

Free surface flow over topography: an inverse approach

Connor Robbins

April, 2023

*Thesis submitted for the degree of
Doctor of Philosophy
in
Applied Mathematics
at
The University of East Anglia
School of Mathematics
and
The University of Adelaide
Faculty of Sciences, Engineering and Technology
School of Computer and Mathematical Sciences*



THE UNIVERSITY
of ADELAIDE

UEA
University of East Anglia

This degree of Doctor of Philosophy was undertaken under a cotutelle agreement between the University of East Anglia and the University of Adelaide, and subject to successful examination will result in a degree award from both Universities (double degree). This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

The problem of steady two-dimensional open channel free surface flow over topography is studied. The focus here is on the inverse problem of recovering the profile of the topography given prescribed surface data. This thesis explains the ill-posed nature of the inverse problem and develops a method based on the truncated singular value decomposition to obtain regularised inverse solutions for the topography given free surface data. It is shown how discretisation of the inverse problem yields a linear system to be solved and how regularisation can be applied to temper the ill-posed nature of the problem such that useful solutions can be obtained. This method is much less computationally expensive than previous approaches using the Newton method which enables a rapid exploration of the solution space. The developed method is trialled against input data from computed solutions to the forward problem to give a benchmark against which the performance of the model can be assessed, and it is found that it is able to accurately reconstruct the topography. We then show that the method can recover the topography even with substantial noise added to the surface. Finally we use the model to explore the solution space of the inverse problem.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Contents

Abstract	3
List of Figures	9
List of Tables	25
Acknowledgements	26
1 Introduction	27
2 Background	33
2.1 Notation	33
2.2 Fluid Mechanics	35
2.2.1 Potential flow model	35
2.2.2 The forced Korteweg-De Vries equations (fKDV)	38
2.2.3 Derivation of the governing equations for the boundary integral method	38
2.3 Linear algebra	46

2.3.1	Solutions to linear equations	46
2.4	Newton Method	52
2.4.1	One equation in one variable	52
2.4.2	Systems of equations	53
2.5	Inverse problems	54
2.5.1	Matrix equations and examples	54
2.5.2	Fredholm equations of the first kind and examples	55
3	Forward and Inverse Methods	68
3.1	Discretisation of Boundary Integral equations	68
3.2	Forward Problem Methods	70
3.2.1	Solutions as a perturbation to the uniform stream	72
3.2.2	Solutions as a perturbation to the solitary wave	73
3.2.3	Hydraulic falls and jumps	73
3.2.4	Trapped wave solutions	76
3.2.5	Forward solutions in critical flow	78
3.3	Inverse Problem Methods	85
3.3.1	Inverse Pressure Problem	86
3.3.2	Inverse topography problem as a linear problem	87
3.4	Unforced Forward and Inverse Problems	95
3.4.1	Unforced forward problem	95

3.4.2	Unforced inverse problem	97
4	Results: Inversely Retrieving Topography from Computed Forward Solutions	105
4.1	Topography consisting of one Gaussian	105
4.1.1	Supercritical flow as a perturbation of the uniform stream .	106
4.1.2	Supercritical flow as a perturbation of the solitary wave . .	107
4.1.3	Unforced solitary wave	112
4.1.4	Hydraulic falls	112
4.1.5	Subcritical flow with a train of waves	114
4.2	Topography consisting of two Gaussians	116
4.2.1	Supercritical flow	116
4.2.2	Subcritical flow with a train of waves	116
4.2.3	Subcritical flow with trapped waves	117
4.3	Discussion	120
5	Results: Inversely Retrieving Topography from Perturbed Computed Forward Solutions	123
5.1	Inverse results for a single surface measurement	125
5.2	Inverse results for multiple surface measurements	128
5.3	Discussion	134
6	Inverse Results for Prescribed Surfaces	136

6.1	Inverse results for a Gaussian free surface	137
6.1.1	Response of the topography to changing Froude	137
6.2	Inverse results for artificial trapped waves	143
6.2.1	KdV	143
7	Conclusion	152
	Appendices	155
A	Norm of inverse fKdV forcing for a Gaussian free surface	155
B	Repeated integration by parts of a Gaussian	160

