Solving Inverse Problems in Partial Differential Equations using higher degree Fréchet derivatives

William Rundell
Department of Mathematics
Texas A&M University

Joint work with Frank Hettlich (Karlsruhe)

What were some of the hot mathematical topics at the end of the seventeenth century?

What were some of the hot mathematical topics at the end of the seventeenth century?

Halley (1694) was impressed by a result published a few years earlier by Thomas de Lagney.

$$a + \frac{\mathrm{ab}}{3\mathrm{a}^3 + \mathrm{b}} < \sqrt[3]{a^3 + b} < \sqrt{\frac{\mathrm{a}^2}{4} + \frac{\mathrm{b}}{3\mathrm{a}}} \tag{1}$$

for $a^3 >> b > 0$, calling these the *rational formula* and the *irrational formula*. Out of these ideas evolved a numerical approximation scheme now known as *Halley's method*.

What were some of the hot mathematical topics at the end of the seventeenth century?

Halley (1694) was impressed by a result published a few years earlier by Thomas de Lagney.

$$a + \frac{\mathrm{ab}}{3\mathrm{a}^3 + \mathrm{b}} < \sqrt[3]{a^3 + b} < \sqrt{\frac{\mathrm{a}^2}{4} + \frac{\mathrm{b}}{3\mathrm{a}}} \tag{1}$$

for $a^3 >> b > 0$, calling these the *rational formula* and the *irrational formula*. Out of these ideas evolved a numerical approximation scheme now known as *Halley's method.*

For the standard example from school calculus, calculate $\sqrt[3]{28}$, a single step of Newton's method applied to $g(x) := x^3 - (a^3 + b)$ gives the approximation $\sqrt[3]{28} = 3.037037$.

What were some of the hot mathematical topics at the end of the seventeenth century?

Halley (1694) was impressed by a result published a few years earlier by Thomas de Lagney.

$$a + \frac{\mathrm{ab}}{3\mathrm{a}^3 + \mathrm{b}} < \sqrt[3]{a^3 + b} < \sqrt{\frac{\mathrm{a}^2}{4} + \frac{\mathrm{b}}{3\mathrm{a}}} \qquad (1)$$

for $a^3 >> b > 0$, calling these the *rational formula* and the *irrational formula*. Out of these ideas evolved a numerical approximation scheme now known as *Halley's method.*

For the standard example from school calculus, calculate $\sqrt[3]{28}$, a single step of Newton's method applied to $g(x) := x^3 - (a^3 + b)$ gives the approximation $\sqrt[3]{28} = 3.037037$.

In comparison, (1) gives the estimates

$$3.036585 < \sqrt[3]{28} < 3.036591$$

What were some of the hot mathematical topics at the end of the seventeenth century?

Halley (1694) was impressed by a result published a few years earlier by Thomas de Lagney.

$$a + \frac{\mathrm{ab}}{3\mathrm{a}^3 + \mathrm{b}} < \sqrt[3]{a^3 + b} < \sqrt{\frac{\mathrm{a}^2}{4} + \frac{\mathrm{b}}{3\mathrm{a}}} \qquad (1)$$

for $a^3 >> b > 0$, calling these the *rational formula* and the *irrational formula*. Out of these ideas evolved a numerical approximation scheme now known as *Halley's method.*

Brook Taylor recognised the derivatives implicit in Halley's method in 1714 (this should not be surprising since it was not until 1740 that Simpson showed the connection between derivatives and Newton's method) so, let's follow that lead . . . .

In the linear equation $f(x+h) = f(x) + f'(x)h$
set $x_n = x$, $h = x_{n+1} - x_n$, with target $f(x+h) = 0$, to give

$$x_{n+1} - x_n = h = -\frac{f(x_n)}{f'(x_n)} \qquad \text{(Newton's Method)}$$

In the linear equation $f(x+h) = f(x) + f'(x)h$
set $x_n = x$, $h = x_{n+1} - x_n$, with target $f(x+h) = 0$, to give

$$x_{n+1} - x_n = h = -\frac{f(x_n)}{f'(x_n)} \qquad \text{(Newton's Method)}$$

If we want greater accuracy the next step should be obvious

$$f(x+h) = f(x) + f'(x)h + \tfrac{1}{2} f''(x)h^2.$$

In the linear equation $f(x+h) = f(x) + f'(x)h$
set $x_n = x$, $h = x_{n+1} - x_n$, with target $f(x+h) = 0$, to give

$$x_{n+1} - x_n = h = -\frac{f(x_n)}{f'(x_n)} \qquad \text{(Newton's Method)}$$

If we want greater accuracy the next step should be obvious

$$f(x+h) = f(x) + f'(x)h + \tfrac{1}{2}f''(x)h^2.$$

This is a quadratic equation and the quadratic formula gives

$$h = \frac{-\mathrm{f}'(\mathrm{x}) - \sqrt{\mathrm{f}'(\mathrm{x})^2 - 2\mathrm{f}(\mathrm{x})\mathrm{f}''(\mathrm{x})}}{\mathrm{f}''(\mathrm{x})}$$

(evaluate $g(x)$ at $x = a$ to obtain the *irrational* formula).

An alternative method might be

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)\, h.\tilde{h}$$

where $\tilde{h}$ is predicted using Newton's method, $\tilde{h} = -\frac{f(x_n)}{f'(x_n)}$.

An alternative method might be

$$f(x+h) = f(x) + f'(x)h + \tfrac{1}{2} f''(x)\, h.\tilde{h}$$

where $\tilde{h}$ is predicted using Newton's method, $\tilde{h} = -\frac{f(x_n)}{f'(x_n)}$.

Inserting this prediction into the above gives the correction

$$f(x_n+h) = f(x_n) + f'(x_n)h - \tfrac{1}{2}\frac{\mathrm{f}''(\mathrm{x_n})\mathrm{f}(\mathrm{x_n})}{\mathrm{f}'(\mathrm{x_n})} h$$

$$= f(x_n) + \tfrac{1}{2}\frac{2\mathrm{f}'(\mathrm{x_n})^2 - \mathrm{f}''(\mathrm{x_n})\mathrm{f}(\mathrm{x_n})}{\mathrm{f}'(\mathrm{x_n})}\, h$$

An alternative method might be

$$f(x+h) = f(x) + f'(x)h + \tfrac{1}{2}f''(x)\,h.\tilde{h}$$

where $\tilde{h}$ is predicted using Newton's method, $\tilde{h} = -\frac{f(x_n)}{f'(x_n)}$.

Inserting this prediction into the above gives the correction

$$f(x_n+h) = f(x_n) + f'(x_n)h - \tfrac{1}{2}\frac{\mathrm{f}''(\mathrm{x_n})\mathrm{f}(\mathrm{x_n})}{\mathrm{f}'(\mathrm{x_n})}h$$

$$= f(x_n) + \tfrac{1}{2}\frac{2\mathrm{f}'(\mathrm{x_n})^2 - \mathrm{f}''(\mathrm{x_n})\mathrm{f}(\mathrm{x_n})}{\mathrm{f}'(\mathrm{x_n})}\,h$$

This is again a linear equation in h and can be easily solved

$$x_{n+1} - x_n = h = \frac{-2f''(x_n)f(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)} \qquad \text{(Halley's Method)}$$

(Apply this to $g(x)$, evaluate at $x = a$ to obtain the *rational* formula)

An alternative method might be

$$f(x+h) = f(x) + f'(x)h + \tfrac{1}{2}f''(x)\,h.\tilde{h}$$

where $\tilde{h}$ is predicted using Newton's method, $\tilde{h} = -\frac{f(x_n)}{f'(x_n)}$.

Inserting this prediction into the above gives the correction

$$\begin{aligned} f(x_n+h) &= f(x_n) + f'(x_n)h - \tfrac{1}{2}\frac{\mathrm{f''(x_n)f(x_n)}}{\mathrm{f'(x_n)}}h \\ &= f(x_n) + \tfrac{1}{2}\frac{\mathrm{2f'(x_n)^2 - f''(x_n)f(x_n)}}{\mathrm{f'(x_n)}}\,h \end{aligned}$$

This is again a linear equation in h and can be easily solved

$$x_{n+1} - x_n = h = \frac{-2f''(x_n)f(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)} \qquad \text{(Halley's Method)}$$

(Apply this to $g(x)$, evaluate at $x = a$ to obtain the *rational* formula)

Halley's Method can be viewed as a predictor-corrector scheme that uses Newton's method as the predictor. Both predictor and corrector steps require solving only *linear* equations.

What about solving a nonlinear equation in function space?
The paradigm is: let X, Y be Banach spaces and $x \in U \subseteq X$, then for $g \in Y$ we wish to solve

$$F(x) = g$$

What about solving a nonlinear equation in function space?
The paradigm is: let X, Y be Banach spaces and $x \in U \subseteq X$, then for $g \in Y$ we wish to solve

$$F(x) = g$$

There are many possible approaches; we can attempt to solve this "point-wise" in X, or as a minimisation problem

$$\min_{x \in U} r(x) \qquad r(x) := \|F(x) - g\|_X .$$

What about solving a nonlinear equation in function space?
The paradigm is: let X, Y be Banach spaces and $x \in U \subseteq X$, then for $g \in Y$ we wish to solve

$$F(x) = g$$

There are many possible approaches; we can attempt to solve this "point-wise" in X, or as a minimisation problem

$$\min_{x \in U} r(x) \qquad\qquad r(x) := \|F(x) - g\|_X .$$

We are interested in iterative solution methods

$$x_{n+1} = x_n + \mathcal{A}_n(F(x_n) - g)$$

$$\mathcal{A}_n = \begin{cases} -(F')^{-1}(x_n) & \text{(Newton)} \\ -\mu(F')^{\star}(x_n) & \text{(Landweber)} \end{cases}$$

What about solving a nonlinear equation in function space?
The paradigm is: let X, Y be Banach spaces and $x \in U \subseteq X$, then for $g \in Y$ we wish to solve

$$F(x) = g$$

There are many possible approaches; we can attempt to solve this "point-wise" in X, or as a minimisation problem

$$\min_{x \in U} r(x) \qquad\qquad r(x) := \|F(x) - g\|_X.$$

We are interested in iterative solution methods

$$x_{n+1} = x_n + \mathcal{A}_n(F(x_n) - g)$$

$$\mathcal{A}_n = \begin{cases} -(F')^{-1}(x_n) & \text{(Newton)} \\ -\mu(F')^\star(x_n) & \text{(Landweber)} \end{cases}$$

What about using higher order derivatives in $\mathcal{A}$?

$$\mathcal{A} = \mathcal{A}(F', F'', \ldots)$$

$$F(x) = g \qquad x_{n+1} = x_n + \mathcal{A}_n(F(x_n) - g)$$

Traditional wisdom would say the following:

Computationally, it is not worthwhile to invoke methods that utilise second derivatives of F since . . . .

$$F(x) = g \qquad x_{n+1} = x_n + \mathcal{A}_n(F(x_n) - g)$$

Traditional wisdom would say the following:

Computationally, it is not worthwhile to invoke methods that utilise second derivatives of F since . . . .

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

$$F(x) = g \qquad x_{n+1} = x_n + \mathcal{A}_n(F(x_n) - g)$$

Traditional wisdom would say the following:

Computationally, it is not worthwhile to invoke methods that utilise second derivatives of F since . . . .

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

2. Calculation of the Hessians of F is very computationally intensive. In fact, even brute force computation of a full first derivative might be infeasible.

$$F(x) = g \qquad x_{n+1} = x_n + \mathcal{A}_n(F(x_n) - g)$$

Traditional wisdom would say the following:

Computationally, it is not worthwhile to invoke methods that utilise second derivatives of F since . . . .

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

2. Calculation of the Hessians of F is very computationally intensive. In fact, even brute force computation of a full first derivative might be infeasible.

3. Have to deal with quadratic equations in function space.

$$F(x) = g \qquad x_{n+1} = x_n + \mathcal{A}_n(F(x_n) - g)$$

Traditional wisdom would say the following:

Computationally, it is not worthwhile to invoke methods that utilise second derivatives of F since . . . .

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

2. Calculation of the Hessians of F is very computationally intensive. In fact, even brute force computation of a full first derivative might be infeasible.

3. Have to deal with quadratic equations in function space.

Here is the response:

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

- For ill-posed problems $r(x)$ is never going to approach zero.

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

- For ill-posed problems $r(x)$ is never going to approach zero.

2. Calculation of the Hessians of F is very computationally intensive. In fact, even brute force computation of a full first derivative might be infeasible.

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

- For ill-posed problems $r(x)$ is never going to approach zero.

2. Calculation of the Hessians of F is very computationally intensive. In fact, even brute force computation of a full first derivative might be infeasible.

- For certain pde's the second is not a valid criticism. Once F has been obtained, derivatives can often be computed with a relatively small additional cost.

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

- For ill-posed problems $r(x)$ is never going to approach zero.

2. Calculation of the Hessians of F is very computationally intensive. In fact, even brute force computation of a full first derivative might be infeasible.

- For certain pde's the second is not a valid criticism. Once F has been obtained, derivatives can often be computed with a relatively small additional cost.

3. Have to deal with quadratic equations in function space.

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

- For ill-posed problems $r(x)$ is never going to approach zero.

2. Calculation of the Hessians of F is very computationally intensive. In fact, even brute force computation of a full first derivative might be infeasible.

- For certain pde's the second is not a valid criticism. Once F has been obtained, derivatives can often be computed with a relatively small additional cost.

3. Have to deal with quadratic equations in function space.

- Our derivation of Halley's method gave a hint of how to deal with this item.

1. Near convergence (residual $r(x)$ near zero) the quadratic term is exceedingly small and no benefit is obtained from including higher derivatives.

- For ill-posed problems $r(x)$ is never going to approach zero.

2. Calculation of the Hessians of F is very computationally intensive. In fact, even brute force computation of a full first derivative might be infeasible.

- For certain pde's the second is not a valid criticism. Once F has been obtained, derivatives can often be computed with a relatively small additional cost.

3. Have to deal with quadratic equations in function space.

- Our derivation of Halley's method gave a hint of how to deal with this item.

Let us look at each of these responses in turn . . . .

We wish to solve the (first kind) integral equation

$$\mathcal{K}f = g \qquad\qquad \int_0^1 K(x,t)f(t)\,dt = g(x)$$

ruept for the function f given the data g.

We wish to solve the (first kind) integral equation

$$\mathcal{K}f = g \qquad\qquad \int_0^1 K(x,t)f(t)\,dt = g(x)$$

ruept for the function f given the data g.

Two common approaches use the Nyström or the Galerkin method:

$$g_i = \sum_1^N A_{ij} f_j \qquad A_{ij} = \begin{cases} K(x_i, t_j) w_j & f_j = f(x_j) \\ \int_0^1 K(x_i, t)\phi_j(t)\,dt & f_j = \langle f, \phi_j \rangle \end{cases}$$

We wish to solve the (first kind) integral equation

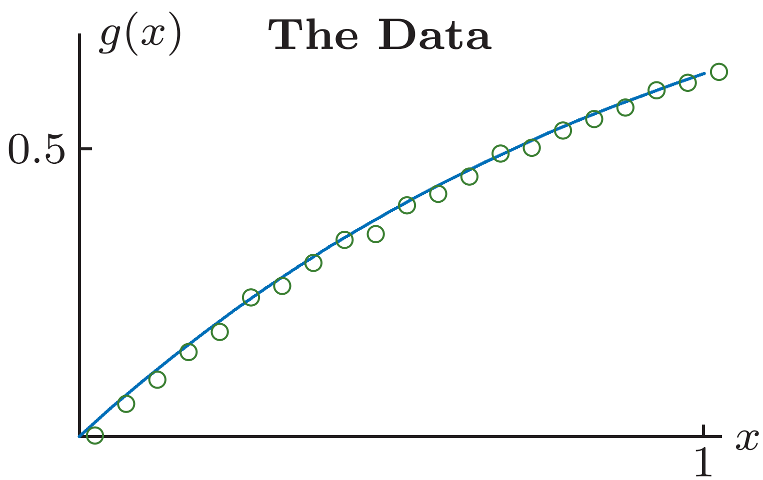
$$\mathcal{K}f = g \qquad\qquad \int_0^1 K(x,t)f(t)\,dt = g(x)$$

ruept for the function f given the data g.

Two common approaches use the Nyström or the Galerkin method:

$$g_i = \sum_1^N A_{ij} f_j \qquad A_{ij} = \begin{cases} K(x_i,t_j)w_j & f_j = f(x_j) \\ \int_0^1 K(x_i,t)\phi_j(t)\,dt & f_j = \langle f, \phi_j\rangle \end{cases}$$

Example: $$Kf = \int_0^1 xe^{-xt} f(t)\,dt = g(x) \qquad x \in [0,1]$$



Here $g(x)$ is an approximation to $1 - e^{-x}$.

Solution should be $f = 1$.

We wish to solve the (first kind) integral equation

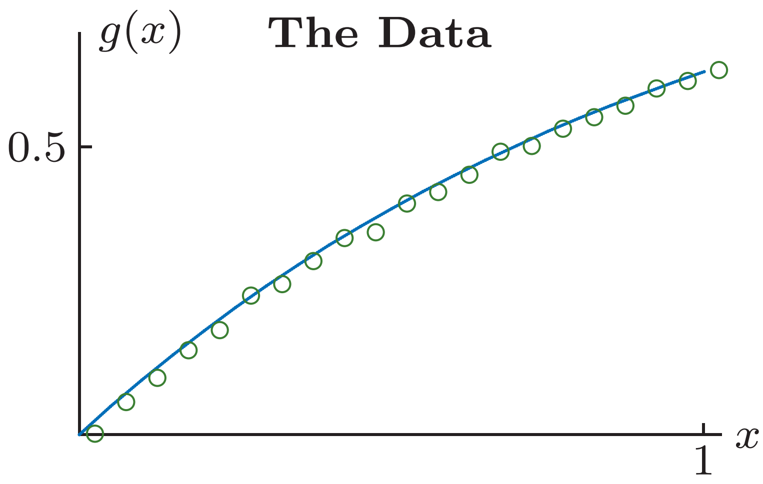
$$\mathcal{K}f = g \qquad\qquad \int_0^1 K(x,t)f(t)\,dt = g(x)$$

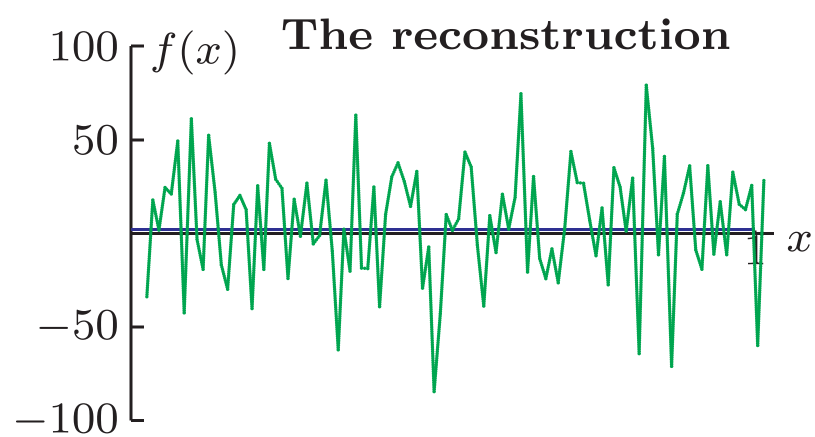
ruept for the function f given the data g.

