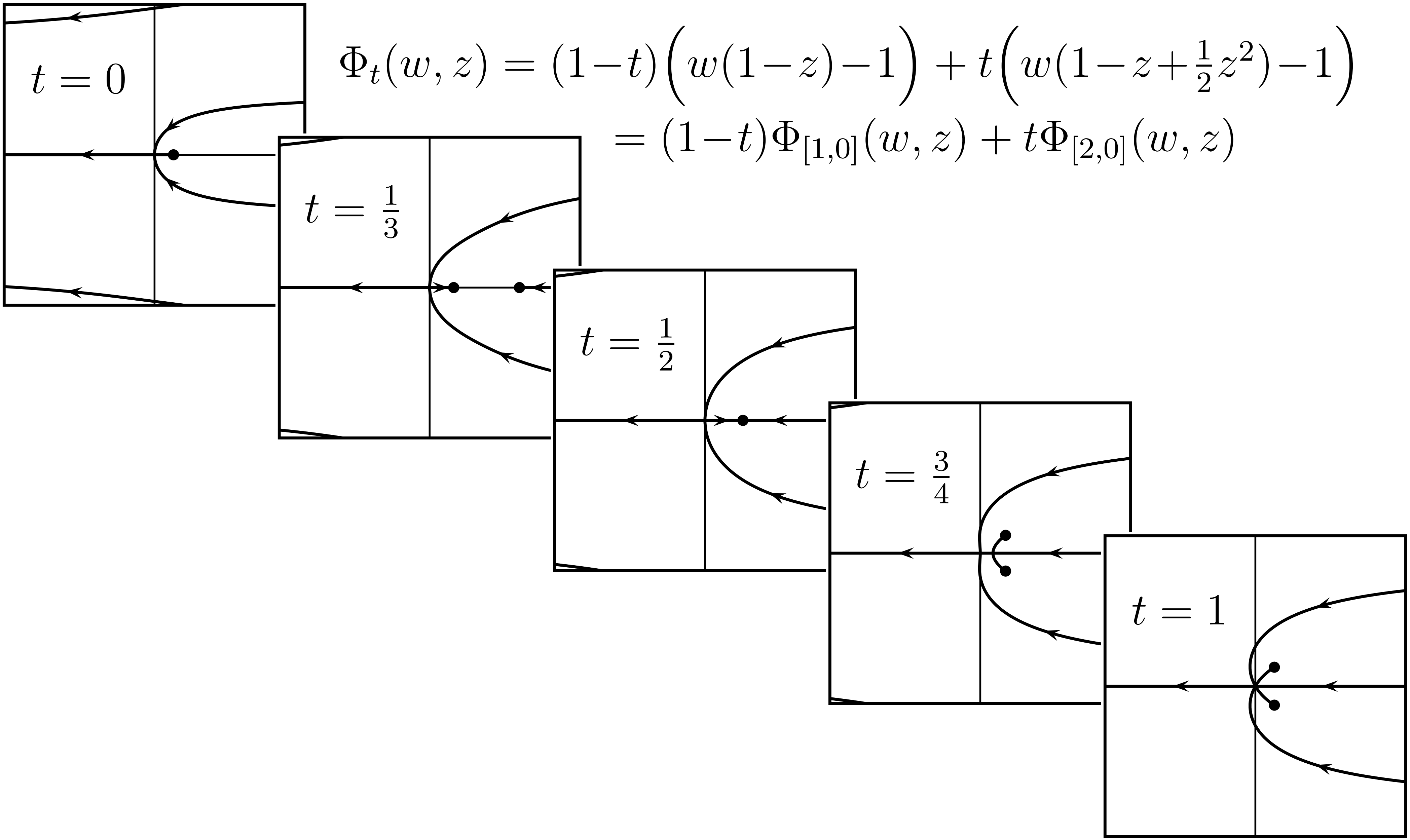
$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$



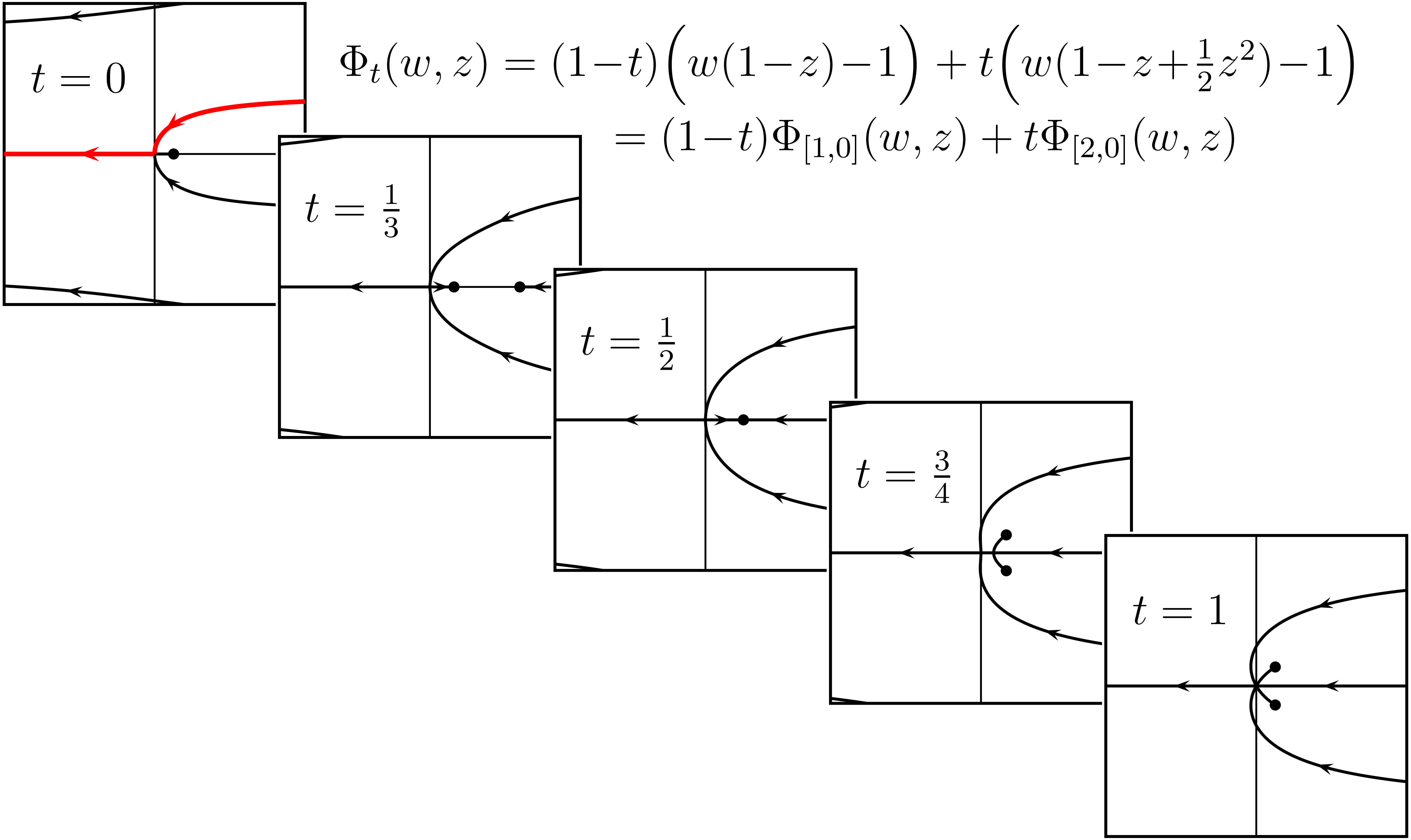
$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$