List of Figures

1.0.1 As the flow moves from left-to-right over uneven topography y_b or past a pressure distribution P the shape of the free surface y_f will be affected. Seeking solutions for one of these quantities, with the other two known simultaneously, defines a family of problems. When the free surface is one of the known quantities we are considering an inverse problem. The focus of this thesis is on the inverse problem for the topography. 28

2.2.1 As the flow moves from left-to-right over uneven topography y_b or past a pressure distribution P the shape of the free surface y_f will be affected. Seeking solutions for one of these quantities, with the other two known simultaneously, defines a family of problems. When the free surface is one of the known quantities we are considering an inverse problem. 39

2.2.2	How the applied complex mappings transform the fluid domain, bounded by the free surface (blue) and the topography (red). Examples of points in the fluid domain are marked and tracked through the mapping: A, B and C are points on the free surface with A representing a point far upstream and C one far downstream; D, E and F are points on the topography analogous to those on the free surface; P represents a point inside flow far upstream, whereas Q and R are points far downstream. (a) The fluid domain in the (ϕ, ψ) plane. (b) The fluid domain in the (α, β) plane.	40
2.2.3	A diagram of the contour Γ and how this contour is then subdivided in order to compute the contour integral around Γ . . .	42
2.5.1	A plot of the singular values of the matrix corresponding to the discretised kernel $K(x, y) = \sin(x - y)$ with $L = \pi, N = 25$	58
2.5.2	$L = \pi, N = 25$ A comparison of the true solution (red dashed) to (2.5.2) to the truncated solutions \underline{f}_κ for truncation ranks $\kappa = 2$ (blue) and $\kappa = 3$ (black).	58
2.5.3	$L = \pi, N = 25$ A comparison of the true solution (red dashed) to (2.5.2) to the truncated solution \underline{f}_κ to the perturbed system for truncation rank $\kappa = 2$ (blue). Top: Standard deviation of added noise $\epsilon = 2 \times 10^{-4}$. Bottom: Standard deviation of added noise $\epsilon = 2$	59
2.5.4	(a) A plot of the singular values of the matrix corresponding to the discretised kernel $K(x, y) = \frac{1}{2} \left(1 + \tanh \left(\frac{\pi(x-y)}{2} \right) \right)$ with $L = \pi, N = 25$. (b) Inverse solutions for the truncation rank $\kappa = 7$ (blue) and $\kappa = 25$ (black) compared to the true solution (red) . . .	60

- 2.5.5 $L = \pi, N = 81$ (a) A plot of the singular values of the matrix corresponding to the discretised kernel $K(x, y) = \frac{1}{2} \left(1 + \tanh \left(\frac{\pi(x-y)}{2} \right) \right)$ with $L = \pi, N = 81$. (b) Inverse solutions for the truncation rank $\kappa = 7$ (blue) and $\kappa = 25$ (black) compared to the true solution (red). 61
- 2.5.6 An example of an idealised L-curve. Selection of κ corresponding to a solution close to the corner on the horizontal leg of the L-curve attempts to obtain a smooth regularised solution \underline{f}_κ while also ensuring that this regularised solution does actually perform well as an approximate solution to (2.5.3). 63
- 2.5.7 $L = \pi, N = 81$ (a) Picard plot to help consider the satisfaction of the DPC (b) The L-curve for the tanh kernel, the red cross indicates the solution with $\kappa = 35$. (c) A plot of the norm of the solution to the inverse problem against the truncation rank κ . (d) The inverse solution with $\kappa = 35$ (blue) compared to the true solution (red). 64
- 2.5.8 $L = \pi, N = 81, \epsilon = 1 \times 10^{-8}$ (a) Picard plot for the noisy problem showing that the DPC is not satisfied for $i > 27$. (b) The L-curve with the red cross indicating the solution with $\kappa = 27$. (c) A plot of the norm of the inverse solution against the truncation rank κ . (d) An inverse solution with a truncation rank chosen to satisfy the DPC $\kappa = 27$ (blue) and one with the truncation rank chosen too high $\kappa = 40$ (black) compared to the true solution (red). 66
- 2.5.9 $L = \pi, N = 81, \epsilon = 1 \times 10^{-2}$ (a) Picard plot for the noisy problem showing that the DPC is satisfied for $i < 8$. (b) The L-curve for this problem with the red cross indicating the solution with $\kappa = 8$. (c) A plot of the norm of the inverse solution against the truncation rank κ . (d) Inverse solutions with $\kappa = 8$ (blue) and $\kappa = 12$ (black) compared to the true solution (red). 67