Two common approaches use the Nyström or the Galerkin method:

$$g_i = \sum_1^N A_{ij} f_j \qquad A_{ij} = \begin{cases} K(x_i,t_j)w_j & f_j = f(x_j) \\ \int_0^1 K(x_i,t)\phi_j(t)\,dt & f_j = \langle f, \phi_j \rangle \end{cases}$$

Example: $$Kf = \int_0^1 xe^{-xt} f(t)\,dt = g(x) \qquad x \in [0,1]$$





We wish to solve the (first kind) integral equation

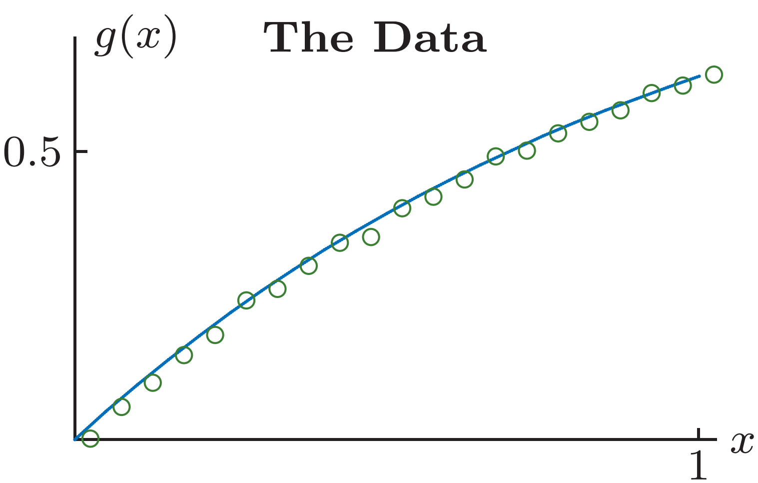
$$\mathcal{K}f = g \qquad\qquad \int_0^1 K(x,t)f(t)\,dt = g(x)$$

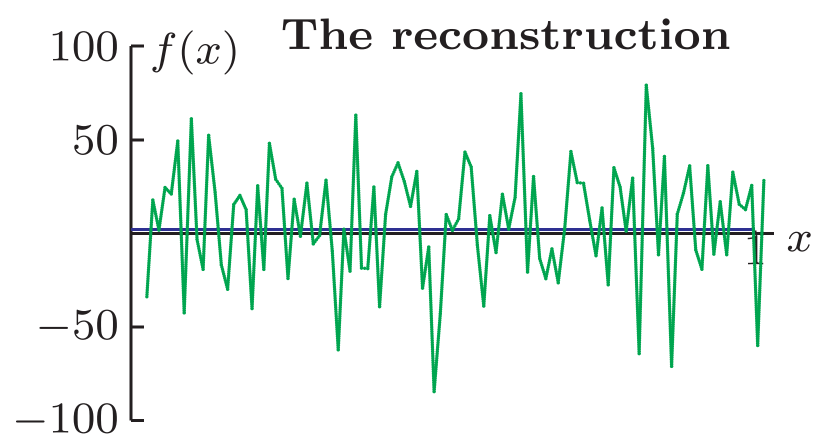
ruept for the function f given the data g.

Two common approaches use the Nyström or the Galerkin method:

$$g_i = \sum_1^N A_{ij} f_j \qquad A_{ij} = \begin{cases} K(x_i, t_j) w_j & f_j = f(x_j) \\ \int_0^1 K(x_i, t)\phi_j(t)\,dt & f_j = \langle f, \phi_j \rangle \end{cases}$$

Example: $$Kf = \int_0^1 x e^{-xt} f(t)\,dt = g(x) \qquad x \in [0,1]$$





Clearly, something has gone horribly wrong

Let's take a more "pde-example",

$$K(x,t) = \begin{cases} t(1-x) & \text{if } t < x \\ x(1-t) & \text{if } t > x \end{cases}$$

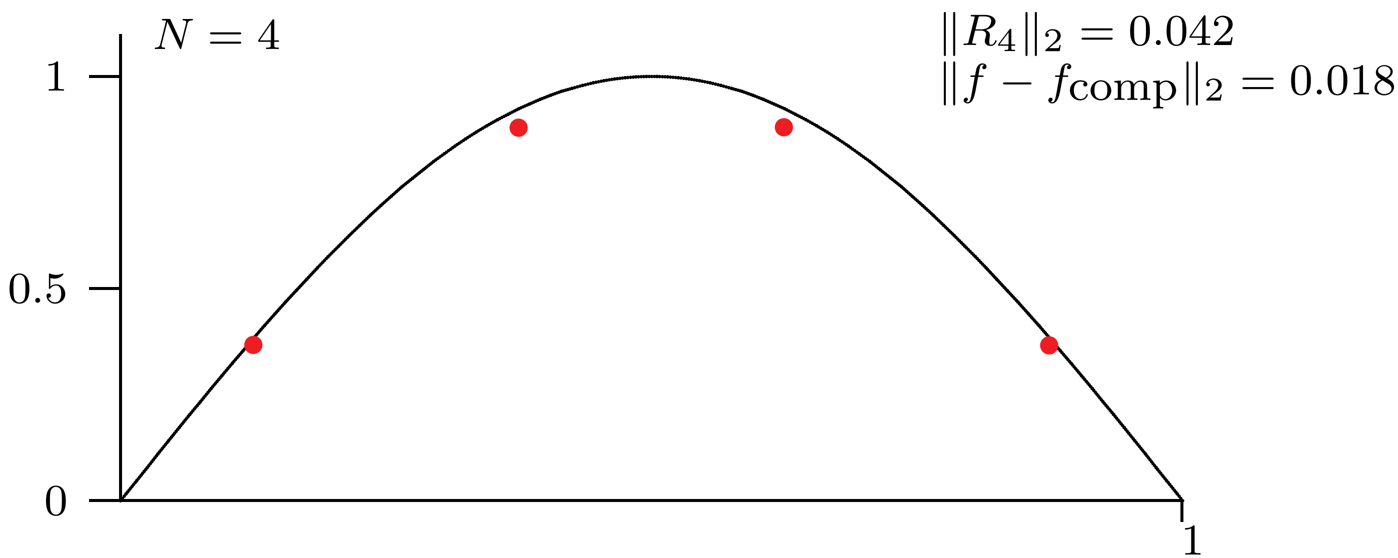
Green's function for $\mathcal{L}y = -y''$

$$g(x) = \frac{\sin(\pi x)}{\pi^2}, \ (f = \sin \pi x)$$

Let's take a more "pde-example", Green's function for $\mathcal{L}y = -y''$

$$K(x,t) = \begin{cases} t(1-x) & \text{if } t < x \\ x(1-t) & \text{if } t > x \end{cases} \qquad g(x) = \frac{\sin(\pi x)}{\pi^2}, \ (f = \sin \pi x)$$

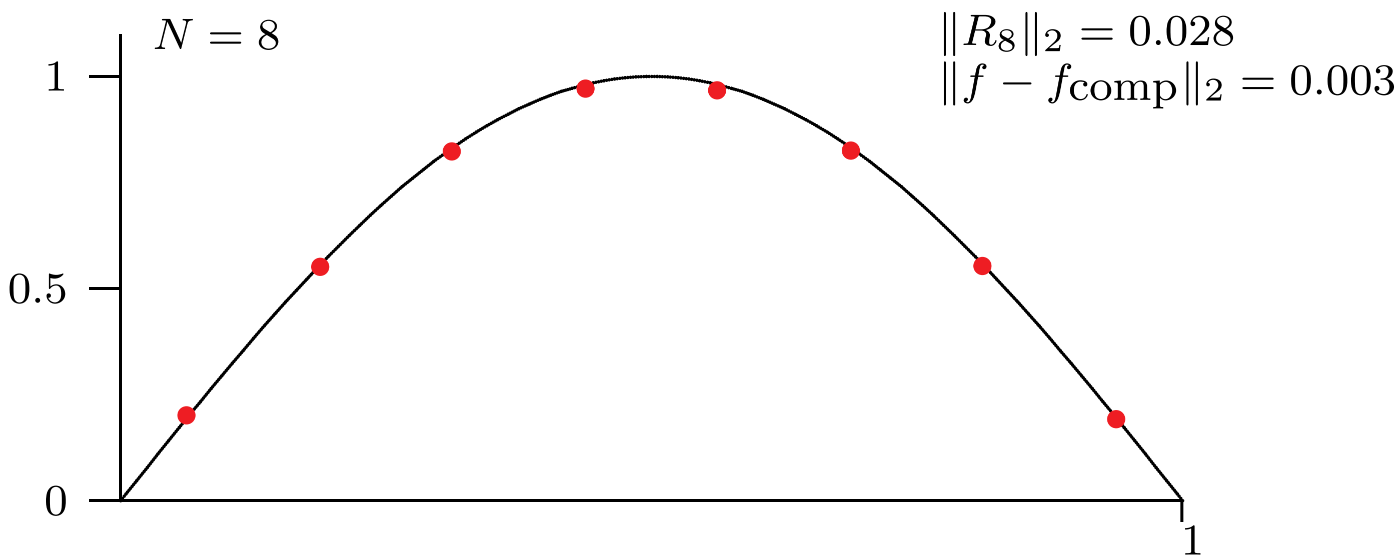
Using the Nyström method, trapezoid rule:



Let's take a more "pde-example", Green's function for $\mathcal{L}y = -y''$

$$K(x,t) = \begin{cases} t(1-x) & \text{if } t < x \\ x(1-t) & \text{if } t > x \end{cases} \qquad g(x) = \frac{\sin(\pi x)}{\pi^2},\ (f = \sin \pi x)$$

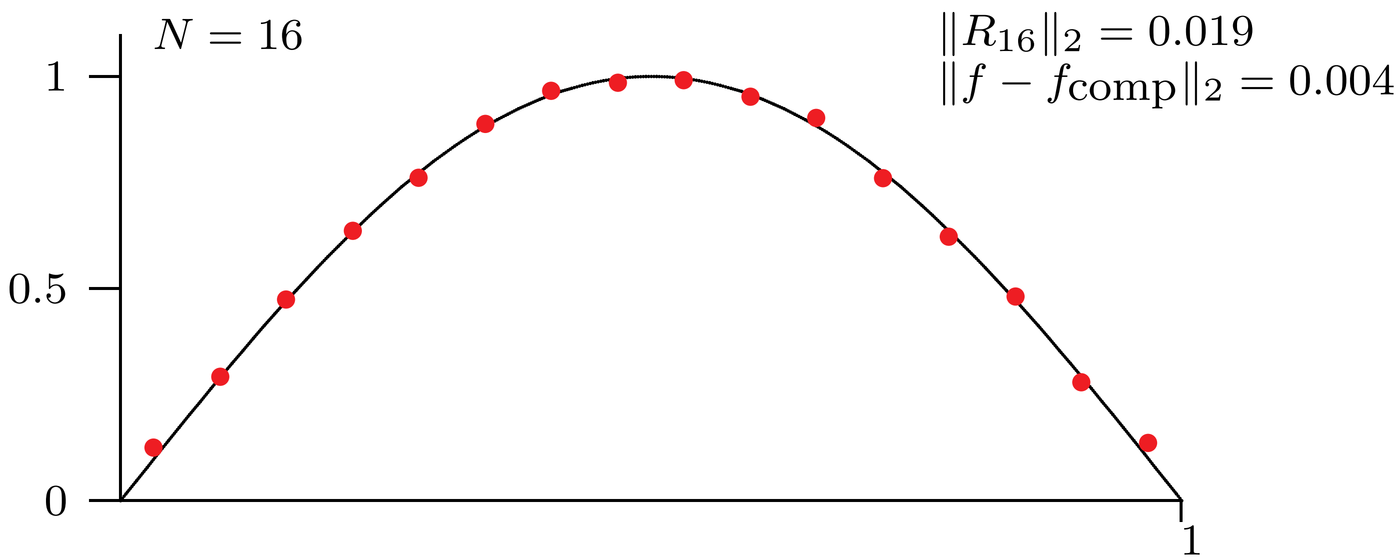
Using the Nyström method, trapezoid rule:



Let's take a more "pde-example", Green's function for $\mathcal{L}y = -y''$

$$K(x,t) = \begin{cases} t(1-x) & \text{if } t < x \\ x(1-t) & \text{if } t > x \end{cases} \qquad g(x) = \frac{\sin(\pi x)}{\pi^2}, \ (f = \sin \pi x)$$

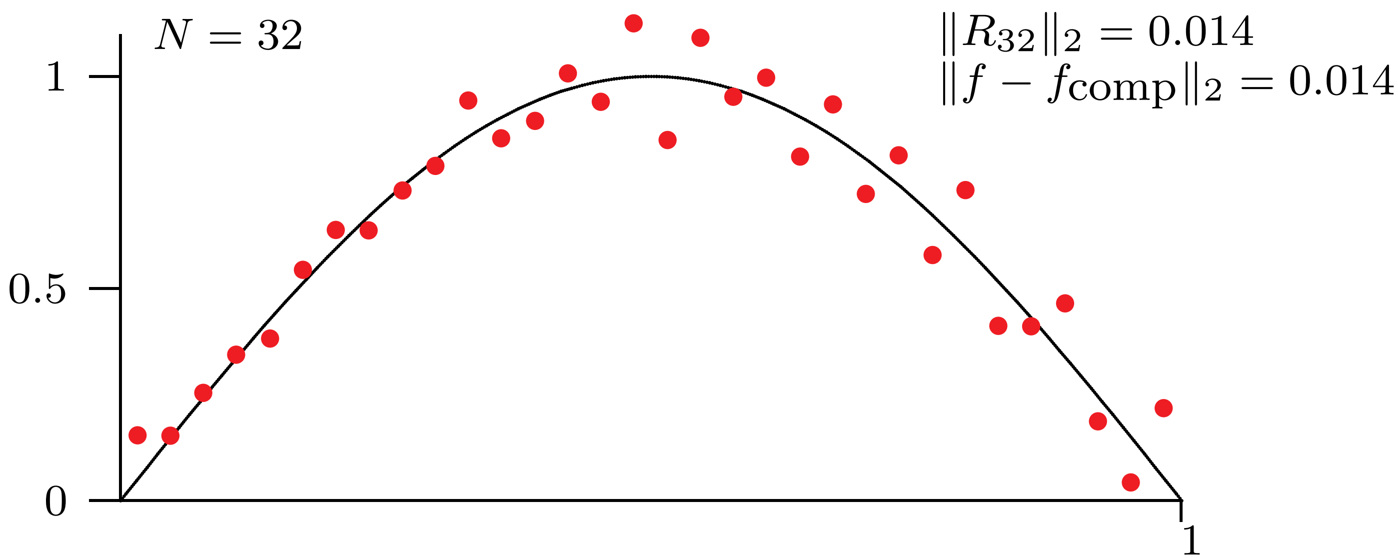
Using the Nyström method, trapezoid rule:



Let's take a more "pde-example", Green's function for $\mathcal{L}y = -y''$

$$K(x,t) = \begin{cases} t(1-x) & \text{if } t < x \\ x(1-t) & \text{if } t > x \end{cases} \qquad g(x) = \frac{\sin(\pi x)}{\pi^2},\ (f = \sin \pi x)$$

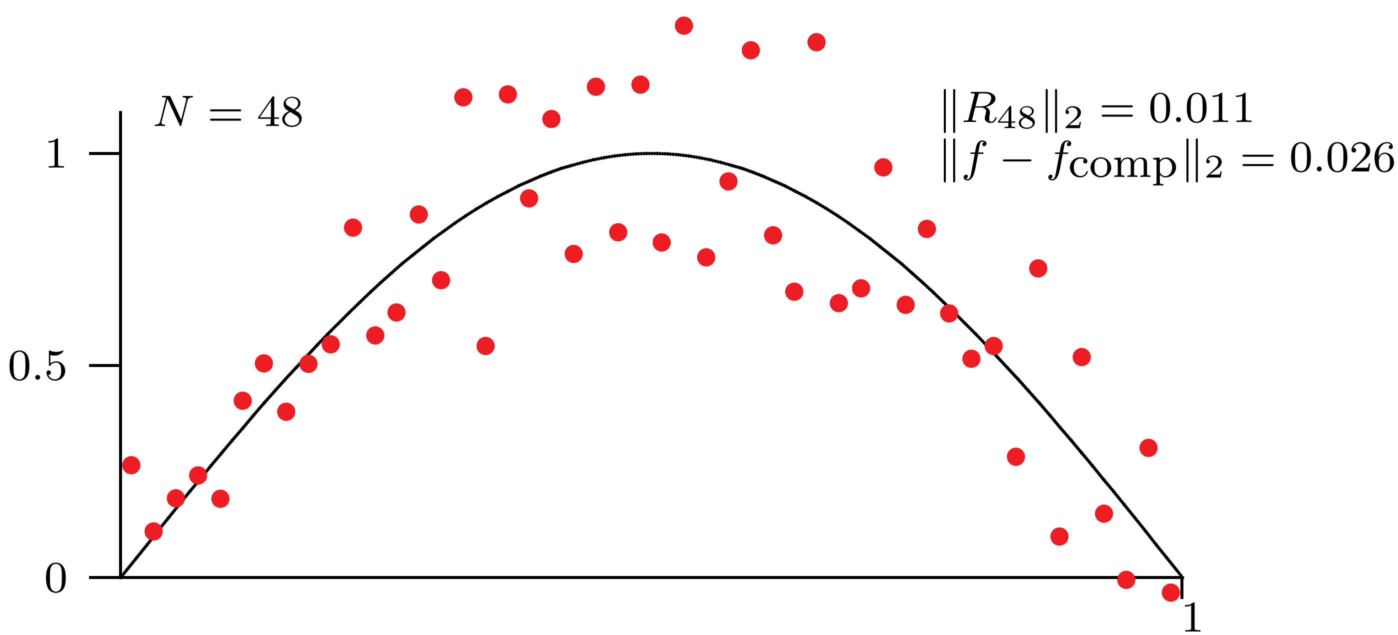
Using the Nyström method, trapezoid rule:



Let's take a more "pde-example", Green's function for $\mathcal{L}y = -y''$

$$K(x,t) = \begin{cases} t(1-x) & \text{if } t < x \\ x(1-t) & \text{if } t > x \end{cases} \qquad g(x) = \frac{\sin(\pi x)}{\pi^2}, \ (f = \sin \pi x)$$

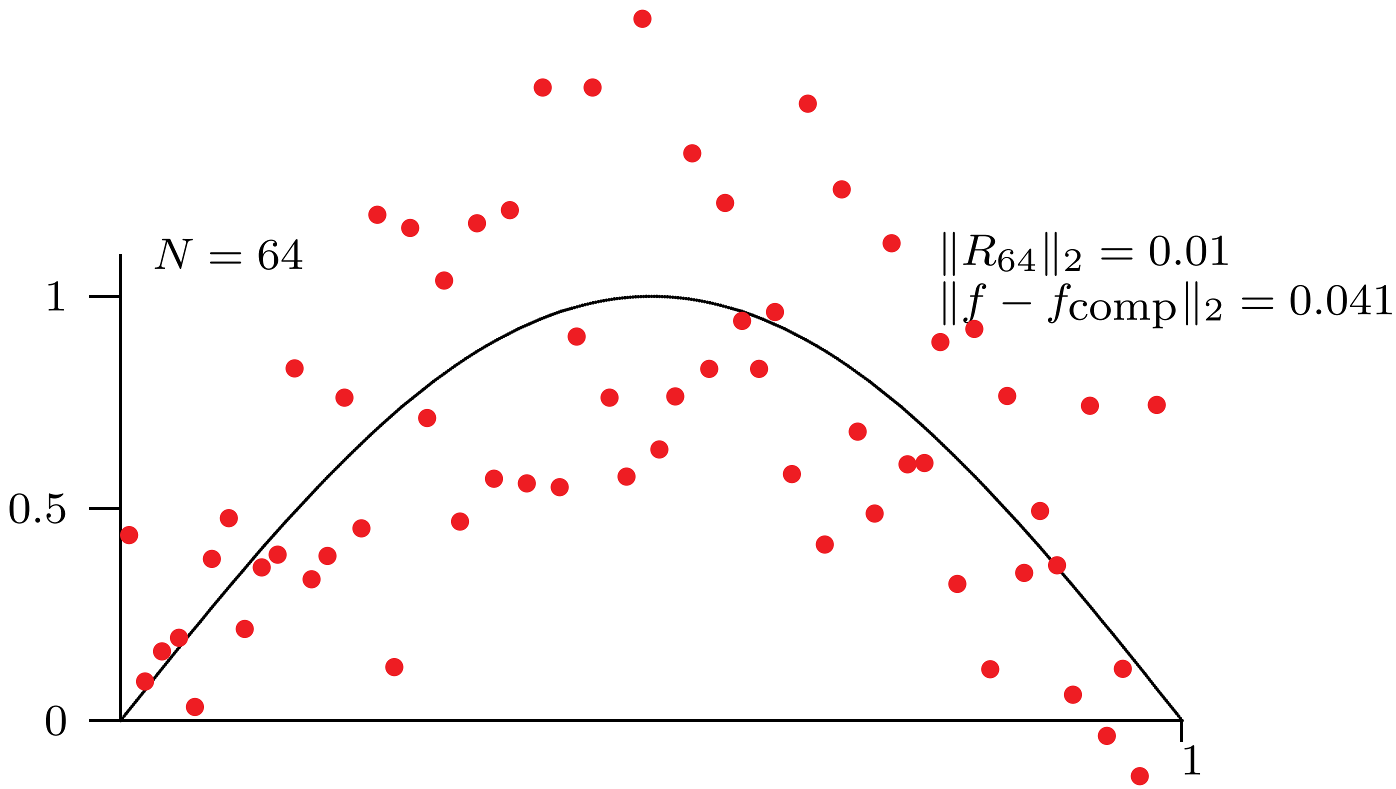
Using the Nyström method, trapezoid rule:



Let's take a more "pde-example", Green's function for $\mathcal{L}y = -y''$

$$K(x,t) = \begin{cases} t(1-x) & \text{if } t < x \\ x(1-t) & \text{if } t > x \end{cases} \qquad g(x) = \frac{\sin(\pi x)}{\pi^2},\ (f = \sin \pi x)$$

Using the Nyström method, trapezoid rule:



The problem is the operator $\mathcal{K}$ is *compact* and even if one-to-one then its inverse is not bounded.

The problem is the operator $\mathcal{K}$ is *compact* and even if one-to-one then its inverse is not bounded.

Here are the first 10 singular values of $\mathcal{K}$ when $K(x,t) = xe^{-xt}$:

0.188, 0.090, 1.62×10^{-3}, 1.45×10^{-6}, 7.82×10^{-9}, 2.77×10^{-11}, 9.41×10^{-13}, 2.63×10^{-13}, 1.67×10^{-13}, 1.32×10^{-13}, 8.34×10^{-14}

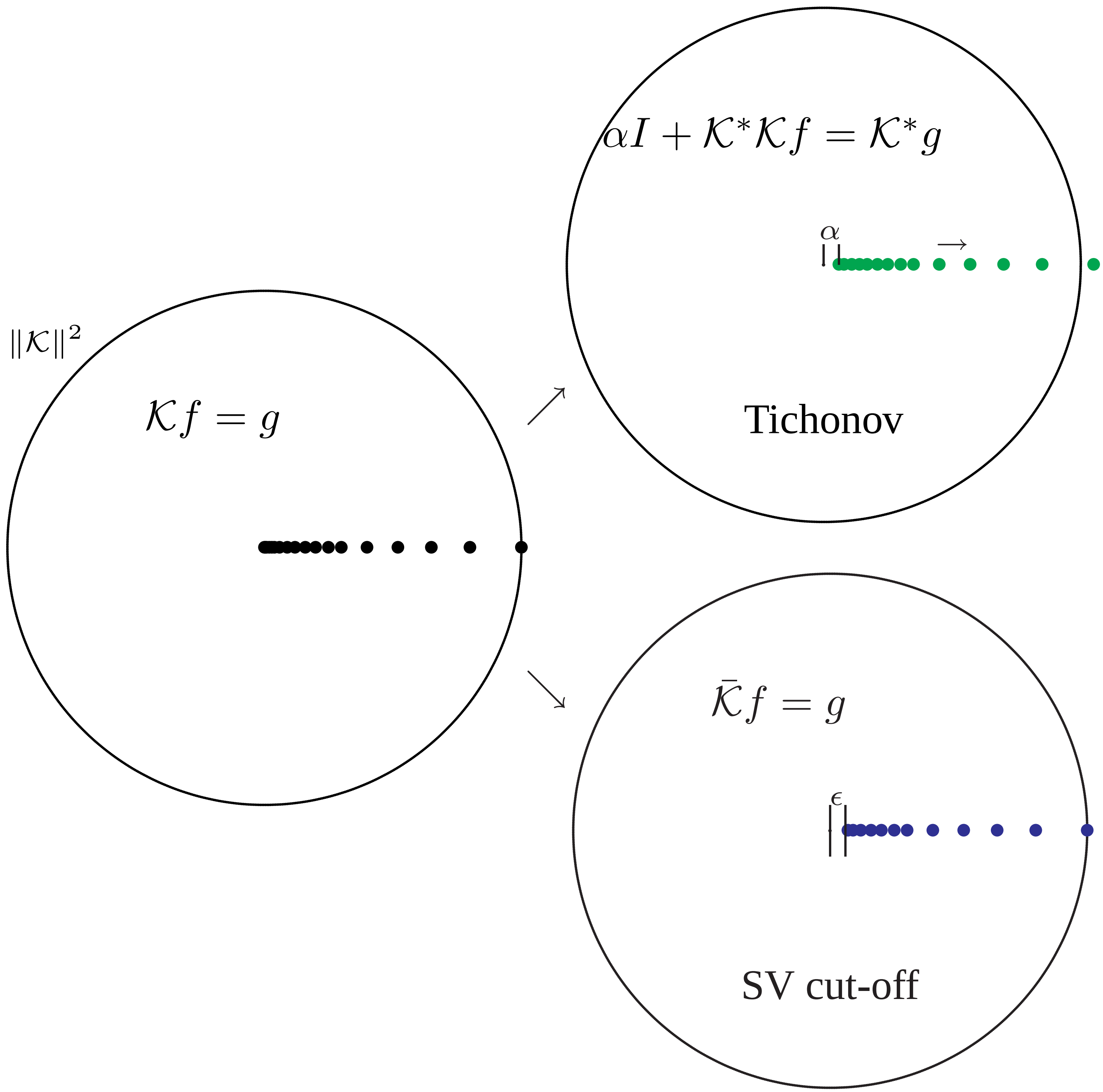
The problem is the operator $\mathcal{K}$ is *compact* and even if one-to-one then its inverse is not bounded.

Here are the first 10 singular values of $\mathcal{K}$ when $K(x,t) = xe^{-xt}$:

0.188, 0.090, 1.62×10^{-3}, 1.45×10^{-6}, 7.82×10^{-9}, 2.77×10^{-11}, 9.41×10^{-13}, 2.63×10^{-13}, 1.67×10^{-13}, 1.32×10^{-13}, 8.34×10^{-14}

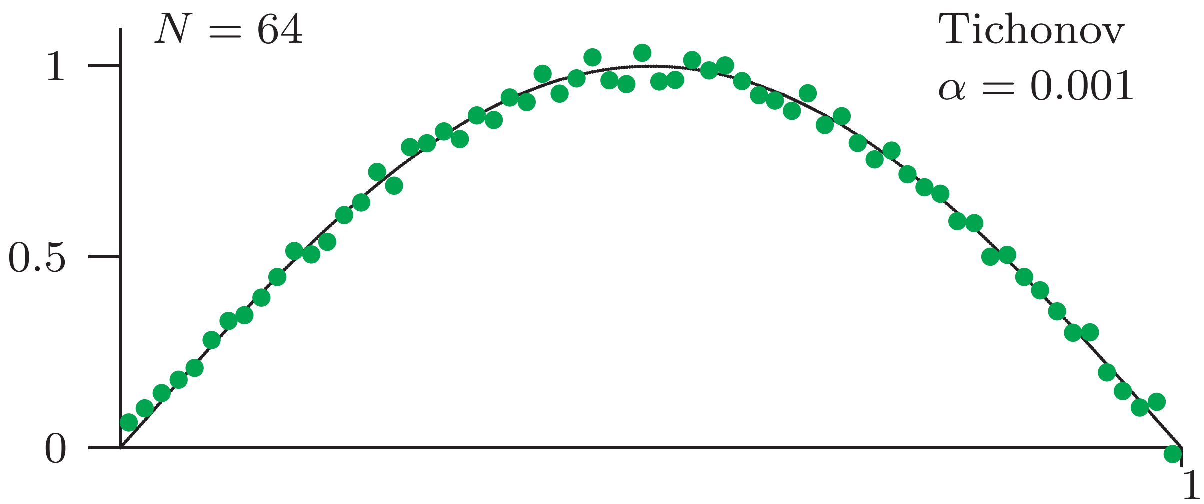
There are some common resolutions:

- Truncate the Fourier series for f $\quad f_N(x) = \sum_0^N f_k \, cos \, k\pi x$.
- Replace $Kf = g$ by $\tilde{K}f = g$, $\tilde{K} = \epsilon I + KK^{\star}$ (Tichonov)
- Stop iterating when the residual increases (stopping condition).

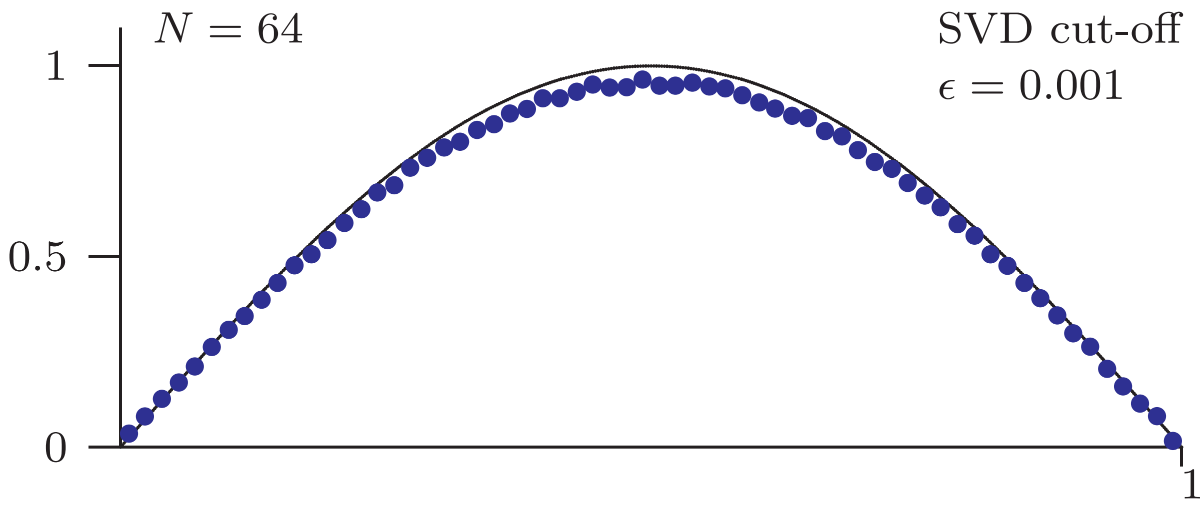

$\|\mathcal{K}\|^2$
$\mathcal{K}f = g$
$\alpha I + \mathcal{K}^*\mathcal{K}f = \mathcal{K}^*g$
α
Tichonov
$\bar{\mathcal{K}}f = g$
ϵ
SV cut-off

Need for **regularisation**.

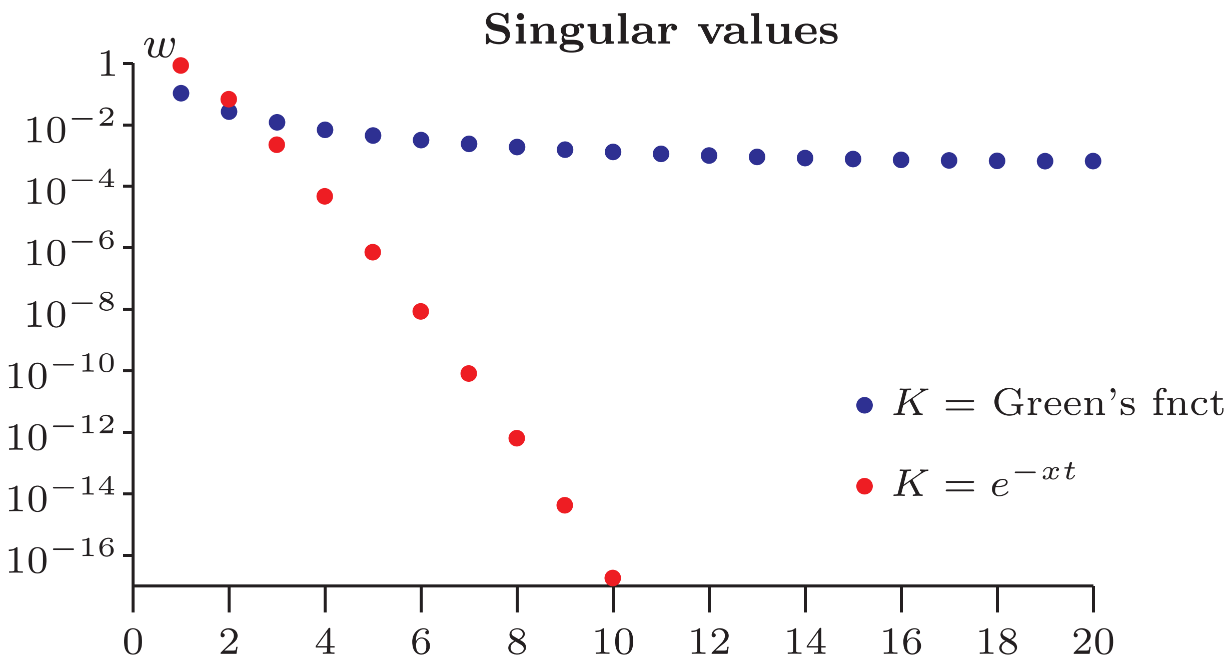
Need for **regularisation**.



Need for **regularisation**.



Need for **regularisation**.



What have we learned?

- It is essential to regularise an ill-posed problem
  - The type and degree of regularisation depends on the error and the problem.
  - A common strategy combines two regularisation methods; one such as Tichonov together with a “stopping condition”.
  - Experience shows it is better to terminate the iteration process *before* the residual $\|r\|$ reaches its minimum value.
  - Despite an extensive theory on the topic, choosing the correct regularisation parameters can be an art in itself.

What have we learned?

- It is essential to regularise an ill-posed problem
  - The type and degree of regularisation depends on the error and the problem.
  - A common strategy combines two regularisation methods; one such as Tichonov together with a “stopping condition”.
  - Experience shows it is better to terminate the iteration process *before* the residual $\|r\|$ reaches its minimum value.
  - Despite an extensive theory on the topic, choosing the correct regularisation parameters can be an art in itself.
- The residual will never go to zero.

Suppose we are interested in recovering the coefficient q in

$$-\triangle u + qu = f(x) \qquad \text{in } \Omega$$

$$u = 0 \qquad \text{on } \partial\Omega$$

where we have the additional measured data $\dfrac{\partial \mathrm{u}}{\partial \nu} = g$ on $\partial\Omega$ so that F is the map $F : q \to \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}$.

Suppose we are interested in recovering the coefficient q in

$$-\triangle u + qu = f(x) \qquad \text{in } \Omega$$

$$u = 0 \qquad \text{on } \partial\Omega$$

where we have the additional measured data $\dfrac{\partial \mathrm{u}}{\partial \nu} = g$ on $\partial\Omega$ so that F is the map $F : q \longrightarrow \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}$.