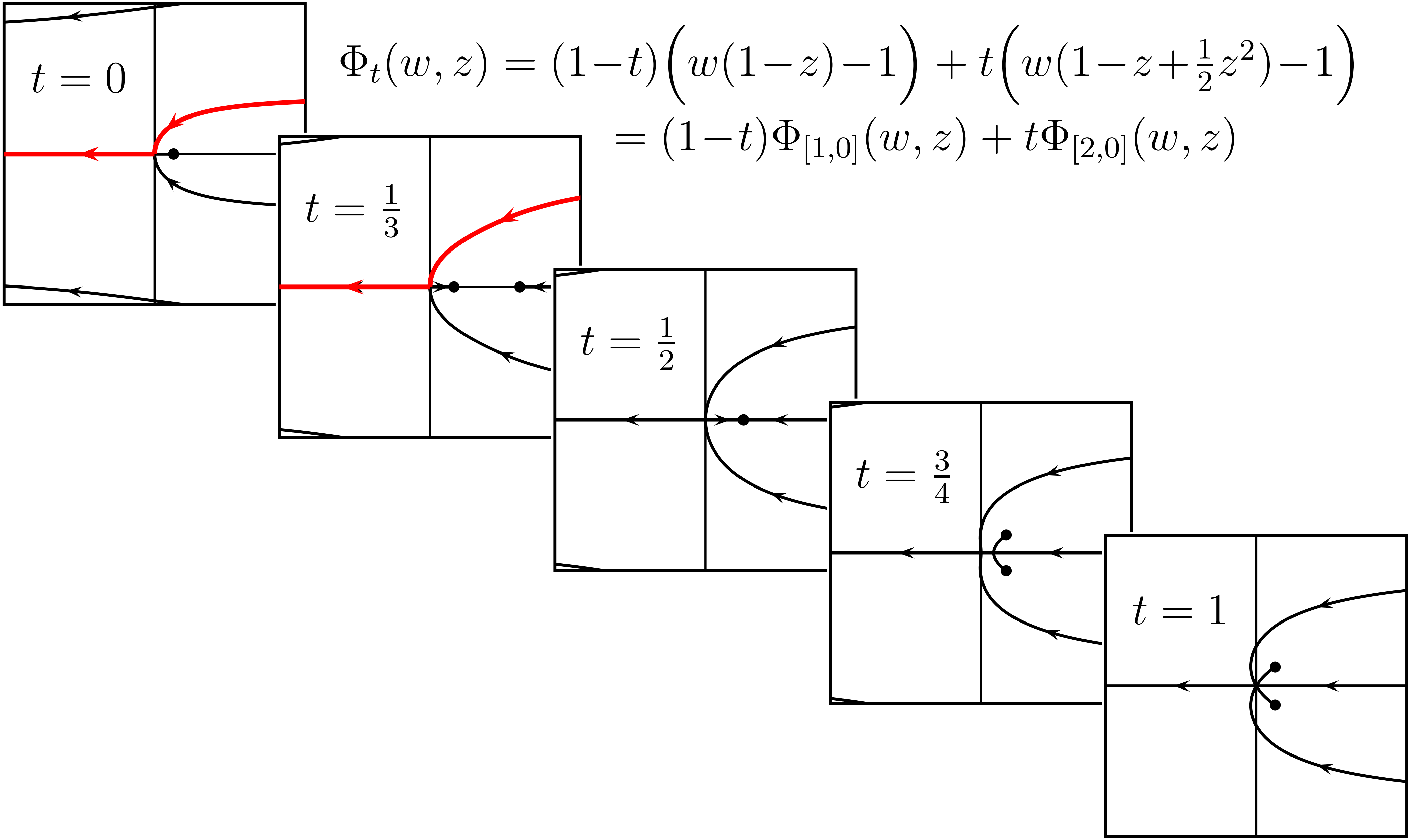
$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$



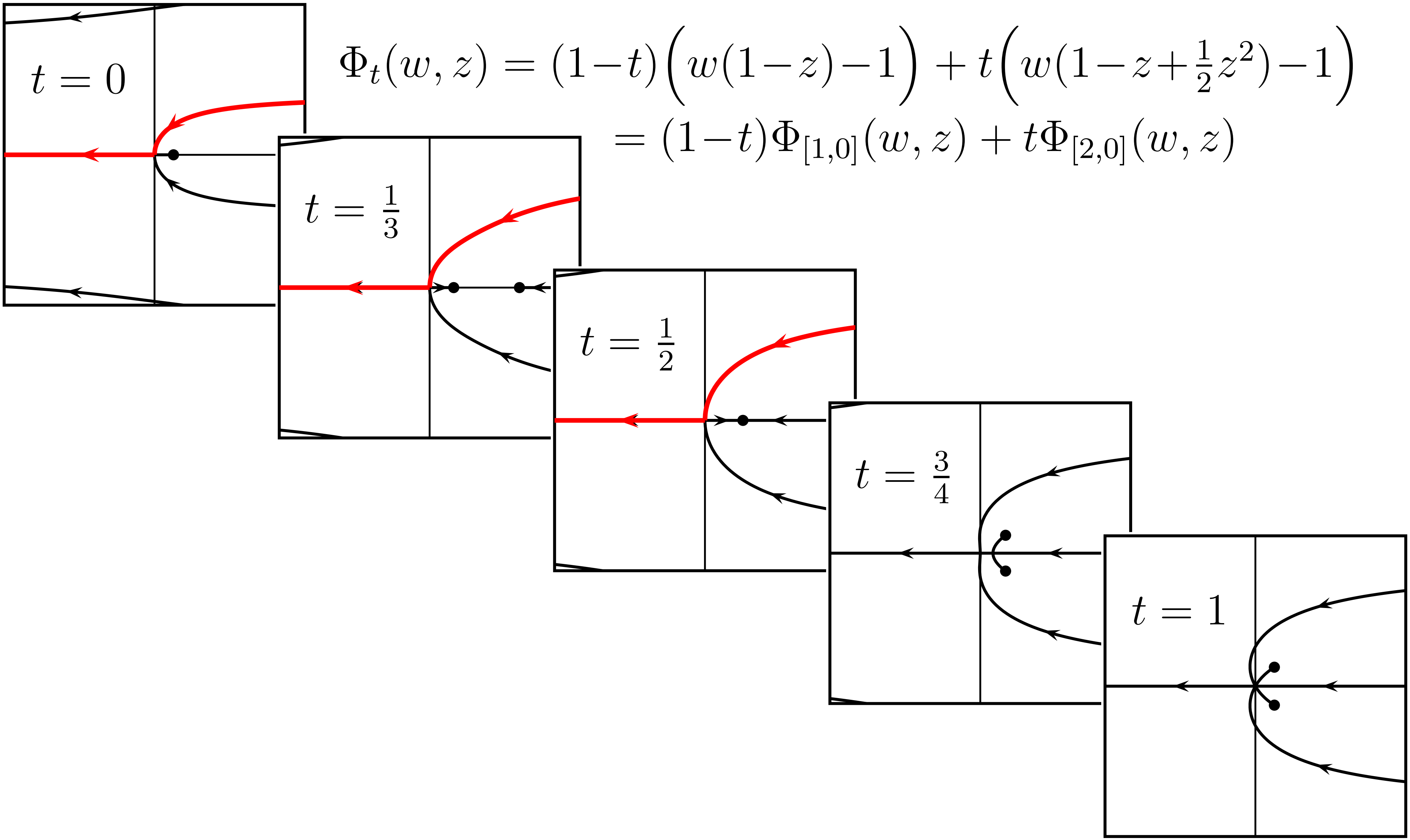
$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$