- 3.2.1 $L = 20, N = 641, F = 1.2, b = 0.3$. Uniform stream perturbation solutions to the forward problem for supercritical flow. These solutions $\eta = y_f - 1$ follow approximately the shape of the topography y_T . The solid lines are plotted as a function of the potential ϕ and the dashed lines are plotted as a function of the physical variable x . (a) Gaussian topography of form (3.2.5) with $a = -0.1$. (b) Decaying cosine topography of form (3.2.6) with $a = -0.03$ 72
- 3.2.2 $L = 20, N = 641, F = 1.2, b = 0.3$. Uniform stream perturbation solutions to the forward problem for subcritical flow. The solid lines are plotted as a function of the potential ϕ and the dashed lines are plotted as a function of the physical variable x . (a) Gaussian topography of form (3.2.5) with $a = -0.1$. (b) Decaying cosine topography of form (3.2.6) with $a = -0.03$ 73
- 3.2.3 Bifurcation curve for forward solutions to forced flow showing the solitary wave and uniform stream branches for multiple amplitudes of forcing with $L = 20, N = 641, b = 1$ 74
- 3.2.4 $L = 20, N = 641, F = 1.1, b = 1$. Nonuniqueness in the forward problem over a topography of the form (3.2.5) results in free surface solutions that are like a perturbation to the uniform stream (orange) or a perturbation to the solitary wave (black). Solutions are plotted both as $y_f(\phi)$ (solid line) and as $y_f(x)$ (dashed line). (a) Unforced flow with $a = 0$. (b) Forced flow with $a = 2 \times 10^{-2}$ 74

- 3.2.5 $L = 20, N = 721, b = 1$. Examples of hydraulic fall solutions to flow past a Gaussian topography of the form (3.2.5). The solid lines are plotted as a function of the potential ϕ and the dashed lines are plotted as a function of the physical variable x . (a) $a = 0.2$. The Froude number was calculated as $F \approx 1.4326$. (b) $a = 0.4$. The Froude number was calculated as $F \approx 1.4326$ 75
- 3.2.6 $L = 30, N = 641, F = 0.8, b = 1, a = -0.04$. Examples of trapped wave free surface solutions (blue) in subcritical flow over a two bumped topography (red) given by (3.2.7) where the separation comes as part of the solution. The solid lines are plotted as a function of the potential ϕ and the dashed lines are plotted as a function of the physical variable x . Different profiles are obtained for different initial guesses for c . (a) Solution with two wave peaks with $c \approx 3.1334$. (b) Solution with three wave peaks with $c \approx 5.4871$. (c) Solution with seven wave peaks with $c \approx 14.9021$. (d) Solution with eight wave peaks with $c \approx 17.2559$ 77
- 3.2.7 (a) Solutions for the free surface for different amplitudes of forcing. (b) Profiles of the forcing for different amplitudes. (c) Phase portrait for the solutions for u showing how they approach the termination point with algebraic decay for all solutions except that with $\alpha = 1$ 84
- 3.3.1 Forward and inverse numerical solutions obtained by Newton's method, $F = 1.2$. (a) Solutions to the forward problem for the prescribed topography y_T given (3.3.1). (b) Prescribed topography, y_T , compared to Tam, Yu, Kelso, and Binder (2015) inverse Newton's method, y_b 86