Any numerical scheme designed to solve this equation will come down to solving a matrix equation of the form

$$Ax = b$$

where the matrix A will depend on the domain Ω and the coefficient q; the vector b will depend on the right hand side.

Suppose we are interested in recovering the coefficient q in

$$-\triangle u + qu = f(x) \qquad \text{in } \Omega$$

$$u = 0 \qquad \text{on } \partial\Omega$$

where we have the additional measured data $\dfrac{\partial \mathrm{u}}{\partial \nu} = g$ on $\partial\Omega$ so that F is the map $F : q \to \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}$.

Any numerical scheme designed to solve this equation will come down to solving a matrix equation of the form

$$Ax = b$$

where the matrix A will depend on the domain Ω and the coefficient q; the vector b will depend on the right hand side.

- The majority of the effort is in setting up and “inverting” A

Suppose we are interested in recovering the coefficient q in

$$-\triangle u + qu = f(x) \qquad \text{in } \Omega$$

$$u = 0 \qquad \text{on } \partial\Omega$$

where we have the additional measured data $\dfrac{\partial \mathrm{u}}{\partial \nu} = g$ on $\partial\Omega$ so that F is the map $F : q \to \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}$.

The derivative of the map F in the direction h_1, $\frac{\partial F}{\partial q}.h_1$, is given by the normal derivative on $\partial\Omega$ of u', where

$$-\triangle u' + qu' = -u\, h_1 \quad \text{in } \Omega \qquad u' = 0 \quad \text{on } \partial\Omega.$$

To solve this we have the equation $Ax = c$ where c depends on h_1 and the previously computed u.

Suppose we are interested in recovering the coefficient q in

$$-\triangle u + qu = f(x) \qquad \text{in } \Omega$$

$$u = 0 \qquad \text{on } \partial\Omega$$

where we have the additional measured data $\dfrac{\partial \mathrm{u}}{\partial \nu} = g$ on $\partial\Omega$ so that F is the map $F : q \rightarrow \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}$.

The derivative of the map F in the direction h_1, $\frac{\partial F}{\partial q}.h_1$, is given by the normal derivative on $\partial\Omega$ of u', where

$$-\triangle u' + qu' = -u\, h_1 \quad \text{in } \Omega \qquad u' = 0 \quad \text{on } \partial\Omega.$$

To solve this we have the equation $Ax = c$ where c depends on h_1 and the previously computed u.

For the second derivative $\frac{\partial^2 F}{\partial q^2}.(h_1, h_2)$ we obtain

$$-\triangle u'' + qu'' = -u'(h_2)\, h_1 - u'(h_1)\, h_2 \quad \text{in } \Omega \qquad u'' = 0 \quad \text{on } \partial\Omega$$

leading to $Ax = d$ where d depends on (h_1, h_2) and the computed u, u'.

Suppose we are interested in recovering the coefficient q in

$$-\triangle u + qu = f(x) \qquad \text{in } \Omega$$
$$u = 0 \qquad \text{on } \partial\Omega$$

where we have the additional measured data $\dfrac{\partial \mathrm{u}}{\partial \nu} = g$ on $\partial\Omega$ so that F is the map $F : q \to \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}$.

Any numerical scheme designed to solve this equation will come down to solving a matrix equation of the form

$$Ax = b$$

where the matrix A will depend on the domain Ω and the coefficient q; the vector b will depend on the right hand side.

- The majority of the effort is in setting up and “inverting” A
- Most of the effort in computing both $\frac{\partial F}{\partial q}.h$ and $\frac{\partial^2 F}{\partial q^2}.(h_1, h_2)$ has already been done!

Schemes with frozen derivatives

Suppose that the previous benign situation was not possible. Perhaps we can compute the derivative of F about a known solution and hold this fixed during each iteration.
Using the previous notation and freezing about the case $q = 0$, this would compute $\frac{\partial F}{\partial q}(0).h$ as

$$-\triangle u' = -u\,h \qquad \text{in } \Omega$$

$$u' = 0 \qquad \text{on } \partial\Omega$$

If, for example, Ω is a ball in R^n then we can use standard representations to compute u' explicitly in terms of h.

Schemes with frozen derivatives

Suppose that the previous benign situation was not possible. Perhaps we can compute the derivative of F about a known solution and hold this fixed during each iteration.
Using the previous notation and freezing about the case $q = 0$, this would compute $\frac{\partial F}{\partial q}(0).h$ as

$$-\triangle u' = -u\,h \qquad \text{in } \Omega$$

$$u' = 0 \qquad \text{on } \partial\Omega$$

If, for example, Ω is a ball in R^n then we can use standard representations to compute u' explicitly in terms of h.

- The "folk theorem" that solutions of the metaharmonic equation behave very closely to harmonic functions gives credence to the belief that the frozen derivative will retain much of the relevant features of the full derivative.

Schemes with frozen derivatives

Suppose that the previous benign situation was not possible. Perhaps we can compute the derivative of F about a known solution and hold this fixed during each iteration.
Using the previous notation and freezing about the case $q = 0$, this would compute $\frac{\partial F}{\partial q}(0).h$ as

$$-\triangle u' = -u\,h \qquad \text{in } \Omega$$

$$u' = 0 \qquad \text{on } \partial\Omega$$

If, for example, Ω is a ball in R^n then we can use standard representations to compute u' explicitly in terms of h.

- The "folk theorem" that solutions of the metaharmonic equation behave very closely to harmonic functions gives credence to the belief that the frozen derivative will retain much of the relevant features of the full derivative.

The second derivative F'' can be handled similarly.

Consider the nonlinear equation

$$F(x) = g. \tag{1}$$

With starting guess x_0 let $\tilde{h}$ be computed by a Newton step,

$$F'[x_n]\,\tilde{h} = g - F(x_n). \tag{2}$$

Consider the nonlinear equation

$$F(x) = g. \tag{1}$$

With starting guess x_0 let $\tilde{h}$ be computed by a Newton step,

$$F'[x_n]\,\tilde{h} = g - F(x_n). \tag{2}$$

Then the next iteration $x_{n+1} = x_n + h$ is defined from the 2nd degree Taylor remainder by the solution of the linear equation

$$F'[x_n]h + \tfrac{1}{2}F''[x_n](\tilde{h}, h) = g - F(x_n). \tag{3}$$

Consider the nonlinear equation

$$F(x) = g. \tag{1}$$

With starting guess x_0 let $\tilde{h}$ be computed by a Newton step,

$$F'[x_n]\,\tilde{h} = g - F(x_n). \tag{2}$$

Then the next iteration $x_{n+1} = x_n + h$ is defined from the 2nd degree Taylor remainder by the solution of the linear equation

$$F'[x_n]h + \tfrac{1}{2}F''[x_n](\tilde{h}, h) = g - F(x_n). \tag{3}$$

or

$$\tilde{h} = (F'[x_n])^{-1}(g - F(x_n)) \qquad \text{(Predictor)}$$

and

$$\begin{aligned} h &= \left(F'[x_n] + \tfrac{1}{2}F''[x_n](\tilde{h}, \cdot)\right)^{-1}(g - F(x_n)) \\ x_{n+1} &= x_n + h \end{aligned} \qquad \text{(Corrector)}$$

Consider the nonlinear equation

$$F(x) = g. \tag{1}$$

With starting guess x_0 let $\tilde{h}$ be computed by a Newton step,

$$F'[x_n]\,\tilde{h} = g - F(x_n). \tag{2}$$

Theorem. *Let $\hat{x} \in U \subseteq X$ denote a solution of (1). Assume $F'[\hat{x}]$ admits a bounded inverse and F' and F'' are uniformly bounded in U. Then there exists $\delta > 0$ such that the iteration (2), (3) with starting guess $x_0 \in B(\hat{x}, \delta) = \{x \in X : \|\hat{x} - x\| < \delta\}$ converges quadraticly to $\hat{x}$. Additionally, if the second derivative is Lipschitz continuous, i.e.*

$$\|F''[x](h, \tilde{h}) - F''[y](h, \tilde{h})\| \leq L\|x - y\|\,\|h\|\,\|\tilde{h}\|$$

for all $x, y \in U$ with $h, \tilde{h} \in X$ and a constant $L > 0$, then

$$\|x_{n+1} - \hat{x}\| \leq c\,\|x_n - \hat{x}\|^3$$

holds for $n = 0, 1, 2, \ldots$ with a constant $c > 0$.

Consider the nonlinear equation

$$F(x) = g. \tag{1}$$

With starting guess x_0 let $\tilde{h}$ be computed by a Newton step,

$$F'[x_n]\,\tilde{h} = g - F(x_n). \tag{2}$$

Theorem. *Let $\hat{x} \in U \subseteq X$ denote a solution of (1). Assume $F'[\hat{x}]$ admits a bounded inverse and F' and F'' are uniformly bounded in U. Then there exists $\delta > 0$ such that the iteration (2), (3) with starting guess $x_0 \in B(\hat{x}, \delta) = \{x \in X : \|\hat{x} - x\| < \delta\}$ converges quadraticly to $\hat{x}$. Additionally, if the second derivative is Lipschitz continuous, i.e.*

$$\|F''[x](h,\tilde{h}) - F''[y](h,\tilde{h})\| \leq L\|x - y\|\,\|h\|\,\|\tilde{h}\|$$

for all $x, y \in U$ with $h, \tilde{h} \in X$ and a constant $L > 0$, then

$$\|x_{n+1} - \hat{x}\| \leq c\,\|x_n - \hat{x}\|^3$$

holds for $n = 0, 1, 2, \ldots$ with a constant $c > 0$.

To prove convergence under noise we need conditions that are often very difficult to check in practice – for example

$$\|F(y) - F(x) - F'[x](y - x)\| \leq C\|y - x\|\,\|F(y) - F(x)\|$$

The inverse Sturm-Liouville problem

Reconstruct the potential $q(x)$ from the equation

$$-y_n'' + q(x)y_n = \lambda_n y_n \qquad y_n(0) = y_n(1) = 0$$

given the eigenvalue sequence $\{\lambda_n\}$.

The inverse Sturm-Liouville problem

Reconstruct the potential $q(x)$ from the equation

$$-y_n'' + q(x)y_n = \lambda_n y_n \qquad y_n(0) = y_n(1) = 0$$

given the eigenvalue sequence $\{\lambda_n\}$.

The result of Borg is that the sequence $\{\lambda_n\}$ is sufficient to recover a symmetric function q, that is, $q(x) = q(1-x)$.
(A second spectrum $\{\mu_n\}$ arising from the boundary conditions $y_n(0) = y_n'(1) = 0$ will determine a general q).

For simplicity, we will assume that q is symmetric, $q(x) = q(1-x)$.

The inverse Sturm-Liouville problem

Reconstruct the potential $q(x)$ from the equation

$$-y_n'' + q(x)y_n = \lambda_n y_n \qquad y_n(0) = y_n(1) = 0$$

given the eigenvalue sequence $\{\lambda_n\}$.

The Gel'fand-Levitan equation gives

$$y(x) = \frac{1}{\sqrt{\lambda}}\left(\sin\sqrt{\lambda}x + \int_0^x u(x,t)\sin\sqrt{\lambda}t\,dt\right)$$

where $u(x,t;q)$ satisfies

$$u_{tt} - u_{xx} + q(x)u = 0 \quad \text{for } 0 < t \leq x < 1$$

with

$$u(x,0) = 0, \qquad u(x,x) = \tfrac{1}{2}\int_0^x q(s)\,ds$$

The function $y(x)$ satisfies the differential equation, the condition $y(0) = 0$ (and the normalisation $y'(0) = 1$).

We use the "additional" condition $y(1) = 0$ to determine q.

The inverse Sturm-Liouville problem

Reconstruct the potential $q(x)$ from the equation

$$-y_n'' + q(x)y_n = \lambda_n y_n \qquad y_n(0) = y_n(1) = 0$$

given the eigenvalue sequence $\{\lambda_n\}$.

The inverse problem can be reduced to:

- Use the eigenvalues and $x = 1$ in the Gelfand Levitan equation $y(x) = \frac{\sin\sqrt{\lambda}x}{\sqrt{\lambda}} + \int_0^x u(x,t)\frac{\sin\sqrt{\lambda}t}{\sqrt{\lambda}}\,dt$ to obtain

$$\int_0^1 g(t)\sin\sqrt{\lambda_n}t\,dt = -\sin\sqrt{\lambda_n} \qquad n = 1,\ 2,\ \ldots\ .$$

  The asymptotics of the λ_n guarantees this is uniquely solvable.

- Recover q from $u(1,t;q) = g(t)$ where $u(x,t;q)$ satisfies

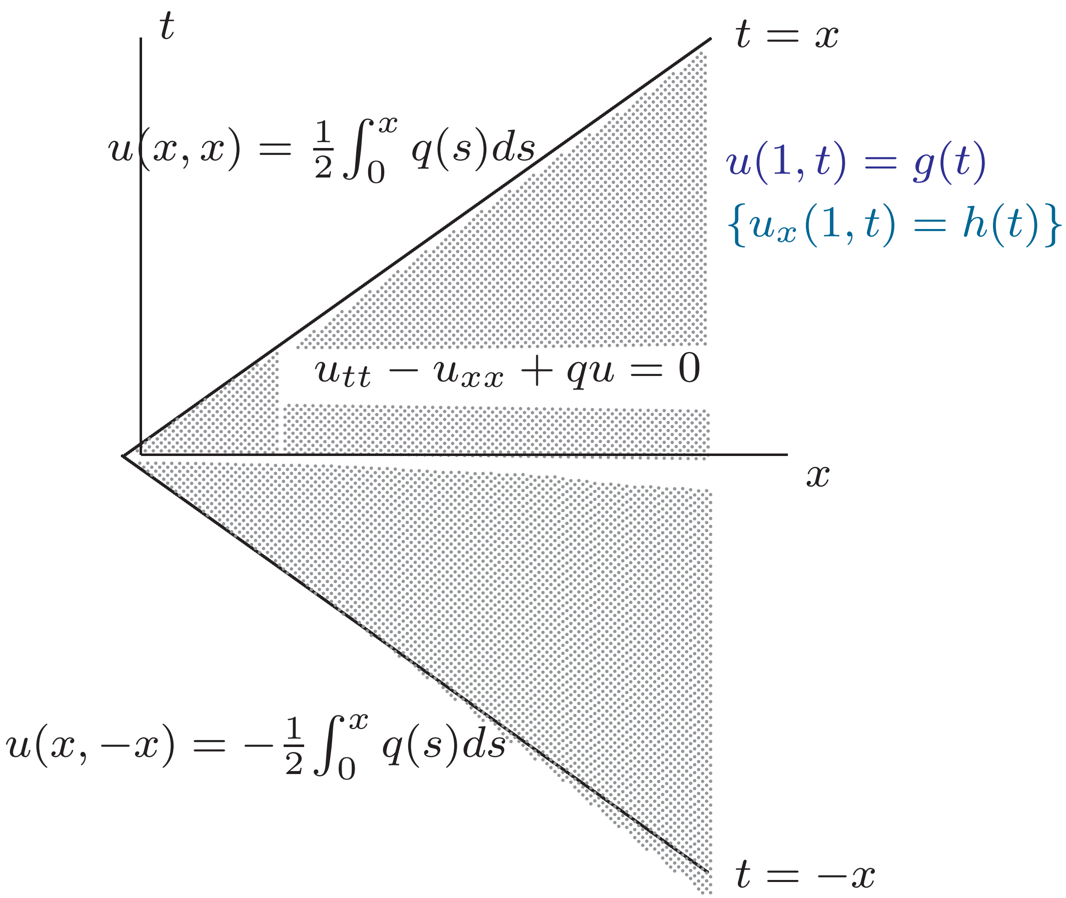
$$u_{tt} - u_{xx} + q(x)u = 0 \quad \text{for } 0 < t \le x < 1$$

  with

$$u(x,0) = 0, \qquad u(x,x) = \tfrac{1}{2}\int_0^x q(s)\,ds$$

Choose the map $F : L^2[0,1] \to L^2[0,1]$ defined by

$$F[q](t) := u_t(1,t;q) - g'(t)$$



Note: $u(x,t;0) = 0$.

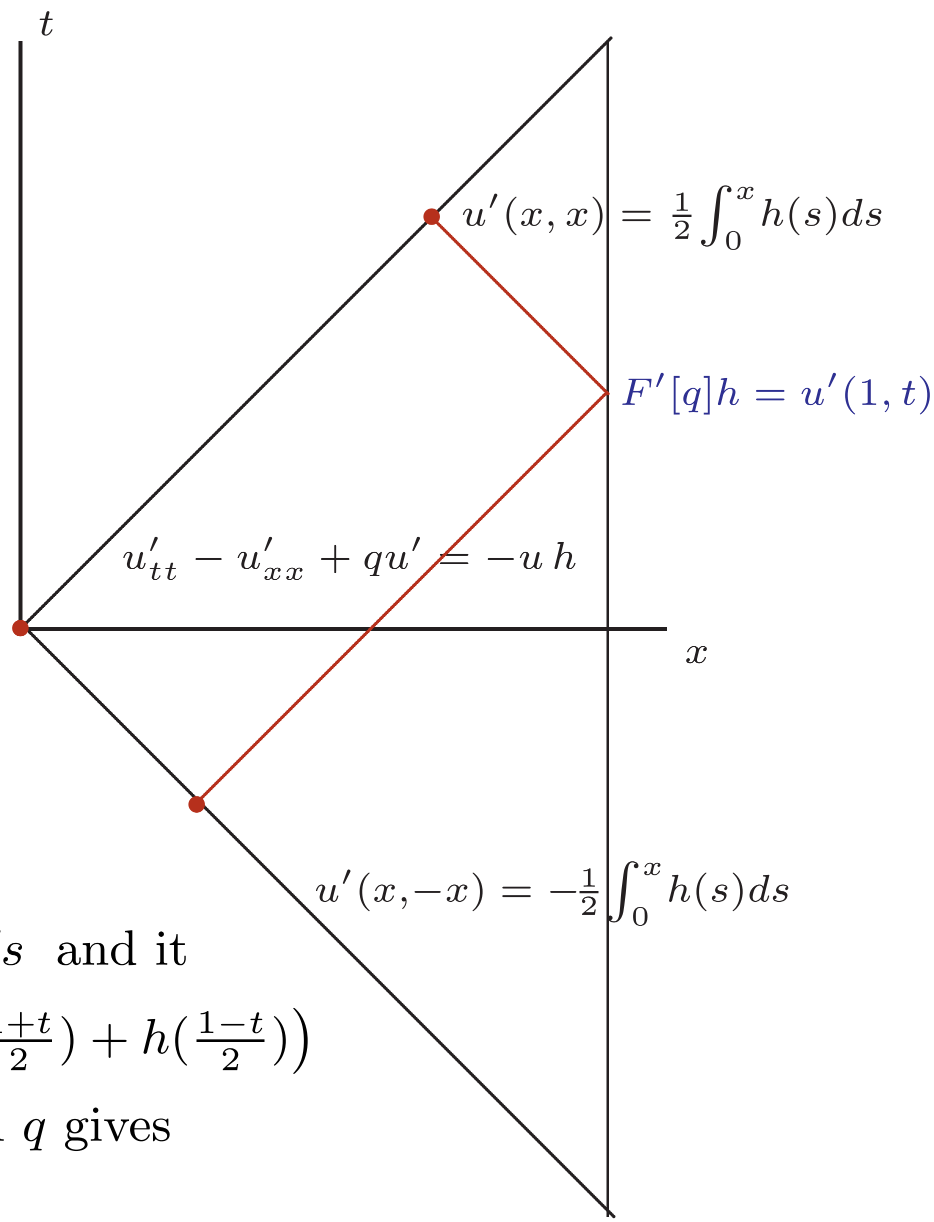
$F'[q].h$ is the solution of the Goursat problem

$u'_{tt} - u'_{xx} + q(x)u' = -hu$

for $0 < |t| \le x < 1$, with

$u'(x, \pm x) = \pm\frac{1}{2}\int_0^x h(s)\,ds$

evaluated at $x = 1$.