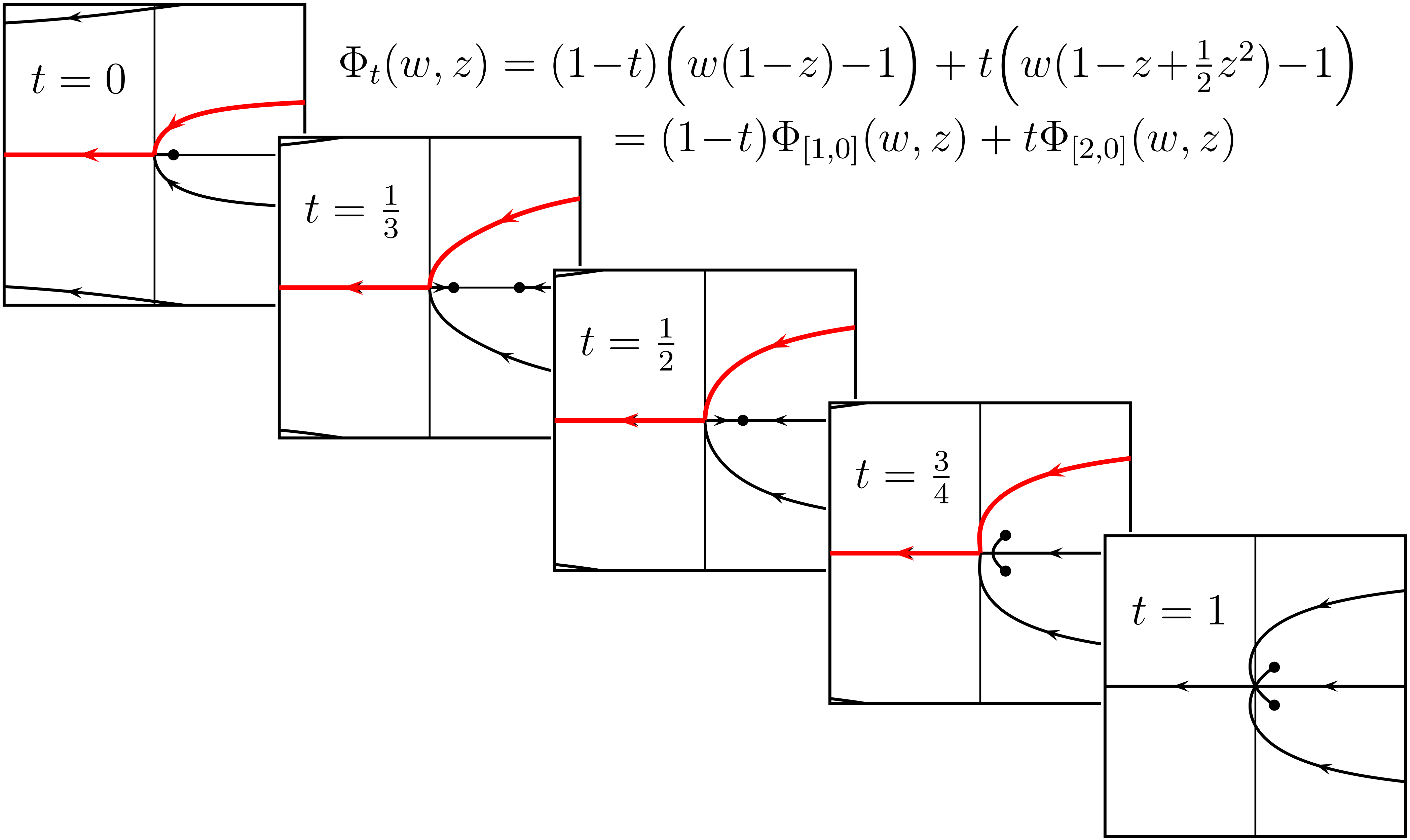
$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$



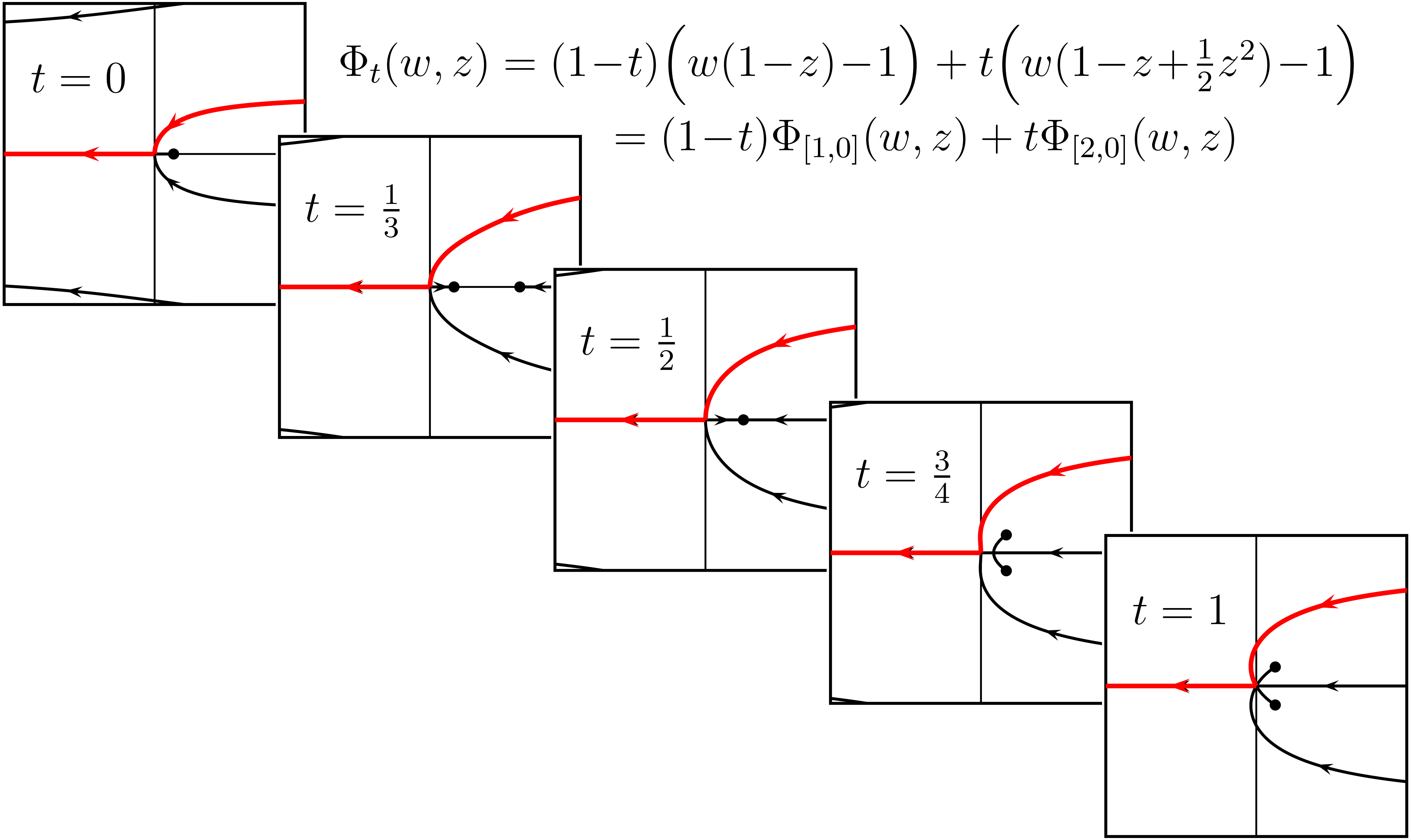
$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$



$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$



$$\Phi_t(w,z) = (1-t)\Big(w(1-z)-1\Big) + t\Big(w(1-z+\tfrac{1}{2}z^2)-1\Big)$$
$$= (1-t)\Phi_{[1,0]}(w,z) + t\Phi_{[2,0]}(w,z)$$



Quadratic approximations to the exponential

Given $[n_0, n_1, n_2]$, there exist polynomials $P_i, i = 0, 1, 2$, of degrees n_0, n_1, n_2 respectively, such that

$$\exp(2z)P_0(z) + \exp(z)P_1(z) + P_2(z) = O(z^{p+1}),$$

where $p = n_0 + n_1 + n_2 + 1$.

Quadratic approximations to the exponential

Given $[n_0, n_1, n_2]$, there exist polynomials $P_i, i = 0, 1, 2$, of degrees n_0, n_1, n_2 respectively, such that

$$\exp(2z)P_0(z) + \exp(z)P_1(z) + P_2(z) = O(z^{p+1}),$$

where $p = n_0 + n_1 + n_2 + 1$.