- 3.3.2 Dependence of the condition number of \mathbf{M} on N when $L = 10$.
 Top: \mathbf{M} with the final row set by the boundary condition $\theta_N = 0$. Bottom: \mathbf{M} with the final row set instead by the boundary condition $\theta_1 = 0$ 90
- 3.3.3 The eigenfunction of \mathbf{M} corresponding to the eigenvalue with magnitude 9.2×10^{-7} , when $L = 10$ and $N = 101$ 92
- 3.3.4 Results for the topography (3.3.1) at $F = 1.2$ when $L = 20$ and $N = 363$. (a) $\|\underline{\theta}_\kappa\|$ rises rapidly to a stable value for a range of the truncation parameter κ before beginning to increase further as the model fails to output smooth solutions. (b) Profiles of $\underline{\theta}_\kappa$ for various κ , showing how the output varies by selecting κ to correspond to the different sections of the $(\kappa, \|\underline{\theta}_\kappa\|)$ curve. 93
- 3.3.5 The case of topography (3.3.1) at $F = 1.2$ when $L = 20$ and $N = 101$. (a) The originally prescribed topography compared to those found by applying the TSVD method to the results of the forward Newton problem at two resolutions ($\kappa = 101; N = 101, 721$). (b) The norm of the error $\|y_b - y_T\|$, where y_T is the prescribed topography (3.3.1), and y_b is the TSVD solution. 95
- 3.4.1 A diagram of how the curve Γ is subdivided for reference in computing the contour integral around Γ . This contour contains one pole at $z = z_0 = \frac{i\pi}{2}$ 100

- 4.1.1 Inverse results for a perturbation to the uniform stream solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.02, b = 1, L = 20, N = 641, F = 1.1$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 89$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 89$ 107
- 4.1.2 Inverse results for a perturbation to the uniform stream solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.2, b = 1, L = 20, N = 641, F = 1.5$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 146$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 146$ 108
- 4.1.3 Inverse results for a perturbation to the solitary wave solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.02, b = 1, L = 20, N = 641, F = 1.1$ (These are the same parameters as for the uniform stream solution of Figure 4.1.1). (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 73$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 73$. . . 109

- 4.1.4 Inverse results for a perturbation to the solitary wave solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.02, b = 1, L = 20, N = 1241, F = 1.3$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 41$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 41$ 110
- 4.1.5 Inverse results for an unforced solitary wave solution with the parameter values $a = 0, L = 20, N = 841, F = 1.1$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 41$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 103$. (e) The maximum displacement of the topography from the zero level decays with N^{-2} (red) Reference line of gradient -2 (black dotted) (f) The norm of the right-hand side of the matrix equation approaches zero like N^{-2} (blue) and a reference line of gradient -2 (black dashed). 111
- 4.1.6 Inverse results for a hydraulic fall solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.1, b = 1, L = 20, N = 721$. The Froude number is calculated as part of the solution and was found to be $F \approx 1.3005$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 115$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 115$ 113

- 4.1.7 Inverse results for a hydraulic fall solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.4, b = 1, L = 20, N = 721$. The Froude number is calculated as part of the solution and was found to be $F \approx 1.6171$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 123$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 123$ 114
- 4.1.8 Inverse results for a subcritical solution with a train of waves over a Gaussian topography (4.1.1) with the parameter values $a = -0.05, b = 1, L = 21, N = 561, F = 0.8$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 25$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 25$ 115
- 4.2.1 Inverse results for a supercritical solution with a train of waves over a Gaussian topography (4.2.1) with the parameter values $a = 0.05, b = 1, L = 20, N = 541, F = 1.2$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 158$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 158$ 117

- 4.2.2 Inverse results for a subcritical solution with a train of waves over a Gaussian topography (4.2.1) with the parameter values $a = -0.04, b = 1, L = 20, N = 641, F = 0.8$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 13$ (black) and $\kappa = 40$ (blue) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 13$. 118
- 4.2.3 Inverse results for a subcritical solution with trapped waves over a Gaussian topography (4.2.1) with the parameter values $a = -0.04, b = 1, L = 30, N = 1241, F = 0.8, c \approx 3.116$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 87$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 87$. (e) Profiles with increasing N showing that as N is increased the inverse solution gets closer to the true solution. 119
- 4.2.4 Inverse results for a subcritical solution with trapped waves over a Gaussian topography (4.2.1) with the parameter values $a = -0.04, b = 1, L = 30, N = 1241, F = 0.8, c \approx 10.147$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 89$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 87$. (e) Profiles with increasing N showing that as N is increased the inverse solution gets closer to the true solution. 121