Now $u(x, t; 0) = 0$ and so $u'(x, t; 0; h)$ must satisfy

$u'_{tt} - u'_{xx} = 0, \quad 0 < |t| \le x < 1.$

Thus $u'(1, t; 0; h) = \frac{1}{2}\int_{\frac{x-t}{2}}^{\frac{x+t}{2}} h(s)\,ds$ and it follows that $u'_t(1, t; 0; h) = \frac{1}{2}\left(h(\frac{1+t}{2}) + h(\frac{1-t}{2})\right)$ and the symmetry assumption on q gives

$u'_t(1, t; 0; h) = h(\frac{1+t}{2})$.

The "frozen" Newton scheme now becomes

$$q_{n+1}(s) - q_n(s) = \tilde{g}'(2s-1) \quad \text{for } s \in [0,1].$$

where $\tilde{g}(t) = \frac{1}{2}(g(2t-1) - u(1, 2t-1; q_n)$.
As an initial approximation we have $\tilde{g}(t) = \frac{1}{2}g(2t-1)$.

We can compute the second derivative in a similar manner.

Implementing the Halley scheme with derivatives evaluated at $q = 0$ gives the predictor step as a Newton update,

$$h_1(t) = \tilde{g}'(t).$$

Our corrector formula is

$$q_{n+1}(s) - q_n(s) = h(t)$$

where $h(t)$ is the solution of the Volterra equation

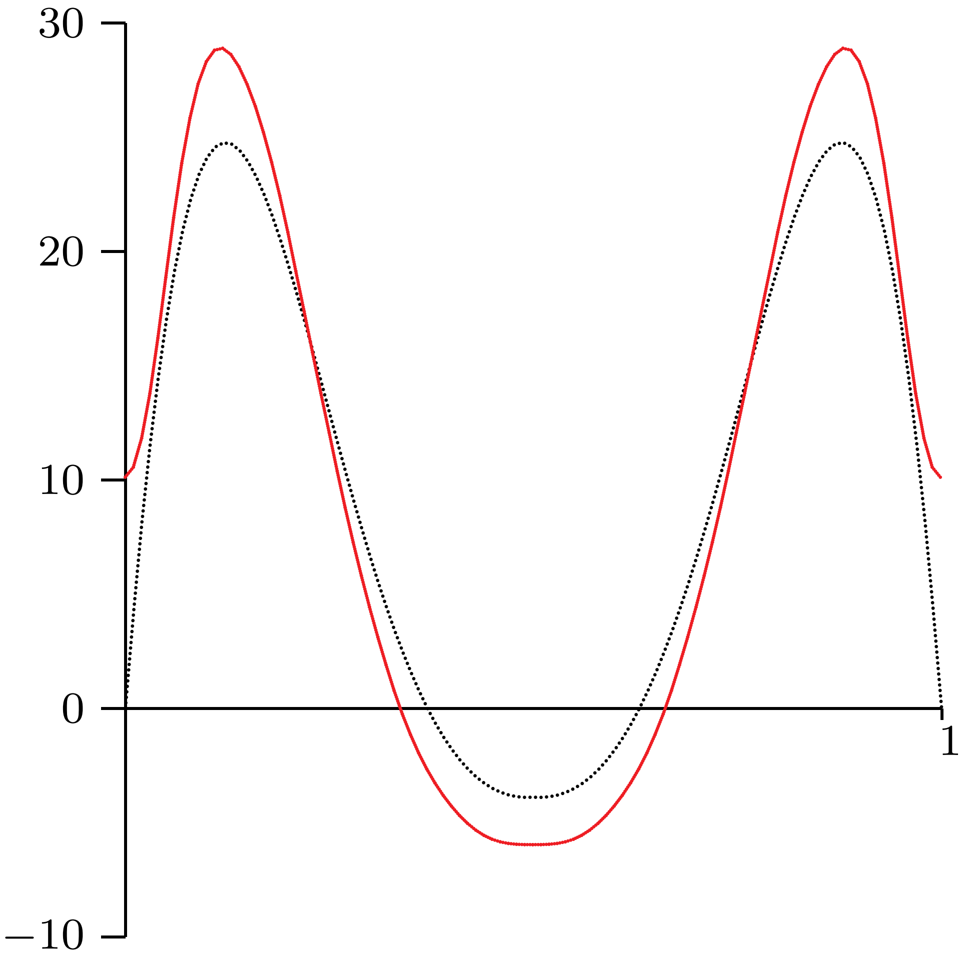
$$\tilde{g}'(t) = h(t) + \frac{1}{4}\int_0^t \Big(\tilde{g}(1-t-s) + \tilde{g}(t-s) - \tilde{g}(1+s-t) - \tilde{g}(t+s)\Big) h(s)\, ds.$$

This step is almost trivial to solve.

Reconstruction of $50\sin(3\pi x)e^{-5x}$, $x \in [0, 1/2]$ from the first 10 Dirichlet eigenvalues

Reconstruction of $50\sin(3\pi x)e^{-5x}$, $x \in [0, 1/2]$

from the first 10 Dirichlet eigenvalues



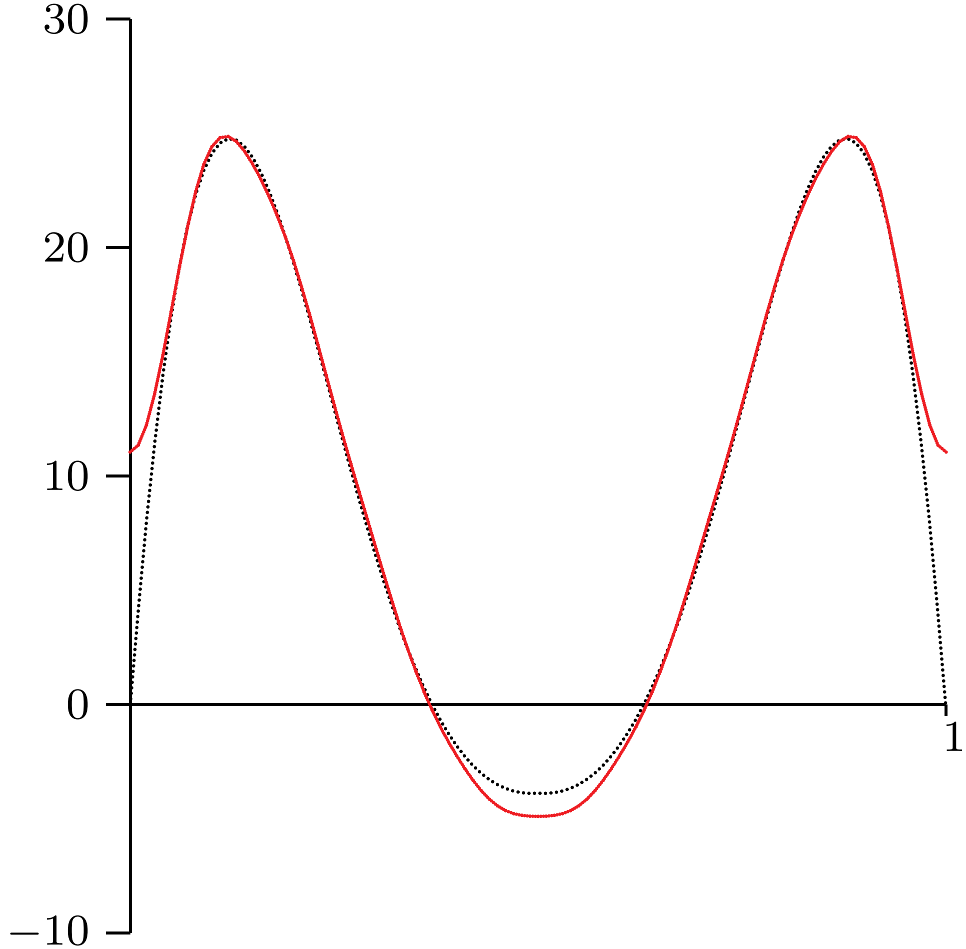
Using the predictor only

(single iteration step)

$$\frac{(1 - \|q_1 - \tilde{q}\|_2)}{\|\tilde{q}\|_2} = 0.83$$

Reconstruction of $50\sin(3\pi x)e^{-5x}$, $x \in [0, 1/2]$

from the first 10 Dirichlet eigenvalues



Using the corrector

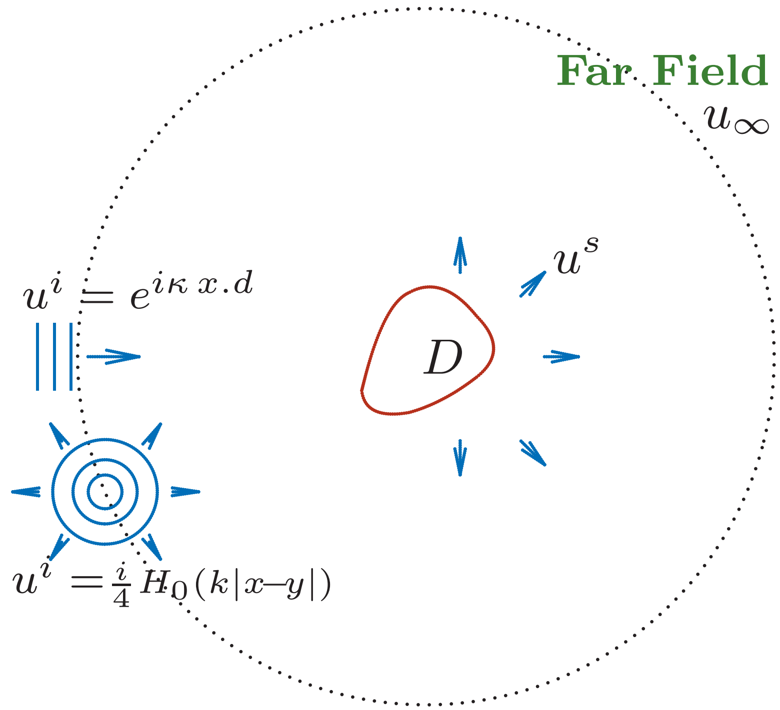
(single iteration step)

$$\frac{(1 - \|q_1 - \tilde{q}\|_2)}{\|\tilde{q}\|_2} = 0.97$$

An Inverse Scattering Problem

A time-harmonic acoustic or electromagnetic plane wave $u^i = e^{i\kappa\, x.d}$ or point source $\frac{i}{4}\mathrm{H}_0(\mathrm{k}|\mathrm{x}-\mathrm{y}|)$ is fired at a cylindrical obstacle D of unknown shape and location.

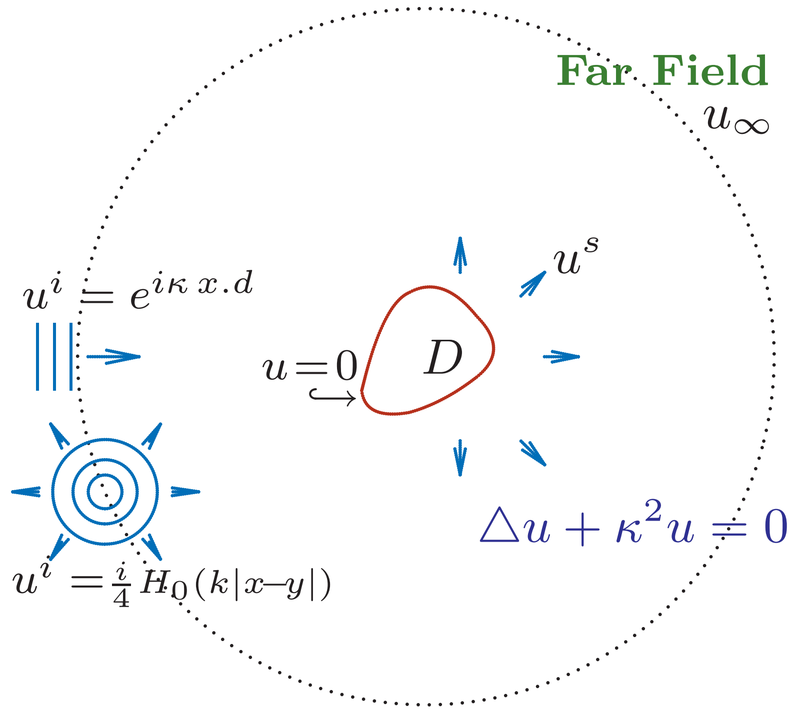
The wave u^s scattered from this object is measured at "infinity" – the far field pattern u_∞.



An Inverse Scattering Problem

A time-harmonic acoustic or electromagnetic plane wave $u^i = e^{i\kappa\, x.d}$ or point source $\frac{i}{4}\mathrm{H}_0(\mathrm{k}|\mathrm{x}-\mathrm{y}|)$ is fired at a cylindrical obstacle D of unknown shape and location.

The wave u^s scattered from this object is measured at "infinity" – the far field pattern u_∞.



The total wave $u = u^i + u^s$ is modeled by an exterior boundary value problem for the Helmholtz equation:

$$\triangle u + \kappa^2 u = 0 \quad \text{in } I\!R^2 \setminus \bar{D}$$

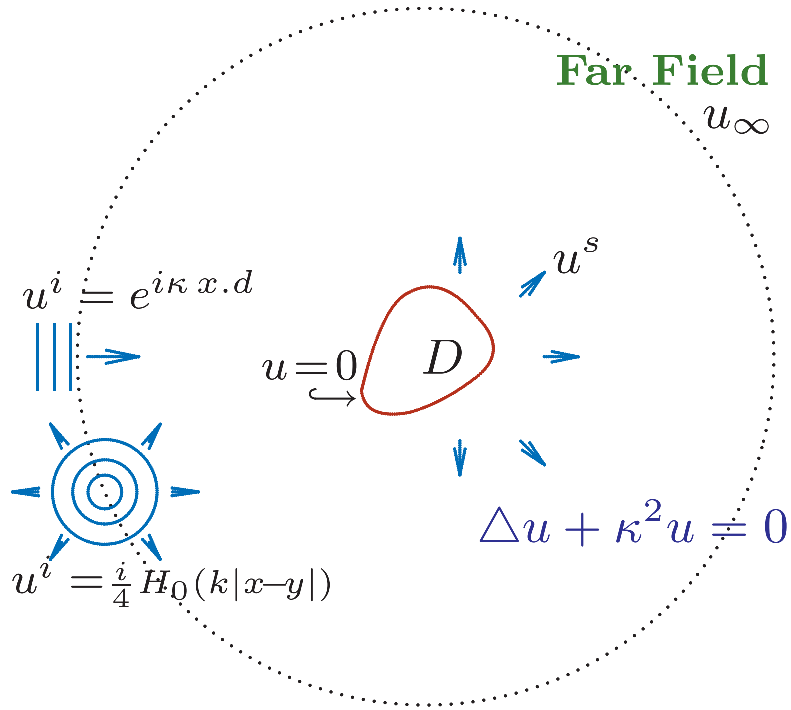
with positive wave number κ and Dirichlet boundary condition

$$u = 0 \quad \text{on } \partial D, \qquad \text{"sound soft object"}$$

An Inverse Scattering Problem

A time-harmonic acoustic or electromagnetic plane wave $u^i = e^{i\kappa\, x.d}$ or point source $\frac{i}{4}\mathrm{H}_0(\mathrm{k}|\mathrm{x}-\mathrm{y}|)$ is fired at a cylindrical obstacle D of unknown shape and location.

The wave u^s scattered from this object is measured at “infinity” – the far field pattern u_∞.



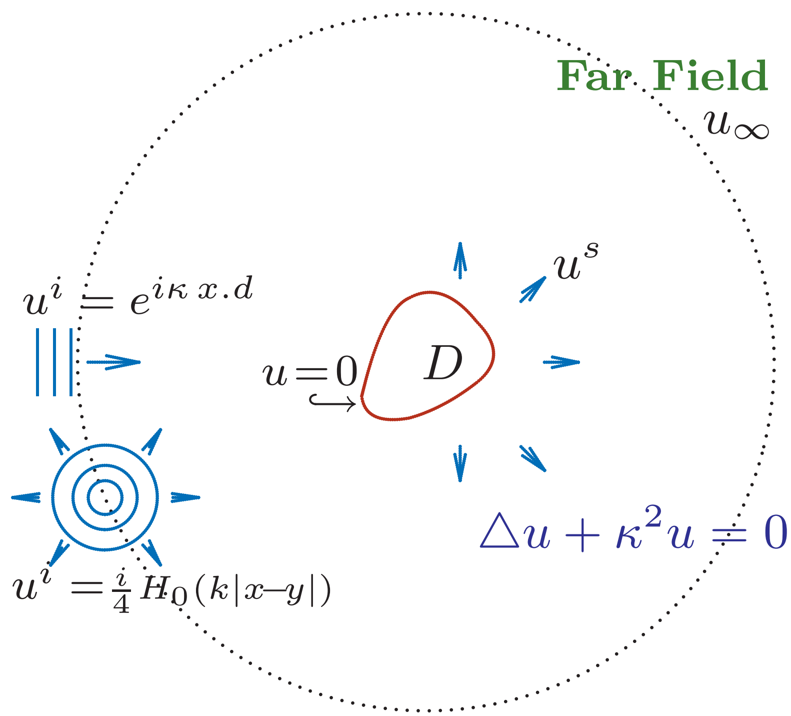
The scattered wave u^s is required to satisfy the Sommerfeld radiation condition uniformly in all directions $\hat{x} = x/|x|$

$$\frac{\partial u^s}{\partial r} - i\kappa u^s = o\left(\frac{1}{\sqrt{r}}\right), \quad r = |x| \to \infty, .$$

An Inverse Scattering Problem

A time-harmonic acoustic or electromagnetic plane wave $u^i = e^{i\kappa\, x.d}$ or point source $\frac{i}{4}\mathrm{H}_0(\mathrm{k}|\mathrm{x}-\mathrm{y}|)$ is fired at a cylindrical obstacle D of unknown shape and location.