Theorem. *Let*

$$(1+t)^{-n_1-1}(2+t)^{-n_2-1} = a_0 + a_1 t + \cdots + a_{n_0} t^{n_0} + O(t^{n_0+1}),$$

$$(-1+t)^{-n_0-1}(1+t)^{-n_2-1} = b_0 + b_1 t + \cdots + b_{n_1} t^{n_1} + O(t^{n_1+1}),$$

$$(-2+t)^{-n_0-1}(-1+t)^{-n_1-1} = c_0 + c_1 t + \cdots + c_{n_2} t^{n_2} + O(t^{n_2+1}),$$

then (P_0, P_1, P_2) *is the* $[n_0, n_1, n_2]$ *quadratic Padé approximation to* $\exp$ *if*

Quadratic approximations to the exponential

Given $[n_0, n_1, n_2]$, there exist polynomials $P_i, i = 0, 1, 2,$ of degrees n_0, n_1, n_2 respectively, such that

$$\exp(2z)P_0(z) + \exp(z)P_1(z) + P_2(z) = O(z^{p+1}),$$

where $p = n_0 + n_1 + n_2 + 1$.

Theorem. *Let*

$$(1+t)^{-n_1-1}(2+t)^{-n_2-1} = a_0 + a_1 t + \cdots + a_{n_0} t^{n_0} + O(t^{n_0+1}),$$

$$(-1+t)^{-n_0-1}(1+t)^{-n_2-1} = b_0 + b_1 t + \cdots + b_{n_1} t^{n_1} + O(t^{n_1+1}),$$

$$(-2+t)^{-n_0-1}(-1+t)^{-n_1-1} = c_0 + c_1 t + \cdots + c_{n_2} t^{n_2} + O(t^{n_2+1}),$$

then (P_0, P_1, P_2) *is the* $[n_0, n_1, n_2]$ *quadratic Padé approximation to* $\exp$ *if*

$$P_0(z) = \sum_{i=0}^{n_0} \frac{a_{n_0-i} z^i}{i!},$$

Quadratic approximations to the exponential

Given $[n_0, n_1, n_2]$, there exist polynomials $P_i, i = 0, 1, 2$, of degrees n_0, n_1, n_2 respectively, such that

$$\exp(2z)P_0(z) + \exp(z)P_1(z) + P_2(z) = O(z^{p+1}),$$

where $p = n_0 + n_1 + n_2 + 1$.

Theorem. *Let*

$$(1+t)^{-n_1-1}(2+t)^{-n_2-1} = a_0 + a_1 t + \cdots + a_{n_0} t^{n_0} + O(t^{n_0+1}),$$

$$(-1+t)^{-n_0-1}(1+t)^{-n_2-1} = b_0 + b_1 t + \cdots + b_{n_1} t^{n_1} + O(t^{n_1+1}),$$

$$(-2+t)^{-n_0-1}(-1+t)^{-n_1-1} = c_0 + c_1 t + \cdots + c_{n_2} t^{n_2} + O(t^{n_2+1}),$$

then (P_0, P_1, P_2) *is the* $[n_0, n_1, n_2]$ *quadratic Padé approximation to* $\exp$ *if*

$$P_0(z) = \sum_{i=0}^{n_0} \frac{a_{n_0-i} z^i}{i!}, \qquad P_1(z) = \sum_{i=0}^{n_1} \frac{b_{n_1-i} z^i}{i!},$$

Quadratic approximations to the exponential

Given $[n_0, n_1, n_2]$, there exist polynomials $P_i, i = 0, 1, 2$, of degrees n_0, n_1, n_2 respectively, such that

$$\exp(2z)P_0(z) + \exp(z)P_1(z) + P_2(z) = O(z^{p+1}),$$

where $p = n_0 + n_1 + n_2 + 1$.

Theorem. *Let*

$$(1+t)^{-n_1-1}(2+t)^{-n_2-1} = a_0 + a_1 t + \cdots + a_{n_0} t^{n_0} + O(t^{n_0+1}),$$

$$(-1+t)^{-n_0-1}(1+t)^{-n_2-1} = b_0 + b_1 t + \cdots + b_{n_1} t^{n_1} + O(t^{n_1+1}),$$

$$(-2+t)^{-n_0-1}(-1+t)^{-n_1-1} = c_0 + c_1 t + \cdots + c_{n_2} t^{n_2} + O(t^{n_2+1}),$$

then (P_0, P_1, P_2) *is the* $[n_0, n_1, n_2]$ *quadratic Padé approximation to* exp *if*

$$P_0(z) = \sum_{i=0}^{n_0} \frac{a_{n_0-i} z^i}{i!}, \quad P_1(z) = \sum_{i=0}^{n_1} \frac{b_{n_1-i} z^i}{i!}, \quad P_2(z) = \sum_{i=0}^{n_2} \frac{c_{n_2-i} z^i}{i!}.$$

The Butcher–Chipman conjecture

After extensive searching, Fred Chipman and I formulated the following audacious statement:
For generalized Padé approximations to the exponential function, the necessary condition for A-stability of linear Padé approximations also holds in the general case:

$$2n_0 - p \in \{0, 1, 2\}.$$

The Butcher–Chipman conjecture

After extensive searching, Fred Chipman and I formulated the following audacious statement:
For generalized Padé approximations to the exponential function, the necessary condition for A-stability of linear Padé approximations also holds in the general case:

$$2n_0 - p \in \{0, 1, 2\}.$$

- In the linear case this is also sufficient

The Butcher–Chipman conjecture

After extensive searching, Fred Chipman and I formulated the following audacious statement:
For generalized Padé approximations to the exponential function, the necessary condition for A-stability of linear Padé approximations also holds in the general case:

$$2n_0 - p \in \{0, 1, 2\}.$$

- In the linear case this is also sufficient
- But there are counterexamples in the general case

The Butcher–Chipman conjecture

After extensive searching, Fred Chipman and I formulated the following audacious statement:
For generalized Padé approximations to the exponential function, the necessary condition for A-stability of linear Padé approximations also holds in the general case:

$$\boldsymbol{2n_0 - p \in \{0, 1, 2\}.}$$

- In the linear case this is also sufficient
- But there are counterexamples in the general case
- $2n_0 - p \geq 0$ follows from the Daniel-Moore theorem

The Butcher–Chipman conjecture

After extensive searching, Fred Chipman and I formulated the following audacious statement:

For generalized Padé approximations to the exponential function, the necessary condition for A-stability of linear Padé approximations also holds in the general case:

$$2n_0 - p \in \{0, 1, 2\}.$$

- In the linear case this is also sufficient
- But there are counterexamples in the general case
- $2n_0 - p \geq 0$ follows from the Daniel-Moore theorem
- This leaves $2n_0 - p \leq 2$ as the remaining challenge

The Butcher–Chipman conjecture

After extensive searching, Fred Chipman and I formulated the following audacious statement:
For generalized Padé approximations to the exponential function, the necessary condition for A-stability of linear Padé approximations also holds in the general case:

$$\boldsymbol{2n_0 - p \in \{0, 1, 2\}.}$$

- In the linear case this is also sufficient
- But there are counterexamples in the general case
- $2n_0 - p \geq 0$ follows from the Daniel-Moore theorem
- This leaves $2n_0 - p \leq 2$ as the remaining challenge
- I will concentrate on the quadratic case

We are only interested in "genuine" order. For example, consider the approximations

$$\Phi_{[2,0,2]}(w,z) = w^2\left(1-\tfrac{5}{8}z+\tfrac{1}{8}z^2\right)-2w+\left(1+\tfrac{5}{8}z+\tfrac{1}{8}z^2\right), \quad (1)$$

$$\Phi_{[1,0,1]}(w,z) = w^2\left(1-\tfrac{1}{2}z\right)-2w+\left(1+\tfrac{1}{2}z\right). \quad (2)$$

We are only interested in “genuine” order. For example, consider the approximations

$$\Phi_{[2,0,2]}(w,z)=w^2(1-\tfrac{5}{8}z+\tfrac{1}{8}z^2)-2w+(1+\tfrac{5}{8}z+\tfrac{1}{8}z^2), \quad (1)$$

$$\Phi_{[1,0,1]}(w,z)=w^2(1-\tfrac{1}{2}z)-2w+(1+\tfrac{1}{2}z). \quad (2)$$

(1) is order 5 in the sense that $\Phi(\exp(z),z)=O(z^6)$, but there is a branch-point at $z=0$, and the order seems to be shared between two sheets of the Riemann surface.