5.0.1 (a) Free surface solution for a topography of the form (5.0.1) with $a = 0.2, b = 1, F = 1.5, N = 641, L = 20$ (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-5}$ added (black).
 (b) TSVD solution for y_b in unperturbed problem with $\kappa = 146$ (blue), TSVD solution for y_b in perturbed problem with $\kappa = 146$ (black). 124

5.1.1 Inverse results for a supercritical perturbation to the uniform stream type free surface solution over a topography of the form (5.0.1) with $a = 0.2, b = 1, F = 1.5, N = 641, L = 20$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-5}$ added (black).
 (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-5}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 69$ 125

5.1.2 Inverse results for a supercritical perturbation to the uniform stream type solution over a topography of the form (5.0.1) with $a = 0.2, b = 1, F = 1.5, N = 641, L = 20$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-3}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-3}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 42$ 126

- 5.1.3 Inverse results for a hydraulic fall solution for a topography of the form (5.0.1) with $a = 0.4, b = 1, N = 721, L = 20, F \approx 1.617$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-3}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-3}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 50$ 128
- 5.1.4 Results for a supercritical perturbation to the uniform stream type solution for a topography of the form (5.0.1) with $a = 0.2, b = 1, F = 1.5, N = 641, L = 20$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-2}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-2}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 40$ 129
- 5.1.5 Results for a hydraulic fall solution for a topography of the form (5.0.1) with $a = 0.4, b = 1, N = 721, L = 20, F \approx 1.617$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-2}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-2}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 40$. 130

- 5.2.1 Inverse results for a perturbation to the uniform stream solution over a Gaussian topography (5.0.1) based on 1000 simulated measurements with the parameter values $a = 0.2, b = 1, L = 20, N = 641, F = 1.5, \kappa = 40, \epsilon = 10^{-2}$. Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue). 131
- 5.2.2 Inverse results for a perturbation to the solitary wave solution over a Gaussian topography (5.0.1) based on 1000 simulated measurements with the parameter values $a = 0.02, b = 1, L = 20, N = 641, F = 1.1, \kappa = 40, \epsilon = 10^{-2}$. Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue). 132
- 5.2.3 Inverse results for a perturbation to the hydraulic fall solution over a Gaussian topography (5.0.1) based on 1000 simulated measurements with the parameter values $a = 0.4, b = 1, L = 20, N = 721, F \approx 1.617, \kappa = 40, \epsilon = 10^{-2}$. Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue). 133

- 5.2.4 Inverse results for a perturbation to a trapped wave solution between two Gaussian topographical features (5.0.2) based on 1000 simulated measurements with the parameter values $a = -0.04, b = 1, L = 30, N = 641, F = 0.8, \kappa = 40, \epsilon = 10^{-2}$. Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue). 134
- 5.2.5 Inverse results for a perturbed uniform free stream with $F = 1.2, L = 20, N = 641, \kappa = 40, \epsilon = 10^{-2}$ (a) The true topography (red) and the TSVD result for a single realisation of the perturbed problem. (b) Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue). . 135
- 6.1.1 Inverse results for the topography as the Froude number is varied with parameters $\beta = 0.3, L = 20, N = 641$ (a) $\alpha = 0.1, \kappa = 92$. (b) $\alpha = -0.1, \kappa = 92$. (c) $\alpha = 0.15, \kappa = 104$. (d) $\alpha = -0.15, \kappa = 104$. (e) $\alpha = 0.2, \kappa = 110$. (f) $\alpha = -0.2, \kappa = 96$ 139
- 6.1.2 Inverse solutions for a fixed free surface as the Froude number is varied with $\beta = 0.3, L = 20, N = 641, \kappa = 92$ (a) The norm of the inverse solution against F . (b) The central value $y_b(0)$ of the inverse solution against F 140

6.1.3 Inverse solutions for a fixed free surface as the Froude number is varied with $\beta = 0.3, L = 20, N = 641, \kappa = 92$ (a) The norm of the inverse solution against F (b) The central value $y_b(0)$ of the inverse solution against F 141