The wave u^s scattered from this object is measured at "infinity" – the far field pattern u_∞.



This implies the asymptotic behaviour

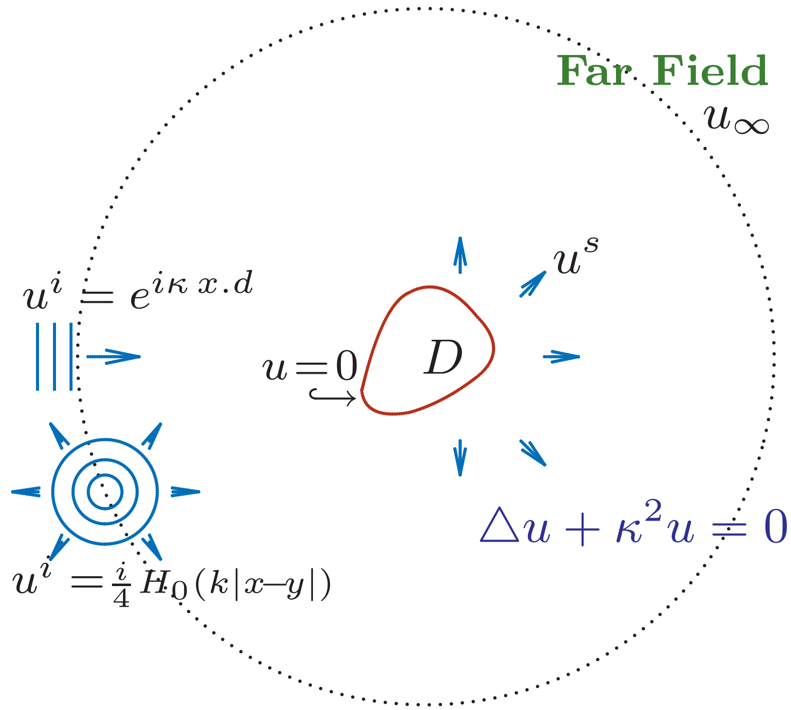
$$u^s(x) = \frac{e^{i\kappa x}}{\sqrt{|x|}}\left(u_\infty(\hat{x}; d) + O\left(\frac{1}{|x|}\right)\right), \quad |x| \to \infty, .$$

The *Amplitude factor* u_∞ is the *far field pattern* of the scattered wave.

An Inverse Scattering Problem

A time-harmonic acoustic or electromagnetic plane wave $u^i = e^{i\kappa\, x.d}$ or point source $\frac{i}{4}\mathrm{H}_0(\mathrm{k}|\mathrm{x}-\mathrm{y}|)$ is fired at a cylindrical obstacle D of unknown shape and location.

The wave u^s scattered from this object is measured at "infinity" – the far field pattern u_∞.



This implies the asymptotic behaviour

$$u^s(x) = \frac{e^{i\kappa x}}{\sqrt{|x|}}\left(u_\infty(\hat{x};d) + O\left(\frac{1}{|x|}\right)\right), \quad |x| \to \infty, .$$

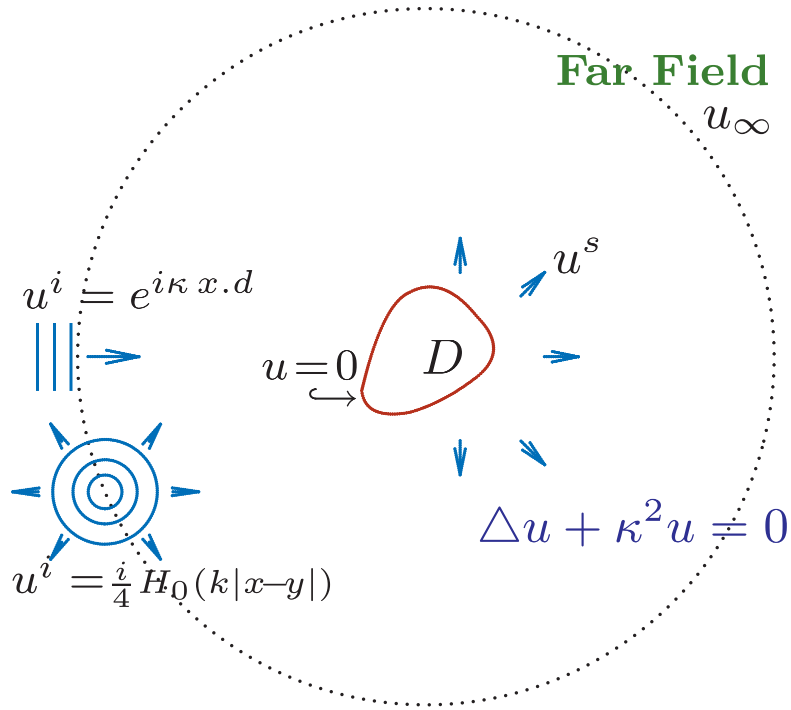
The *Amplitude factor* u_∞ is the *far field pattern* of the scattered wave.

$F : X \to L^2(S^1)$ maps a set X of admissible boundaries onto the far field pattern. F is nonlinear and *ill-posed*.

An Inverse Scattering Problem

A time-harmonic acoustic or electromagnetic plane wave $u^i = e^{i\kappa\, x.d}$ or point source $\frac{i}{4}\mathrm{H}_0(\mathrm{k}|\mathrm{x}-\mathrm{y}|)$ is fired at a cylindrical obstacle D of unknown shape and location.

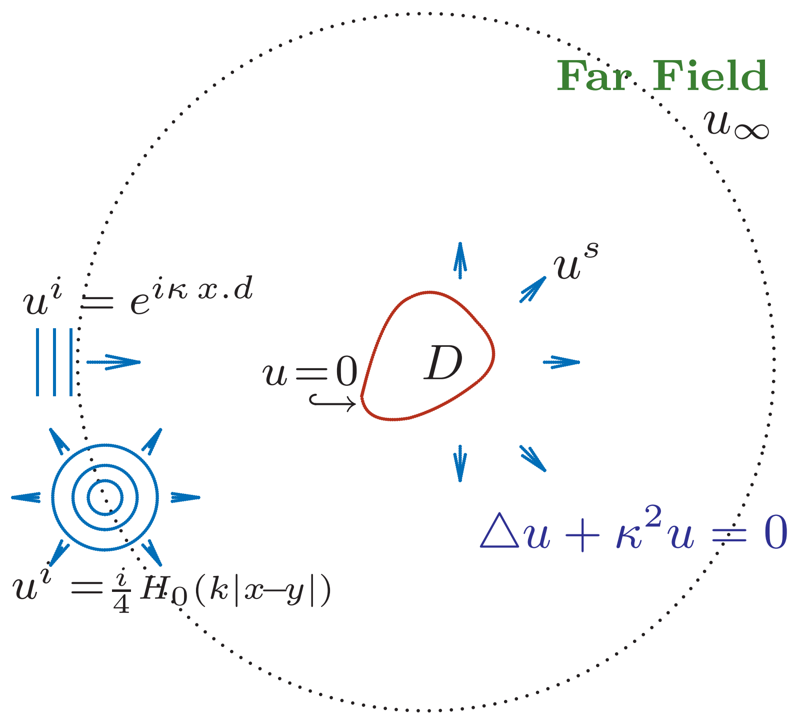
The wave u^s scattered from this object is measured at “infinity” – the far field pattern u_∞.



Assume starlike obstacles: $\partial D := \left\{ x = q(t)\binom{\cos t}{\sin t} : t \in [0, 2\pi) \right\}$ with a 2π–periodic positive function $r \in C^3(R)$.

An Inverse Scattering Problem

A time-harmonic acoustic or electromagnetic plane wave $u^i = e^{i\kappa\, x.d}$ or point source $\frac{i}{4}\mathrm{H}_0(\mathrm{k}|\mathrm{x}-\mathrm{y}|)$ is fired at a cylindrical obstacle D of unknown shape and location. The wave u^s scattered from this object is measured at “infinity” – the far field pattern u_∞.



Assume starlike obstacles: $\partial D := \left\{ x = q(t)\binom{cos\, t}{sin\, t} : t \in [0, 2\pi) \right\}$ with a 2π–periodic positive function $r \in C^3(R)$.

For smooth variations h of ∂D, the derivative is defined by

$$\lim_{\|h\|_{C^1}\to 0} \frac{1}{\|\mathrm{h}\|_{\mathrm{C}^1}} \|F(\partial D_h) - F(\partial D) - F'[\partial D]h\| = 0$$

ν is the unit outward normal, and H the mean curvature of ∂D. $h_\nu := h\cdot\nu$.

Theorem. *F is twice differentiable. F' is represented by the far field pattern $F'[\partial D]\, h = u'_\infty$ of the solution u' of the exterior Dirichlet problem,*

$$\triangle u' + k^2 u' = 0 \quad in\ R^2 \backslash \overline{D} \qquad u' = -h_\nu \frac{\partial u}{\partial \nu} \quad on\ \partial D$$

$F''[\partial D](h_1, h_2) = u''_\infty$ where u''_∞ is the far field pattern of the radiating solution $u'' \in H^1_{loc}(R^2 \backslash \overline{D})$ of the exterior Dirichlet problem

$$\triangle\, u'' + k^2 u'' = 0 \quad \text{in } \mathbf{R}^2 \backslash \overline{D},$$

$$u'' = -h_{1,\nu} \frac{\partial u'_2}{\partial \nu} - h_{2,\nu} \frac{\partial u'_1}{\partial \nu} + \big(h_{1,\nu} h_{2,\nu} - h_{1,\tau} h_{2,\tau}\big) H \frac{\partial u}{\partial \nu}$$
$$+ \Big(h_{1,\tau}(\tau \cdot \nabla(h_{2,\nu})) + h_{2,\tau}(\tau \cdot \nabla(h_{1,\nu}))\Big) \frac{\partial u}{\partial \nu} \quad \text{on } \partial D$$

u is the solution of the scattering problem, u'_j ($j = 1, 2$) is the solution of the boundary value problem with respect to the variation h_j.

Theorem. *F is twice differentiable. F' is represented by the far field pattern $F'[\partial D]\, h = u'_\infty$ of the solution u' of the exterior Dirichlet problem,*

$$\triangle u' + k^2 u' = 0 \quad in\ R^2 \backslash \overline{D} \qquad u' = -h_\nu \frac{\partial u}{\partial \nu} \quad on\ \partial D$$

$F''[\partial D](h_1, h_2) = u''_\infty$ where u''_∞ is the far field pattern of the radiating solution $u'' \in H^1_{loc}(R^2 \backslash \overline{D})$ of the exterior Dirichlet problem

$$\triangle\, u'' + k^2 u'' = 0 \quad \text{in } \mathbf{R}^2 \backslash \overline{D},$$

$$u'' = -h_{1,\nu} \frac{\partial u'_2}{\partial \nu} - h_{2,\nu} \frac{\partial u'_1}{\partial \nu} + \big(h_{1,\nu} h_{2,\nu} - h_{1,\tau} h_{2,\tau} \big) H \frac{\partial u}{\partial \nu}$$
$$+ \Big(h_{1,\tau} (\tau \cdot \nabla (h_{2,\nu})) + h_{2,\tau} (\tau \cdot \nabla (h_{1,\nu})) \Big) \frac{\partial u}{\partial \nu} \quad \text{on } \partial D$$

u is the solution of the scattering problem, u'_j ($j = 1, 2$) is the solution of the boundary value problem with respect to the variation h_j.

Note: We can show that F' is injective, $F'[\partial D] h = 0$ implies $h_\nu = 0$

(the argument uses a combination of Rellich's and Holmgren's theorems).

Theorem. *F is twice differentiable. F' is represented by the far field pattern $F'[\partial D]\, h = u'_\infty$ of the solution u' of the exterior Dirichlet problem,*

$$\triangle u' + k^2 u' = 0 \quad in\ R^2 \backslash \overline{D} \qquad u' = -h_\nu \frac{\partial u}{\partial \nu} \quad on\ \partial D$$

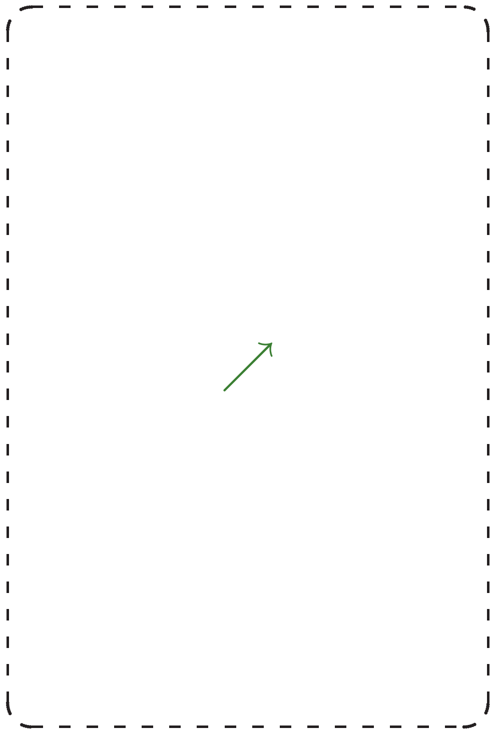
$F''[\partial D](h_1, h_2) = u''_\infty$ where u''_∞ is the far field pattern of the radiating solution $u'' \in H^1_{loc}(R^2 \backslash \overline{D})$ of the exterior Dirichlet problem

$$\triangle\, u'' + k^2 u'' = 0 \quad \text{in } \mathbf{R}^2 \backslash \overline{D},$$

$$u'' = -h_{1,\nu} \frac{\partial u_2'}{\partial \nu} - h_{2,\nu} \frac{\partial u_1'}{\partial \nu} + \big(h_{1,\nu} h_{2,\nu} - h_{1,\tau} h_{2,\tau}\big) H \frac{\partial u}{\partial \nu}$$
$$+ \Big(h_{1,\tau}(\tau \cdot \nabla(h_{2,\nu})) + h_{2,\tau}(\tau \cdot \nabla(h_{1,\nu}))\Big) \frac{\partial u}{\partial \nu} \quad \text{on } \partial D$$

u is the solution of the scattering problem, u'_j ($j = 1, 2$) is the solution of the boundary value problem with respect to the variation h_j.

Note: The largest part of the computation of F, F' and F'' is the common inversion of the matrix representing $\triangle v + k^2 v$.

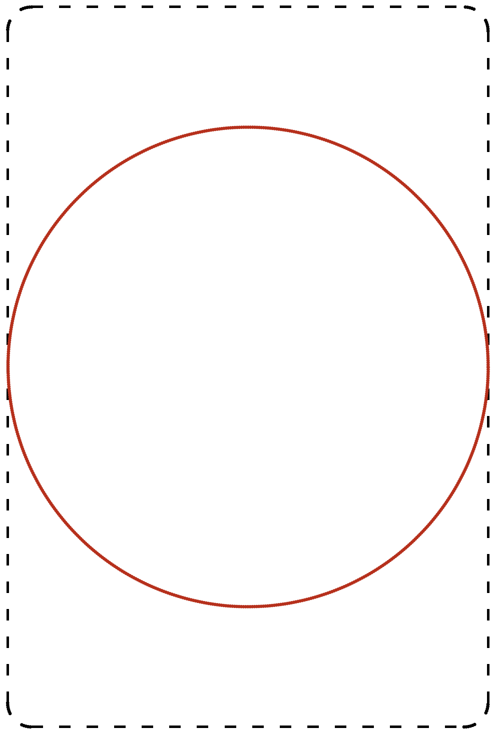


Reconstruction of a sound soft rectangular object from a single incident field using accurate data ($\sim 0.1\%$ noise).

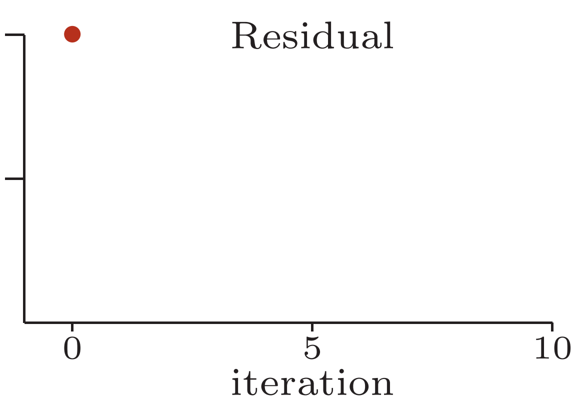
The size of the obstacle is 1.5×1 units and the value of κ is one.