We are only interested in “genuine” order. For example, consider the approximations

$$\Phi_{[2,0,2]}(w,z) = w^2(1-\tfrac{5}{8}z+\tfrac{1}{8}z^2)-2w+(1+\tfrac{5}{8}z+\tfrac{1}{8}z^2), \quad (1)$$

$$\Phi_{[1,0,1]}(w,z) = w^2(1-\tfrac{1}{2}z)-2w+(1+\tfrac{1}{2}z). \quad (2)$$

(1) is order 5 in the sense that $\Phi(\exp(z),z) = O(z^6)$, but there is a branch-point at $z = 0$, and the order seems to be shared between two sheets of the Riemann surface.

(2) factorizes into the product of the zero order approximation $w-1$ and the order 2 approximation $w(1-\frac{1}{2}z)-(1+\frac{1}{2}z)$.

We are only interested in “genuine” order. For example, consider the approximations

$$\Phi_{[2,0,2]}(w,z) = w^2(1-\tfrac{5}{8}z+\tfrac{1}{8}z^2)-2w+(1+\tfrac{5}{8}z+\tfrac{1}{8}z^2), \quad (1)$$

$$\Phi_{[1,0,1]}(w,z) = w^2(1-\tfrac{1}{2}z)-2w+(1+\tfrac{1}{2}z). \quad (2)$$

(1) is order 5 in the sense that $\Phi(\exp(z),z) = O(z^6)$, but there is a branch-point at $z=0$, and the order seems to be shared between two sheets of the Riemann surface.

(2) factorizes into the product of the zero order approximation $w-1$ and the order 2 approximation $w(1-\frac{1}{2}z)-(1+\frac{1}{2}z)$.

To make sure that we never deal with such irrelevancies, we will consider only cases for which $n_0 > n_1 + n_2$.

The final step of the proof of the BC conjecture would be exactly the same as for the Ehle theorem if we could first prove that every pole is at the end of an up-arrow from zero on the “principal sheet”.

The final step of the proof of the BC conjecture would be exactly the same as for the Ehle theorem if we could first prove that every pole is at the end of an up-arrow from zero on the “principal sheet”.

We will use homotopy to connect an order p approximation $[n_0, n_1, n_2]$ to the order $p-1$ approximation $[n_0, n_1, n_2 - 1]$.

The final step of the proof of the BC conjecture would be exactly the same as for the Ehle theorem if we could first prove that every pole is at the end of an up-arrow from zero on the “principal sheet”.

We will use homotopy to connect an order p approximation $[n_0, n_1, n_2]$ to the order $p-1$ approximation $[n_0, n_1, n_2 - 1]$.

This will be illustrated by a single example:

$$\Phi_t = (1-t)\Phi_{[2,1,-1]} + t\Phi_{[2,1,0]},$$

where

$$\Phi_{[2,1,-1]}(w,z) = w^2(1 - \tfrac{2}{3}z + \tfrac{1}{6}z^2) - w(1 + \tfrac{1}{3}z),$$

$$\Phi_{[2,1,0]}(w,z) = w^2(1 - \tfrac{10}{17}z + \tfrac{2}{17}z^2) - w(\tfrac{16}{17} + \tfrac{8}{17}z) - \tfrac{1}{17}.$$

As soon as t becomes positive, the zero at -3 is replaced by a double branch point which breaks into two complex conjugate branch points.

As soon as t becomes positive, the zero at -3 is replaced by a double branch point which breaks into two complex conjugate branch points. Also, as t becomes positive, a stagnation point suddenly appears at $z = -4$.

As soon as t becomes positive, the zero at -3 is replaced by a double branch point which breaks into two complex conjugate branch points. Also, as t becomes positive, a stagnation point suddenly appears at $z = -4$. To see how this happens, we find the w-resultant of $\Phi_t(w, z)$ and

$$\left(\frac{\partial}{\partial z} + w\frac{\partial}{\partial w}\right)\Phi_t(w, z).$$

This has a factor z^3, corresponding to the order 3 stagnation point at 0, and the additional factor

$$(289 + 238t - 95t^2)z + (1156 - 1088t - 68t^2),$$

As soon as t becomes positive, the zero at -3 is replaced by a double branch point which breaks into two complex conjugate branch points. Also, as t becomes positive, a stagnation point suddenly appears at $z = -4$. To see how this happens, we find the w-resultant of $\Phi_t(w, z)$ and

$$(\frac{\partial}{\partial z} + w\frac{\partial}{\partial w})\Phi_t(w, z).$$

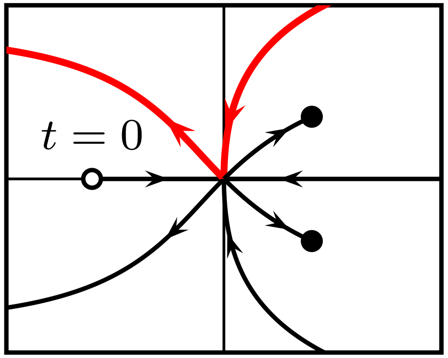
This has a factor z^3, corresponding to the order 3 stagnation point at 0, and the additional factor

$$(289 + 238t - 95t^2)z + (1156 - 1088t - 68t^2),$$

with zero at $z = -4$, when $t = 0$, and at $z = 0$, when $t = 1$.

$$\Phi_{[2,1,-1]}(w,z) = w^2(1 - \tfrac{2}{3}z + \tfrac{1}{6}z^2) - w(1 + \tfrac{1}{3}z)$$
$$\Phi_{[2,1,0]}(w,z) = w^2(1 - \tfrac{10}{17}z + \tfrac{2}{17}z^2) - w(\tfrac{16}{17} + \tfrac{8}{17}z) - \tfrac{1}{17}$$

$$\Phi_t = (1-t)\Phi_{[2,1,-1]} + t\Phi_{[2,1,0]}$$



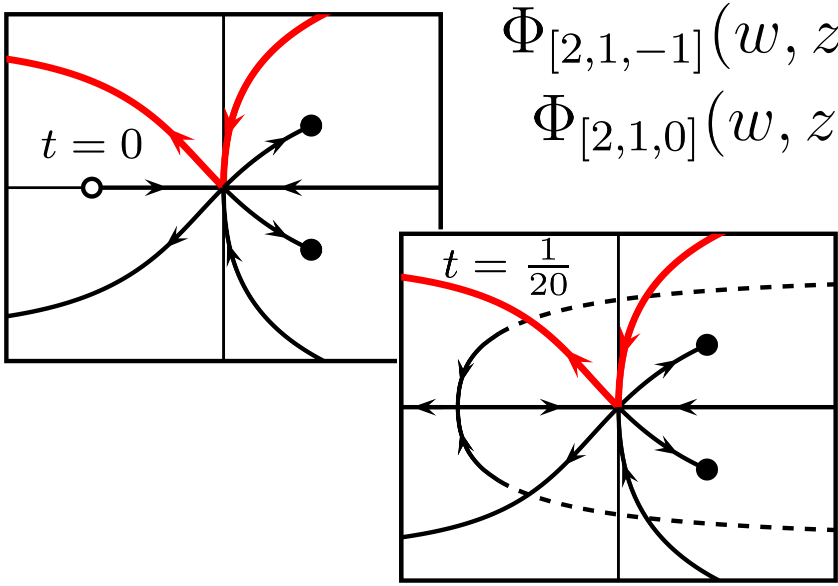
$$\Phi_{[2,1,-1]}(w,z) = w^2(1 - \tfrac{2}{3}z + \tfrac{1}{6}z^2) - w(1 + \tfrac{1}{3}z)$$

$$\Phi_{[2,1,0]}(w,z) = w^2(1 - \tfrac{10}{17}z + \tfrac{2}{17}z^2) - w(\tfrac{16}{17} + \tfrac{8}{17}z) - \tfrac{1}{17}$$

$$\Phi_t = (1-t)\Phi_{[2,1,-1]} + t\Phi_{[2,1,0]}$$

$$\Phi_{[2,1,-1]}(w,z) = w^2(1 - \tfrac{2}{3}z + \tfrac{1}{6}z^2) - w(1 + \tfrac{1}{3}z)$$

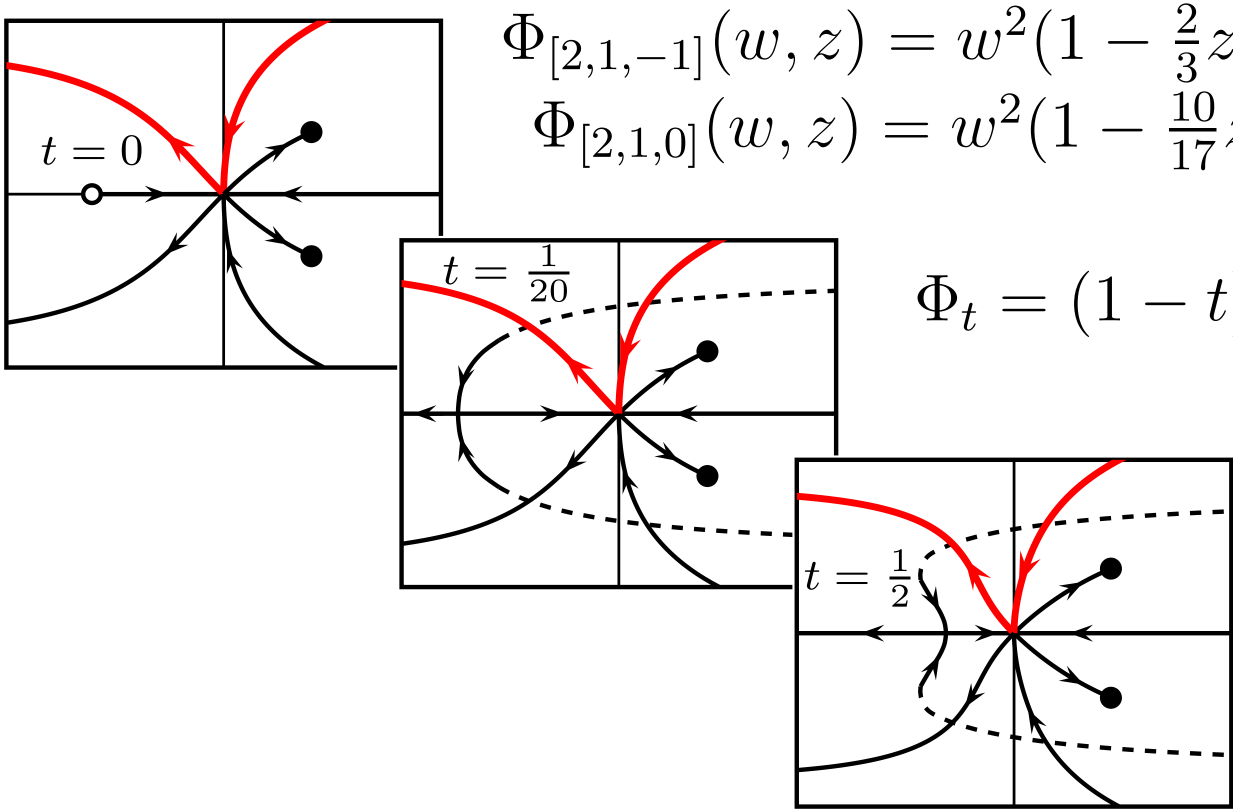
$$\Phi_{[2,1,0]}(w,z) = w^2(1 - \tfrac{10}{17}z + \tfrac{2}{17}z^2) - w(\tfrac{16}{17} + \tfrac{8}{17}z) - \tfrac{1}{17}$$



$$\Phi_t = (1-t)\Phi_{[2,1,-1]} + t\Phi_{[2,1,0]}$$

$$\Phi_{[2,1,-1]}(w,z) = w^2(1 - \tfrac{2}{3}z + \tfrac{1}{6}z^2) - w(1 + \tfrac{1}{3}z)$$

$$\Phi_{[2,1,0]}(w,z) = w^2(1 - \tfrac{10}{17}z + \tfrac{2}{17}z^2) - w(\tfrac{16}{17} + \tfrac{8}{17}z) - \tfrac{1}{17}$$

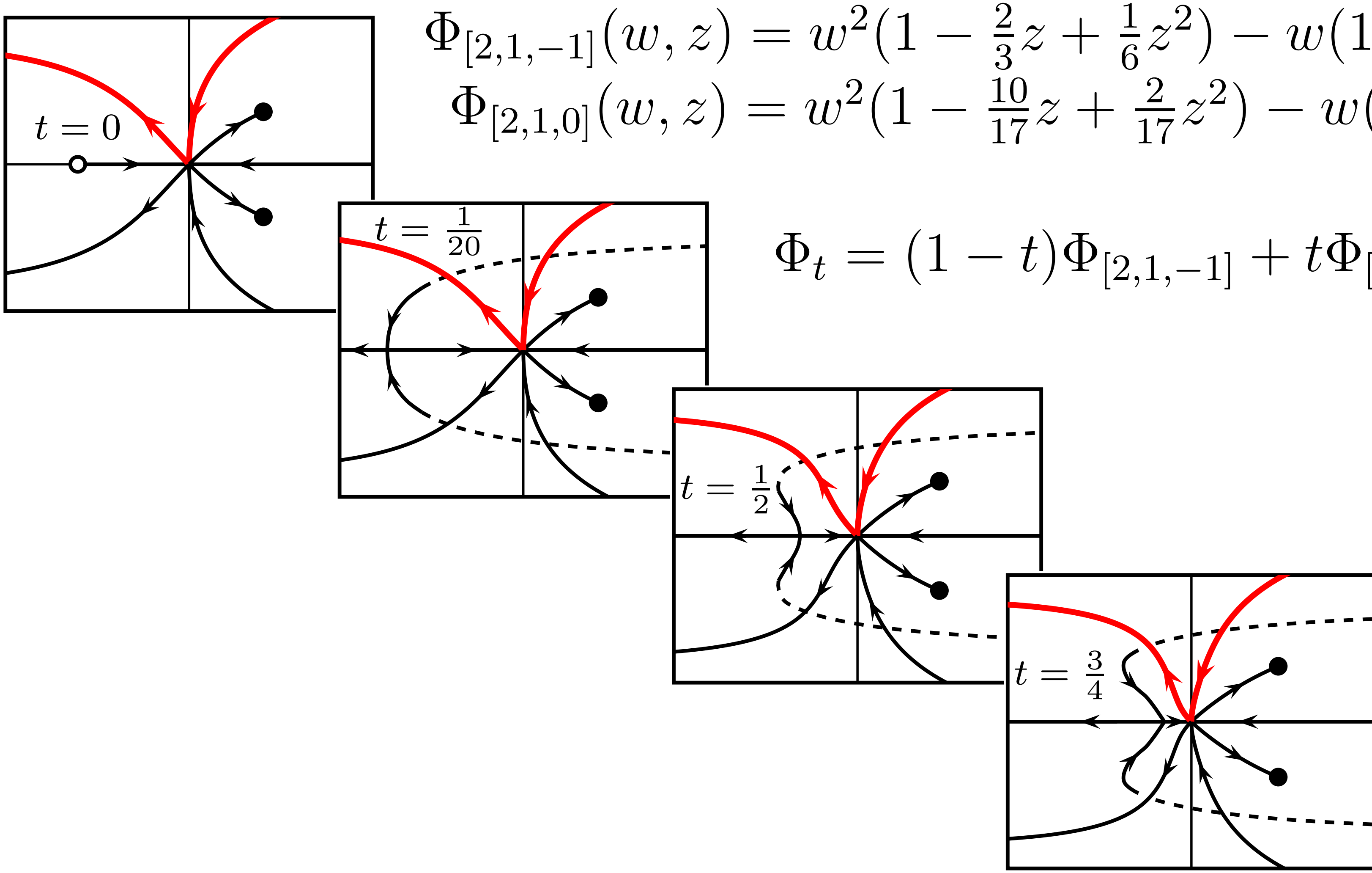


$$\Phi_t = (1-t)\Phi_{[2,1,-1]} + t\Phi_{[2,1,0]}$$

$$\Phi_{[2,1,-1]}(w,z) = w^2\left(1 - \tfrac{2}{3}z + \tfrac{1}{6}z^2\right) - w\left(1 + \tfrac{1}{3}z\right)$$