Exact Obstacle

Direction of Incident Wave

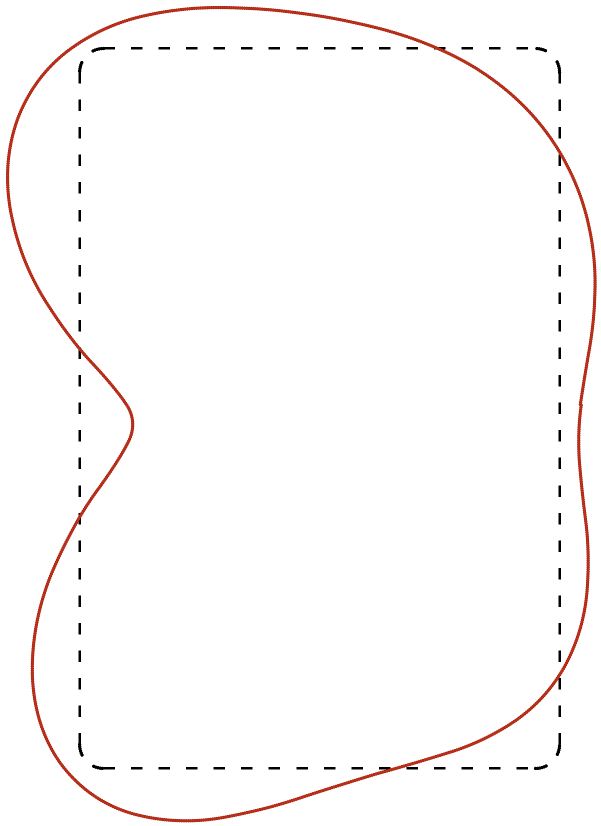


Initial Approximation

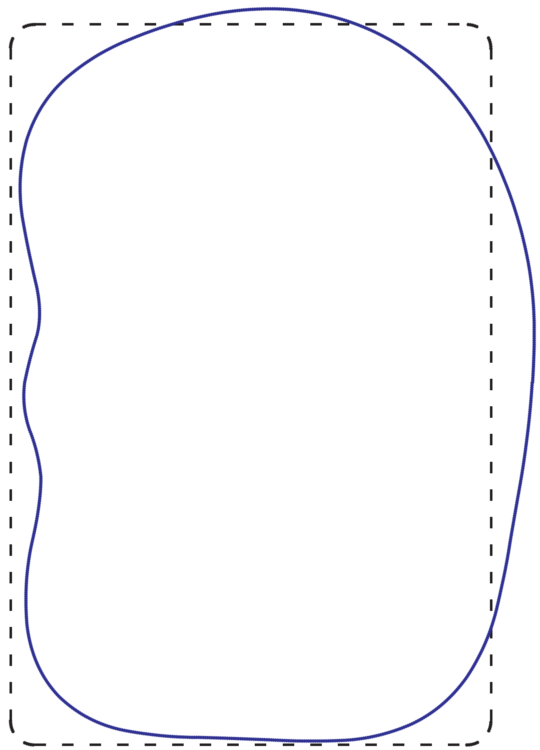


The graphic below plots the relative residual is against iteration number, $\|F(q_n) - u_\infty\| / \|F(q_0) - u_\infty\|$

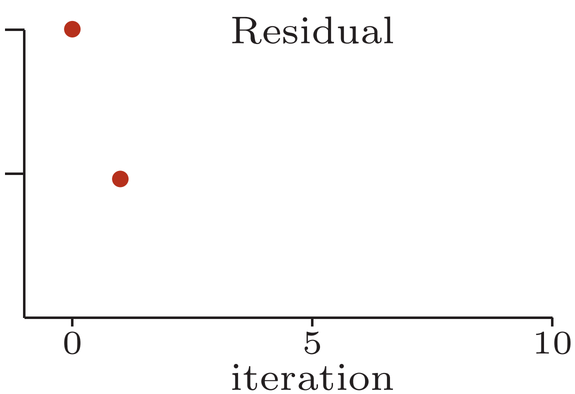
Tichonov regularisation is used.

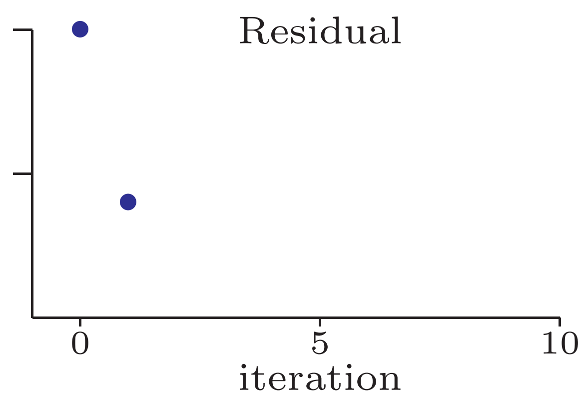


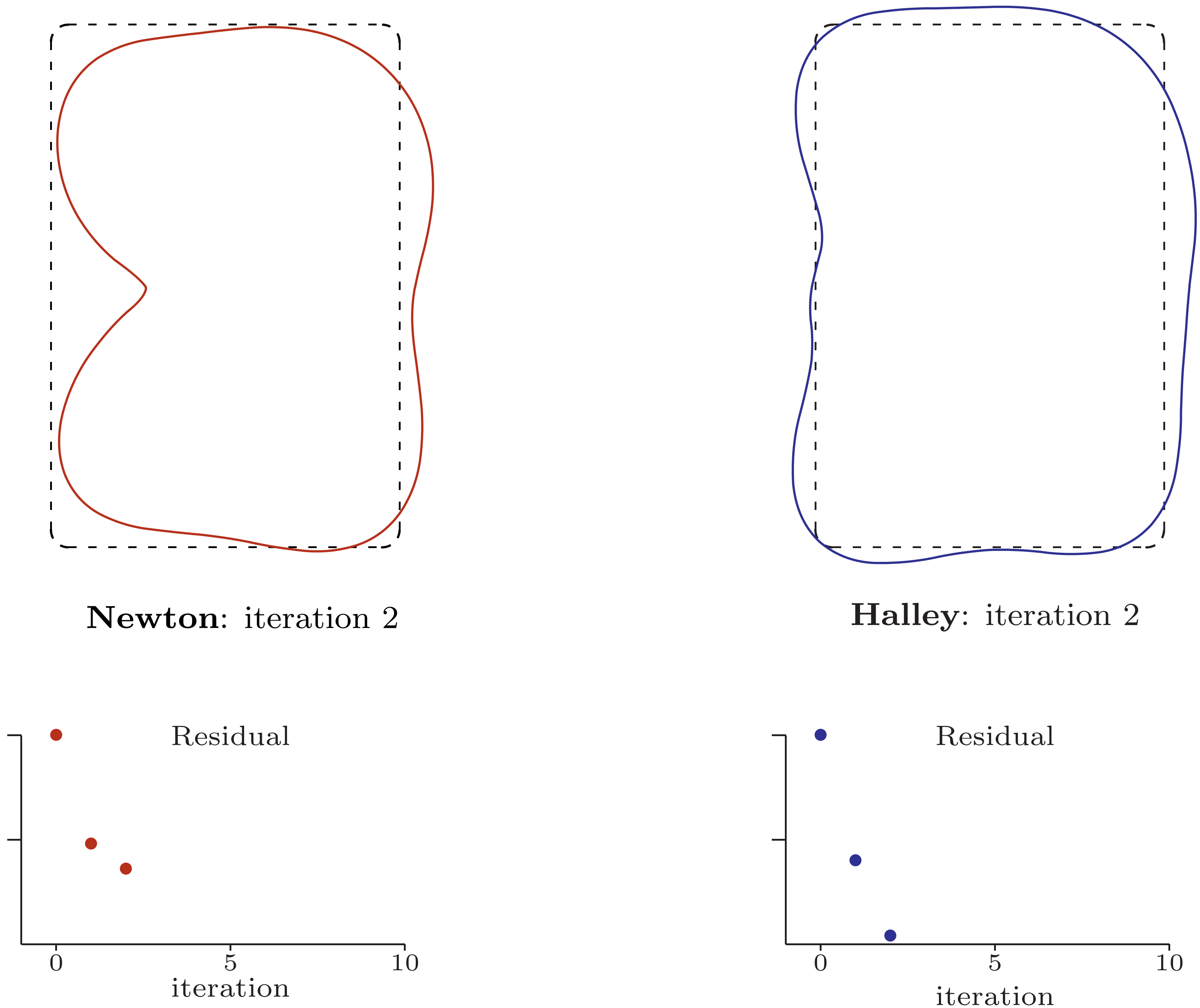
Newton: iteration 1

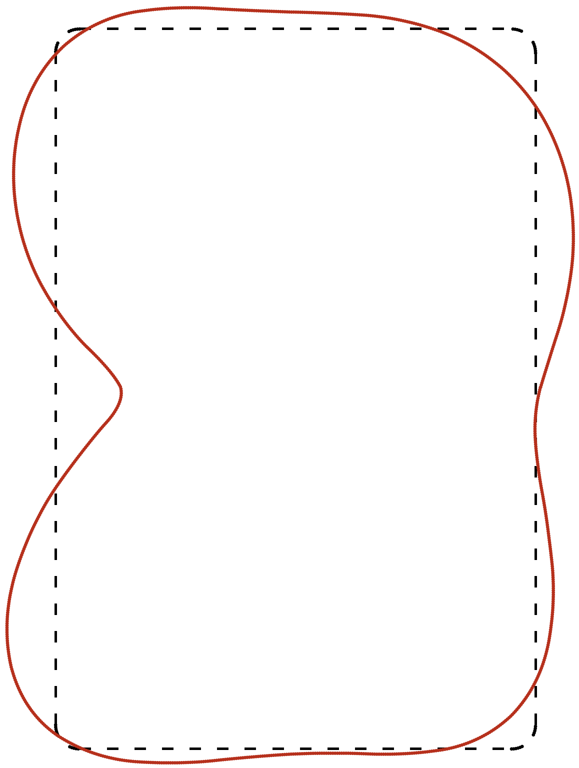


Halley: iteration 1

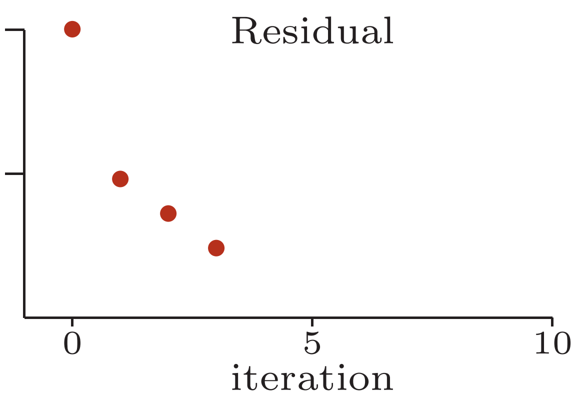


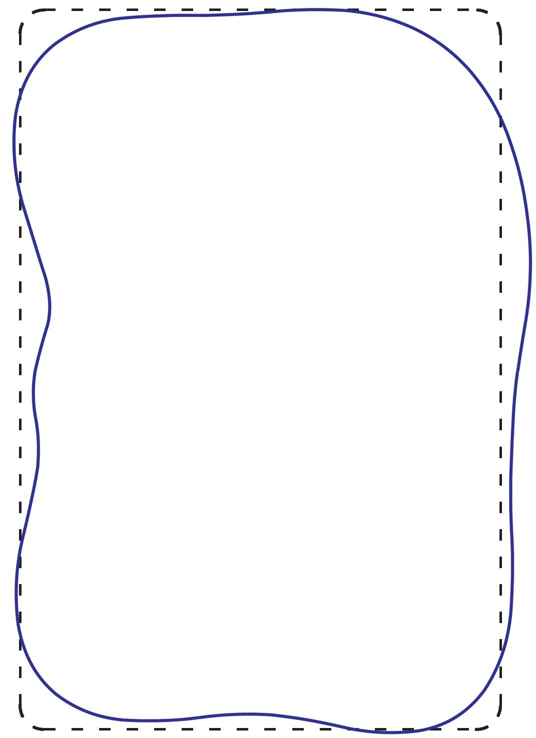



Newton: iteration 2
Halley: iteration 2
Residual
0
5
10
iteration
Residual
0
5
10
iteration