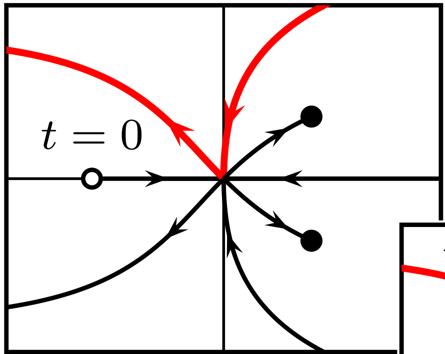
$$\Phi_{[2,1,0]}(w,z) = w^2\left(1 - \tfrac{10}{17}z + \tfrac{2}{17}z^2\right) - w\left(\tfrac{16}{17} + \tfrac{8}{17}z\right) - \tfrac{1}{17}$$

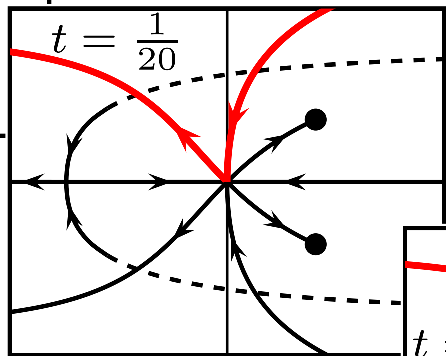
$$\Phi_t = (1-t)\Phi_{[2,1,-1]} + t\Phi_{[2,1,0]}$$



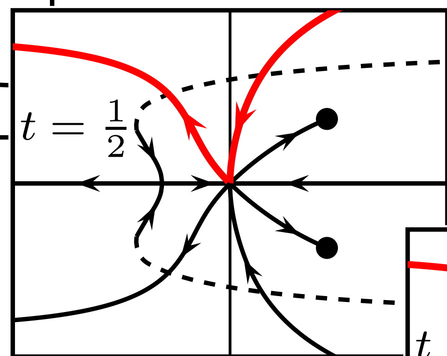
$$\Phi_{[2,1,-1]}(w,z) = w^2\left(1 - \tfrac{2}{3}z + \tfrac{1}{6}z^2\right) - w\left(1 + \tfrac{1}{3}z\right)$$

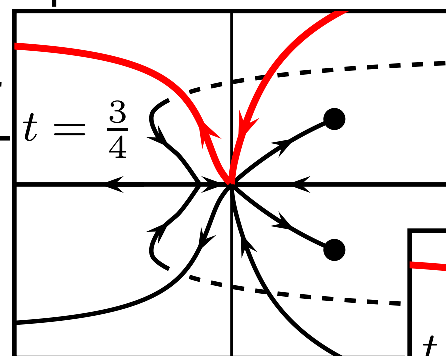
$$\Phi_{[2,1,0]}(w,z) = w^2\left(1 - \tfrac{10}{17}z + \tfrac{2}{17}z^2\right) - w\left(\tfrac{16}{17} + \tfrac{8}{17}z\right) - \tfrac{1}{17}$$

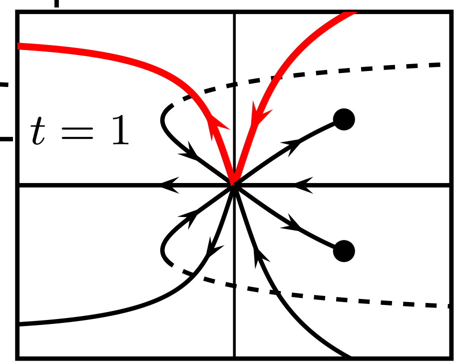


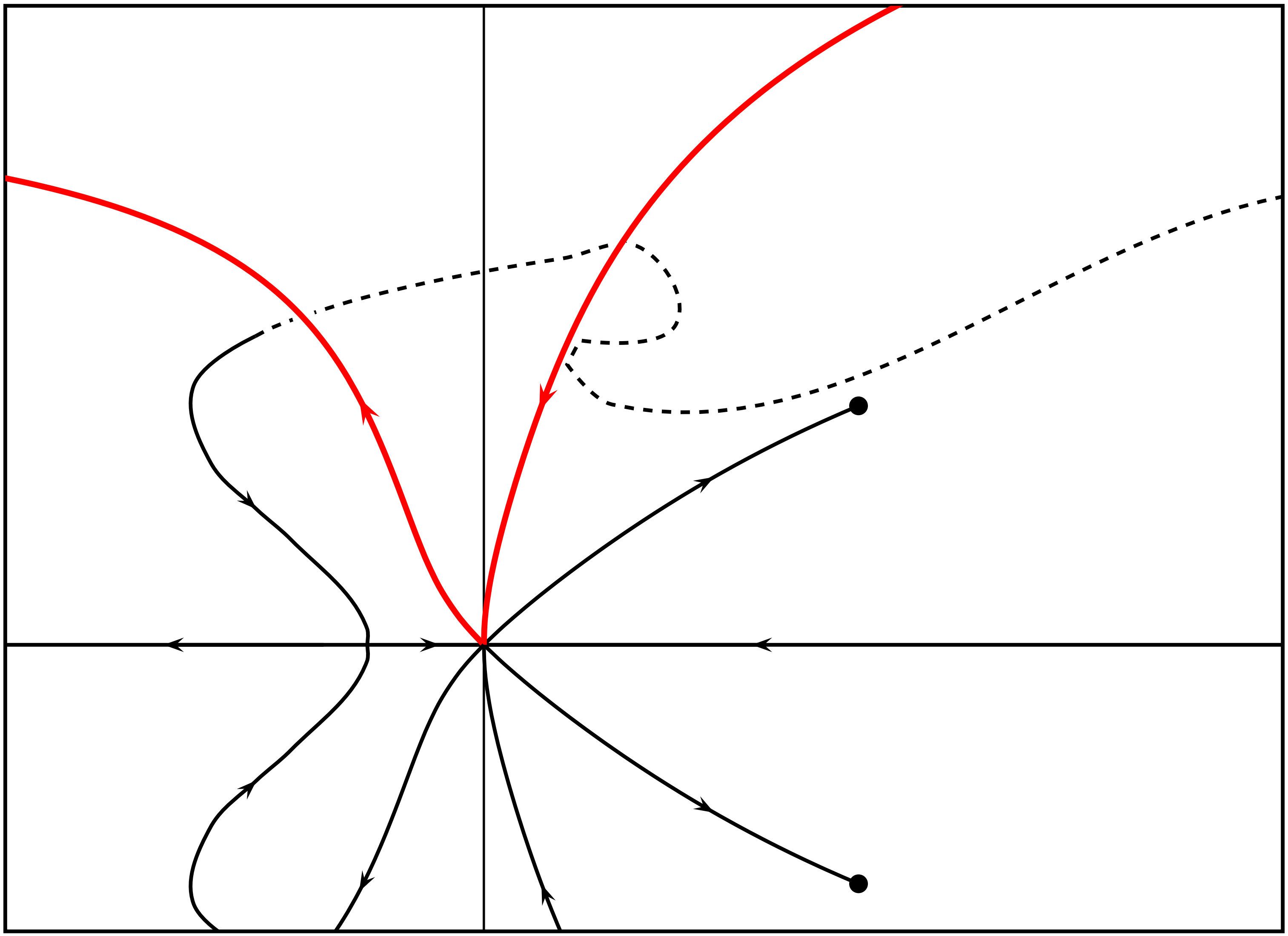


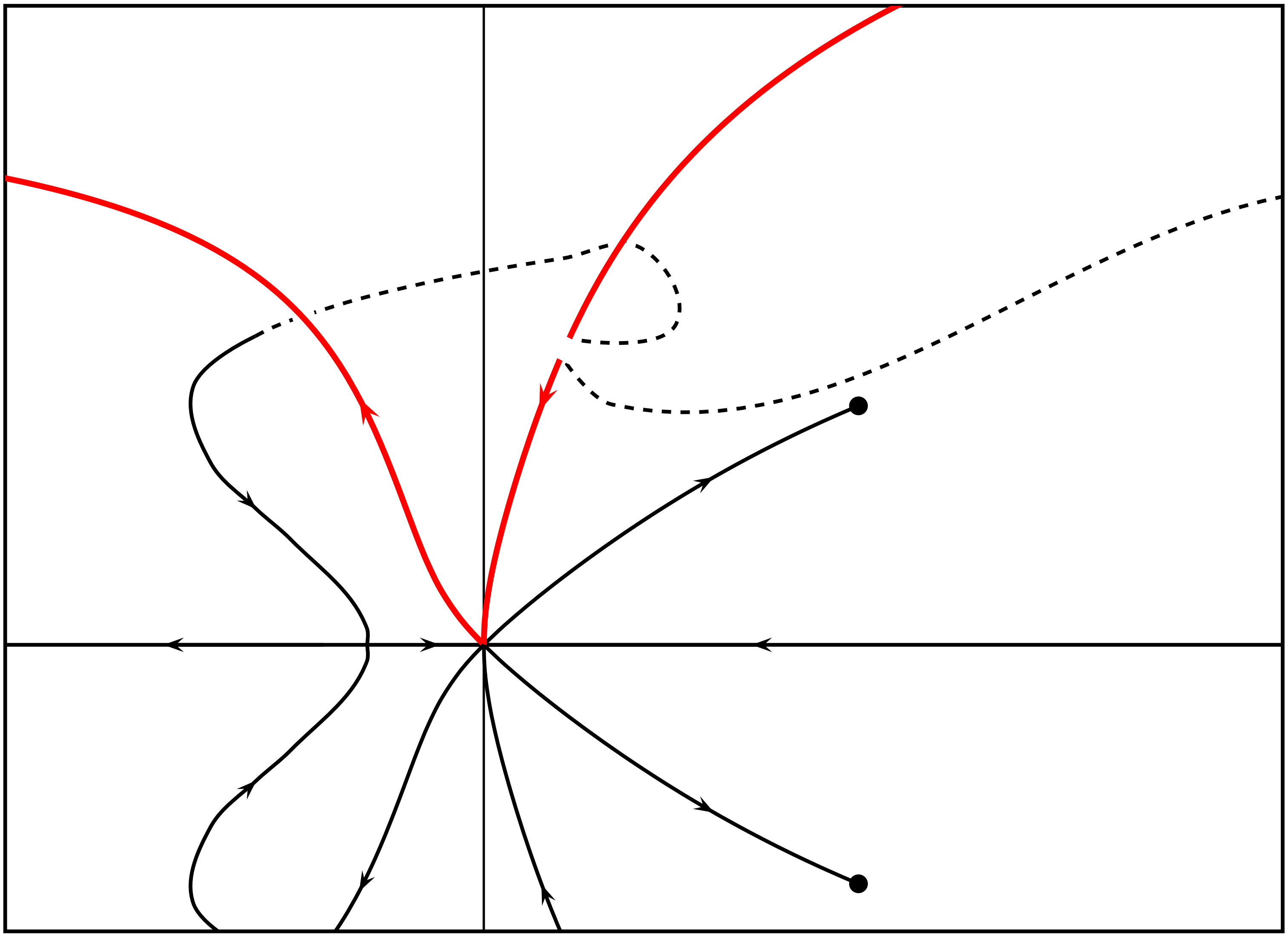
$$\Phi_t = (1-t)\Phi_{[2,1,-1]} + t\Phi_{[2,1,0]}$$

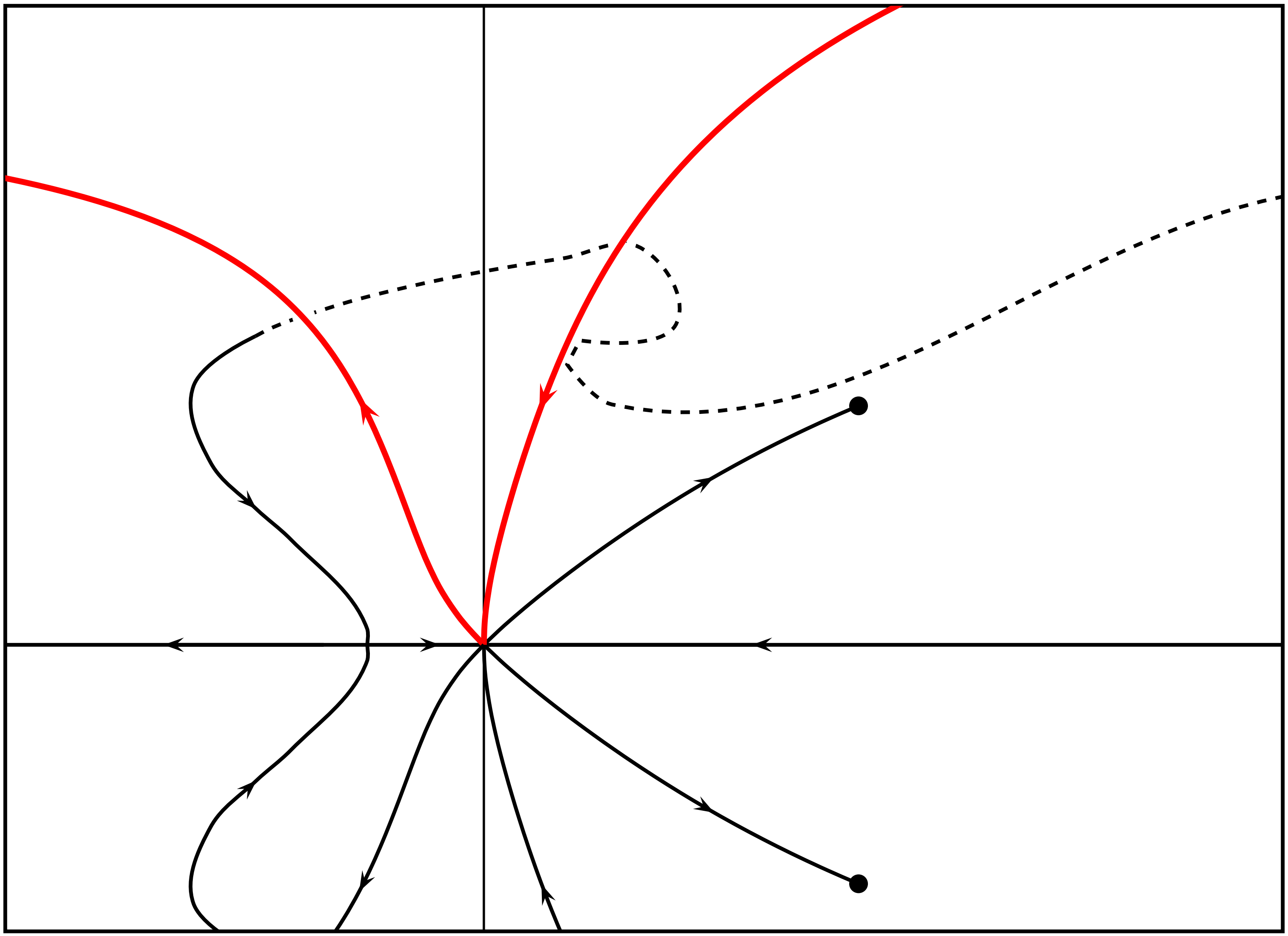












Many thanks