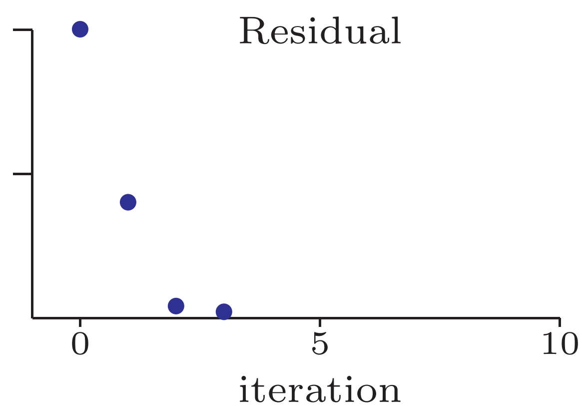


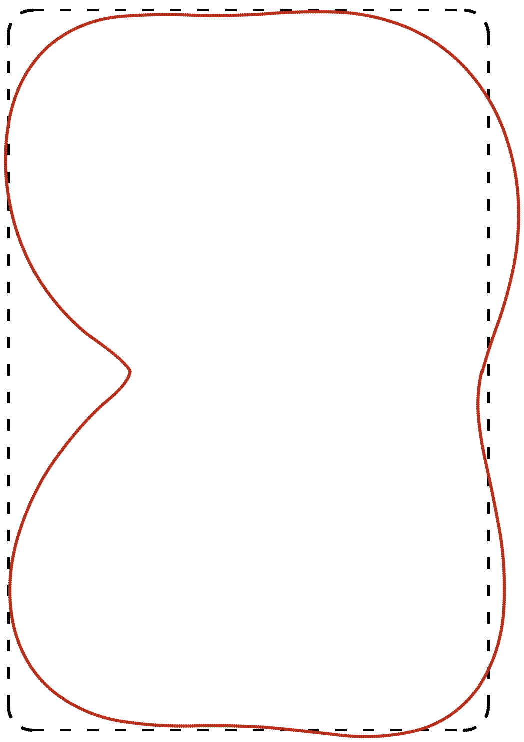
Newton: iteration 3



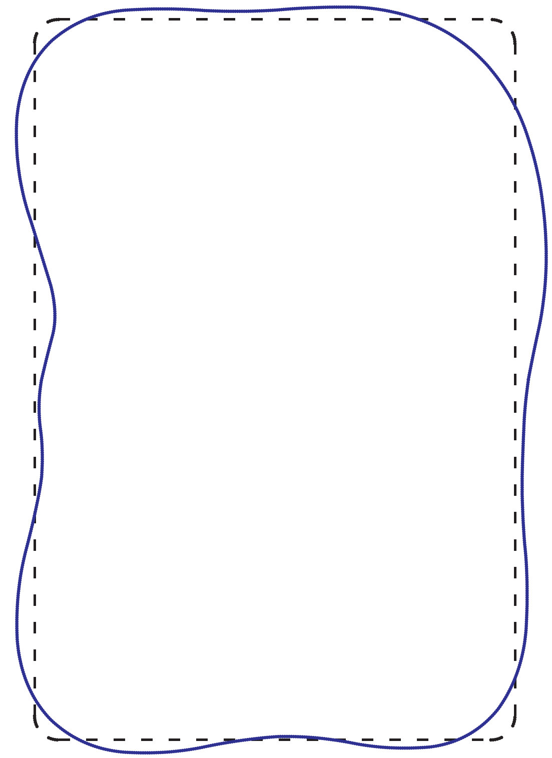


Halley: iteration 3

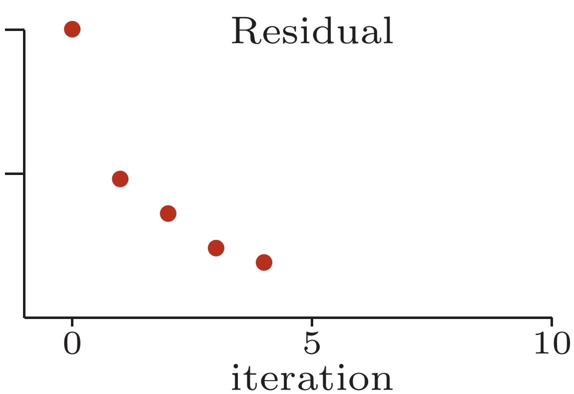


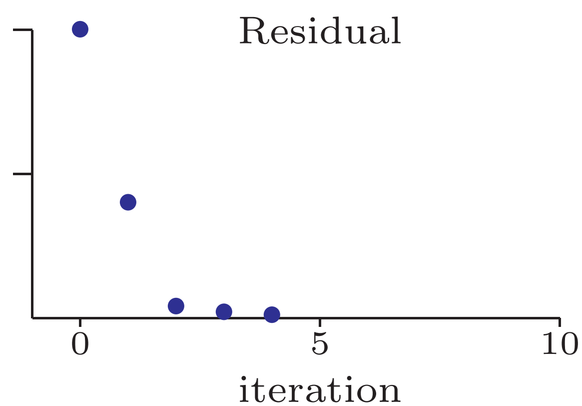


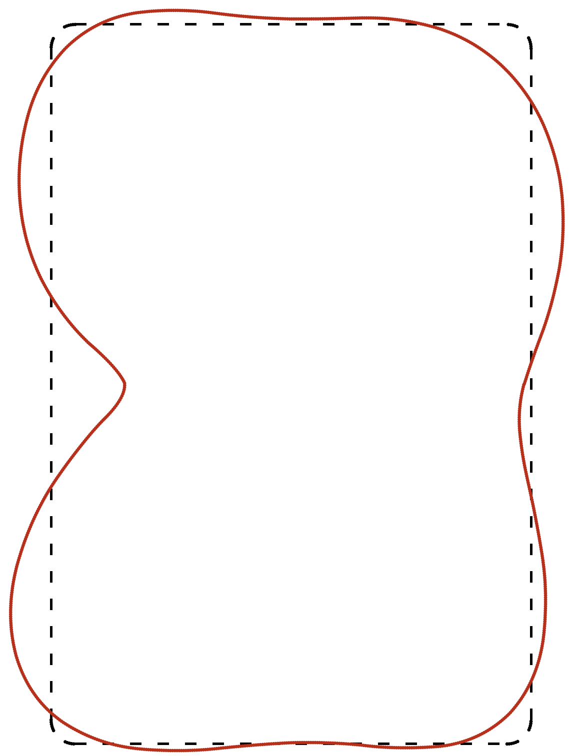
Newton: iteration 4



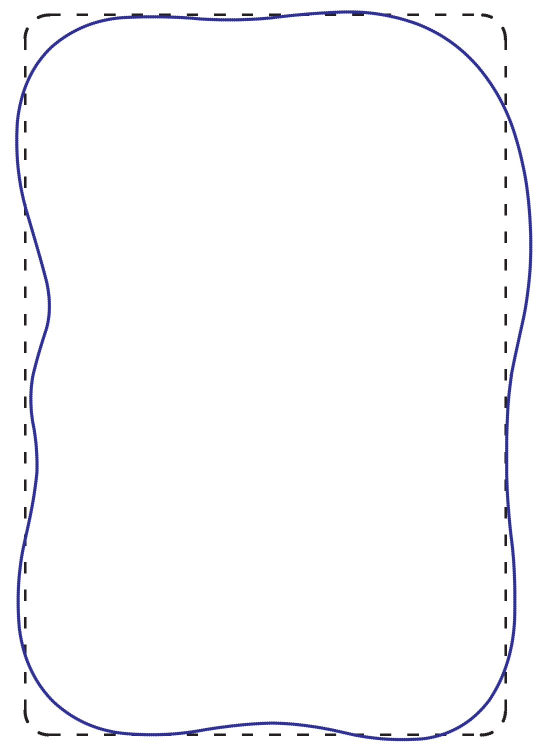
Halley: iteration 4



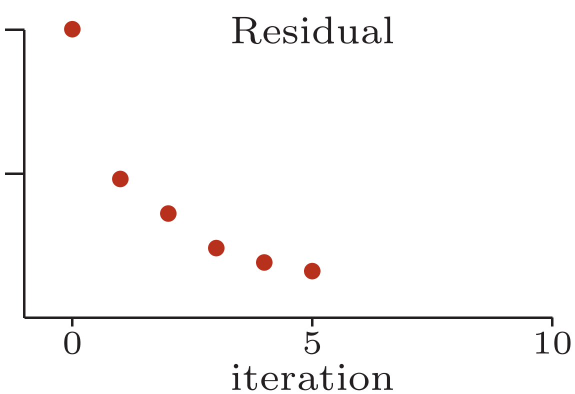


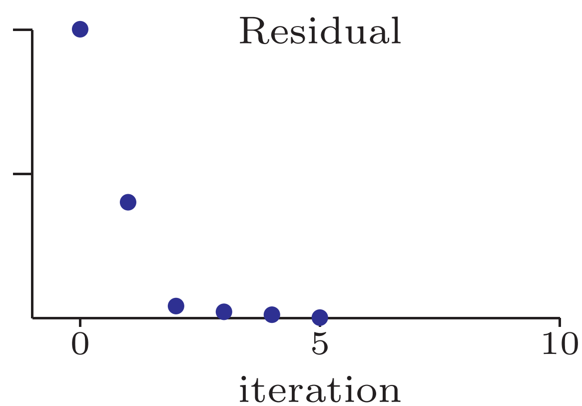


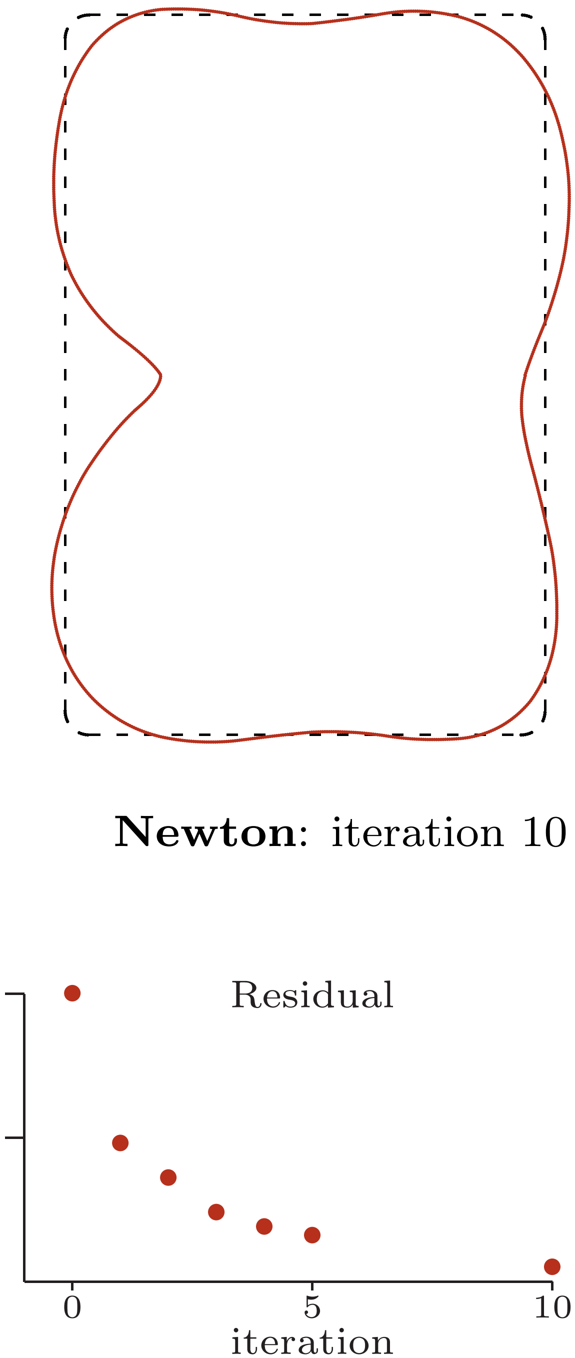
Newton: iteration 5

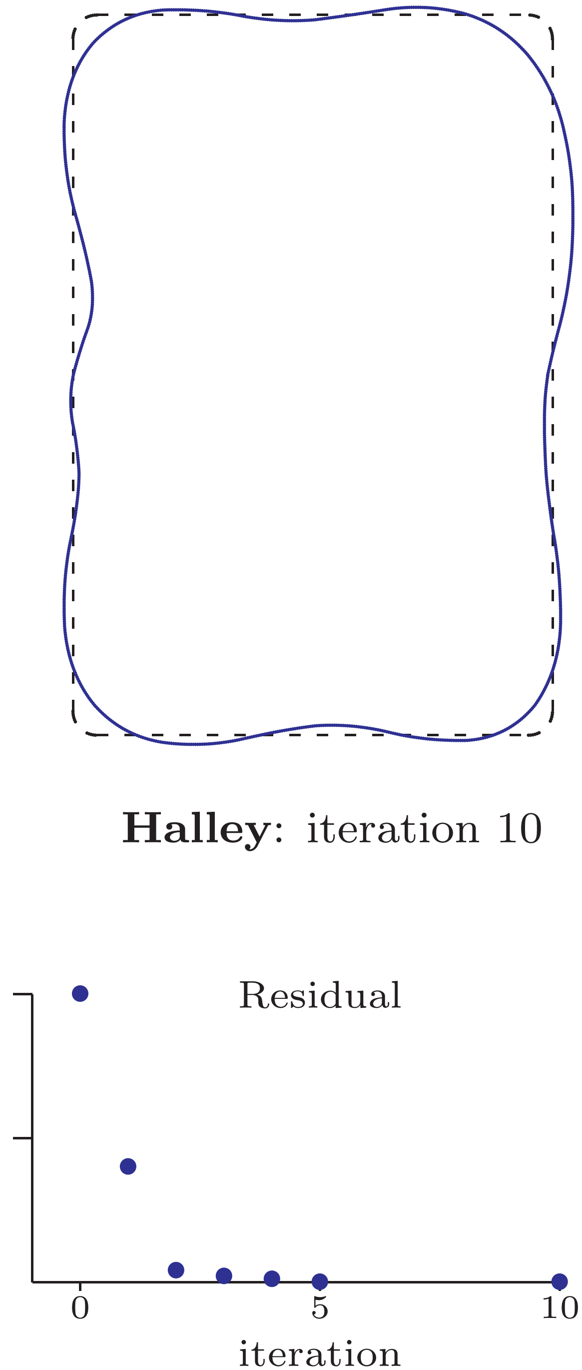


Halley: iteration 5

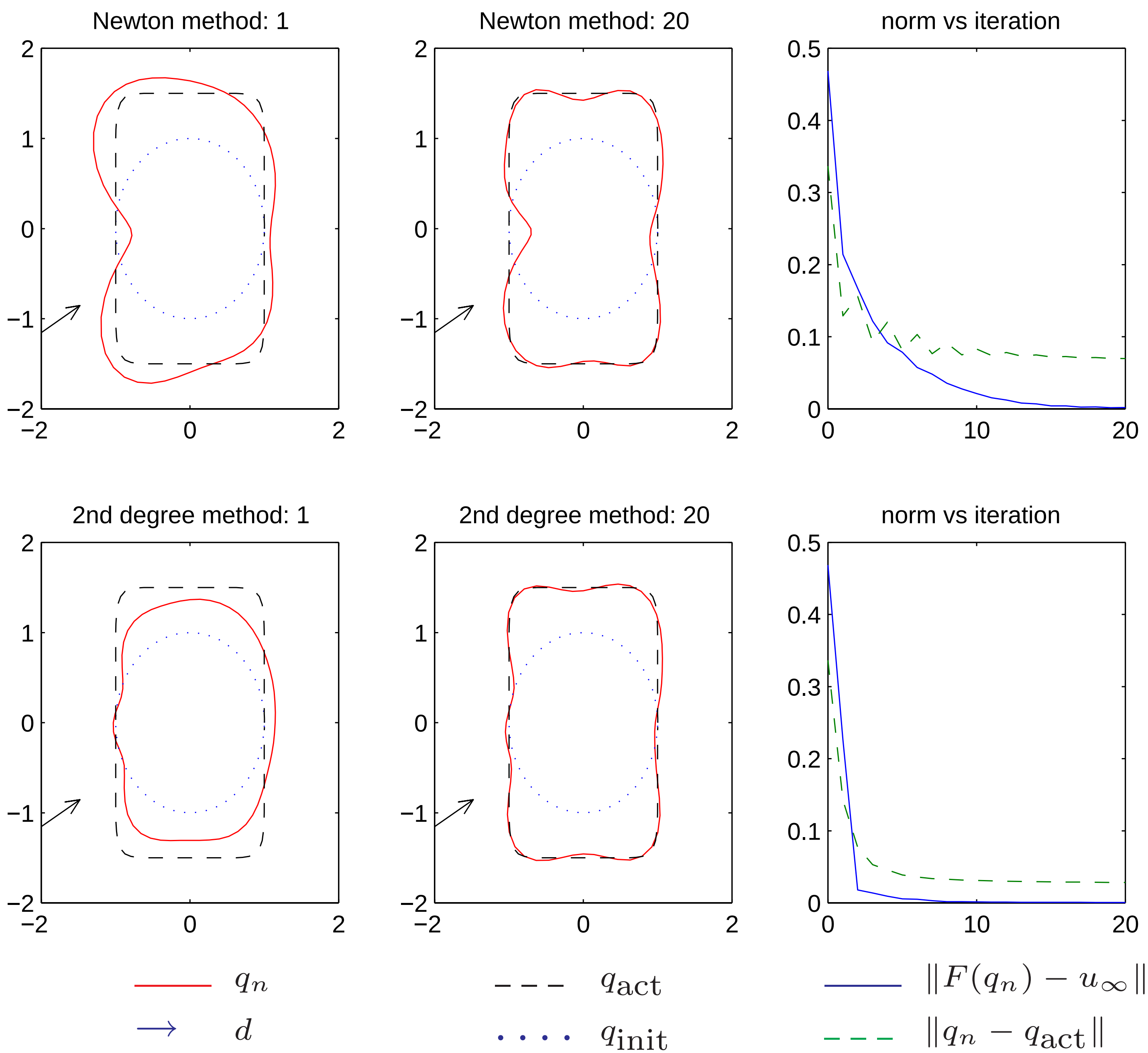




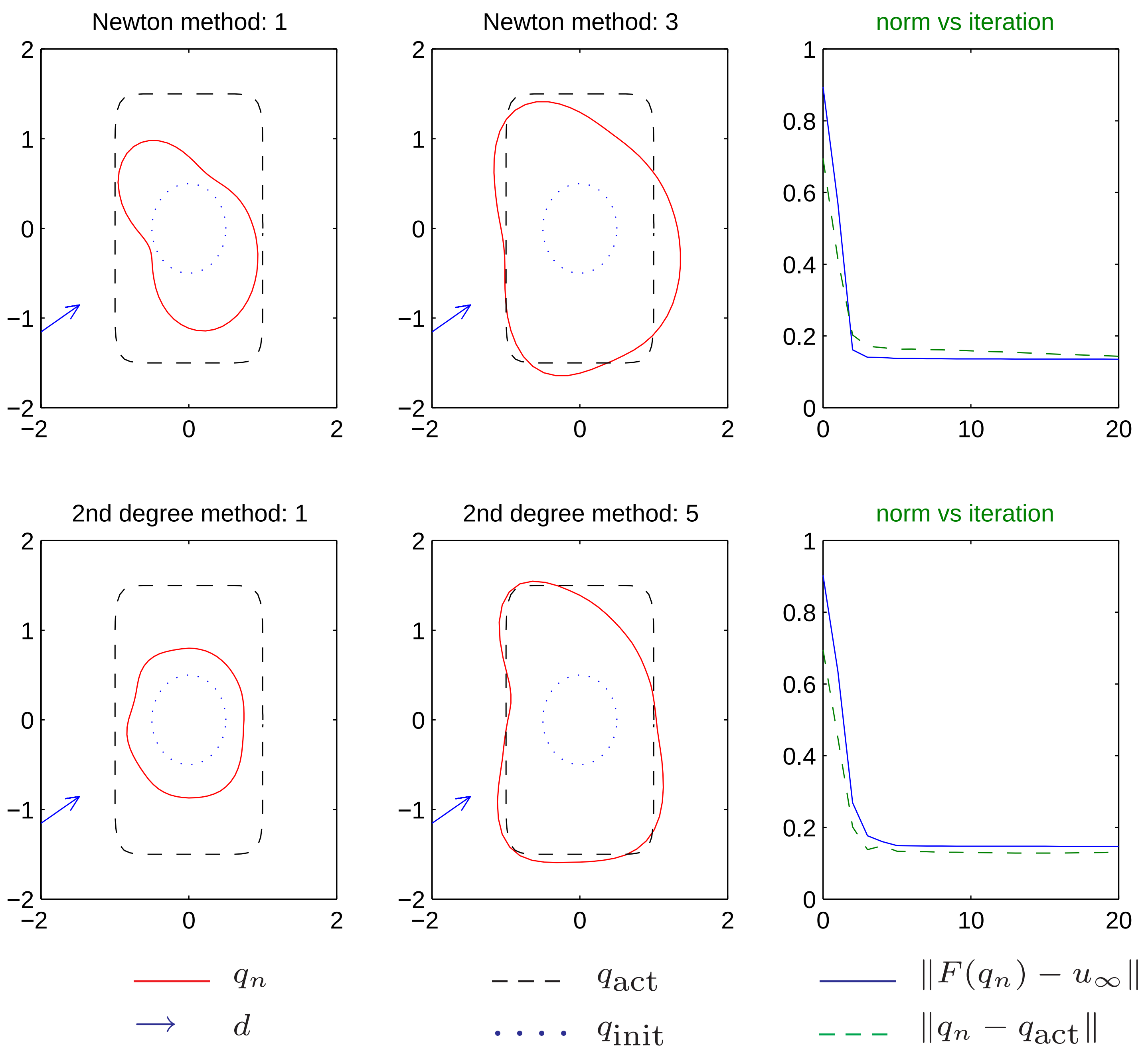

Newton: iteration 10
Residual
0
5
10
iteration


Halley: iteration 10
Residual
0
5
10
iteration

Reconstructions from noise-free data



Reconstructions from data with 10% noise



Some Open Issues

We are solving the Newton equation $F'[x]\,\tilde{h} = g - F(x)$ where $F'[x]$ is singular. Let us suppose that it turns out that F' is a positive semidefinite matrix (this may come from the maximum principle in the underlying pde) and is also symmetric (just to simplify the idea and avoid transposes). Then we might regularise by

$$(F'[x] + \alpha R)\,\tilde{h} = g - F(x)$$

where R is a positive (semi-definite) matrix and $\alpha > 0$.

Some Open Issues

We are solving the Newton equation $F'[x]\,\tilde{h} = g - F(x)$ where $F'[x]$ is singular. Let us suppose that it turns out that F' is a positive semidefinite matrix (this may come from the maximum principle in the underlying pde) and is also symmetric (just to simplify the idea and avoid transposes). Then we might regularise by

$$(F'[x] + \alpha R)\,\tilde{h} = g - F(x)$$

where R is a positive (semi-definite) matrix and $\alpha > 0$.

- What if (again from the underlying pde) it could be shown that the Hessian $F''[x](\,\cdot\,, h)$ was a positive semidefinite matrix? Would then the choice $R = \frac{1}{2}F''[x](\,\cdot\,, h)$ be a regularisation?

  Even if it only partly did so, it might allow a reduction in the level of regularisation required by another scheme.

Some Open Issues

We are solving the Newton equation $F'[x]\,\tilde{h} = g - F(x)$ where $F'[x]$ is singular. Let us suppose that it turns out that F' is a positive semidefinite matrix (this may come from the maximum principle in the underlying pde) and is also symmetric (just to simplify the idea and avoid transposes). Then we might regularise by

$$(F'[x] + \alpha R)\,\tilde{h} = g - F(x)$$

where R is a positive (semi-definite) matrix and $\alpha > 0$.

- What if (again from the underlying pde) it could be shown that the Hessian $F''[x](\,\cdot\,, h)$ was a positive semidefinite matrix? Would then the choice $R = \frac{1}{2}F''[x](\,\cdot\,, h)$ be a regularisation?

  Even if it only partly did so, it might allow a reduction in the level of regularisation required by another scheme.

- Can we effectively use a higher (than two) order MacLaurin expansion in an attempt to better model the nonlinear map? For most problems we think the answer is no. However, we were once convinced that a single derivative was the sensible limit.

Is Halley's method a Panacea?

We have painted a very rosy picture of the power of Halley's method. To be fair we should point out difficulties.

Is Halley's method a Panacea?

We have painted a very rosy picture of the power of Halley's method. To be fair we should point out difficulties.

- There are several important inverse problems where we are unable (or more precisely, we don't know how) to compute derivatives of the map F using the construction from F itself.

Is Halley's method a Panacea?

We have painted a very rosy picture of the power of Halley's method. To be fair we should point out difficulties.

- There are several important inverse problems where we are unable (or more precisely, we don't know how) to compute derivatives of the map F using the construction from F itself.

- Regularisation issues are more complex. With any method there is the problem of choosing the optimal value of the regularisation parameter. With the second degree method we have to select two regularisation parameters, one each for the predictor and the corrector; experience shows that best results are obtained when these are not the same.

Is Halley's method a Panacea?

We have painted a very rosy picture of the power of Halley's method. To be fair we should point out difficulties.

- There are several important inverse problems where we are unable (or more precisely, we don't know how) to compute derivatives of the map F using the construction from F itself.

- Regularisation issues are more complex. With any method there is the problem of choosing the optimal value of the regularisation parameter. With the second degree method we have to select two regularisation parameters, one each for the predictor and the corrector; experience shows that best results are obtained when these are not the same.

- While the scattering example shows that Halley's method can provide superior as well as faster reconstructions, there are problems (even those involving the detection of obstacles) for which this seems not to be the case. This may be due to difficulties in selecting an optimal level of regularisation or it may be inherent in the problem.