6.1.4 Plots of the central response of the inverse solution for a fixed Froude number as the amplitude of the prescribed surface is varied. $\beta = 0.3, L = 20, N = 641, \kappa = 58$ (a) $F = 1.1 > 1 - \frac{\beta^2}{3}$ (b) $F = 0.9 < 1 - \frac{\beta^2}{3}$ (c) $F = 0.97 = 1 - \frac{\beta^2}{3}$ 142

6.1.5 Plots of the norm of the inverse solution for a fixed Froude number as the amplitude of the prescribed surface is varied. $\beta = 0.3, L = 20, N = 641, \kappa = 58$ (a) $F = 1.1$ (b) $F = 0.9$ (c) $F = 0.97$ 144

6.2.1 Schematic of the c_i curves against frequency ν for different regimes of flow. The solid lines represent the possible forms of the c_1 curves, depending on the Froude number, and the broken lines represent c_2 for different arbitrary values of A , plotted only to show how the value of A can alter the number of possible intersections between the c_1 and c_2 curves and the values of ν for which one mode will be dominant over the other. 146

6.2.2 The energies carried in the Fourier modes of the inverse solution for the fully nonlinear (solid) and fKdV (dashed) models with $F = 1.2, N = 1624$. No intersection occurs in supercritical flow when $\mu > \frac{3|A|}{8}$. (a) $A = 0.05$ (b) $A = 0.1$ 147

6.2.3 The energies carried in the Fourier modes of the inverse solution for the fully nonlinear (solid) and fKdV (dashed) models with $A = 0.2, N = 1624, F = 1.05$. An intersection occurs in supercritical flow when $\mu \leq \frac{3|A|}{8}$ 148

- 6.2.4 Comparison of the energy in the Fourier modes between nonlinear (solid) and fkdv (dashed) for critical flow with $A = 0.1, N = 1624$. One intersection occurs between the two energies. 149
- 6.2.5 Comparison of the energy in the Fourier modes between nonlinear (solid) and fkdv (dashed) for subcritical flow with $A = 0.02, N = 1624, F = 0.9$ Two intersections occur in subcritical flow and the energy in the c_1 mode appears to go to zero for a particular value of ν 150

List of Tables

2.1	A comparison of the direct solutions to a perturbed matrix equation and the regularised solutions computed by way truncated singular value decomposition (TSVD).	55
-----	--	----

Acknowledgements

I would like to thank my supervisors Mark and Ben for their guidance, support and unerring patience.

Thank you to everyone who has helped me enjoy myself along the way, from friends on academic side with their help in solving tricky maths problems that ended up being a complete distraction to the friends I made in Australia who spent countless hours just wandering around with me while I was entranced by the wildlife.

A special acknowledgement to my mum and sister for putting up with me moving in with them three times throughout the course of my PhD and for actually remembering what it is I study this time around.

Finally, to Silvana, I could never put into words what the help and care you've given me means.

Introduction

The subject of this thesis is the study of the inverse problem of water waves over topography. The inverse problem is the problem of calculating the topography when the surface profile of the fluid is known, contrasting the forward problem which is to calculate the free surface given knowledge of the underlying topography. The motion of fluids is modelled by the Navier Stokes equations. We shall focus on two dimensional steady flow under gravity of an incompressible, irrotational and inviscid fluid over a localised topographical forcing such that we insist that the topography is level outside of a finite interval. Steady means here that, while the fluid itself is still moving, the surface profile remains fixed in time. Under these assumptions the governing equations are simplified to the steady Euler equations.

An asymptotic expansion of the water wave equations leads to the Korteweg-de Vries (KdV) equation, a weakly nonlinear model based on the assumptions that the surface waves are of a sufficiently long wave length relative to their height. The derivation of this equation was later modified by Akylas, 1984 to include a term that takes into account incident surface pressure that forces the flow, known as the forced KdV (fKdV) equation. Later work showed that the same equation could be derived for fluid forced instead by the topography or even by a combination of the two types of forcing (Cole, 1985; Grimshaw and Smyth, 1986; Shen, 1995; Wu, 1987). Binder, 2019 identified 11 different basic flow types for the fKdV over a two parameter solution space, these parameters were the

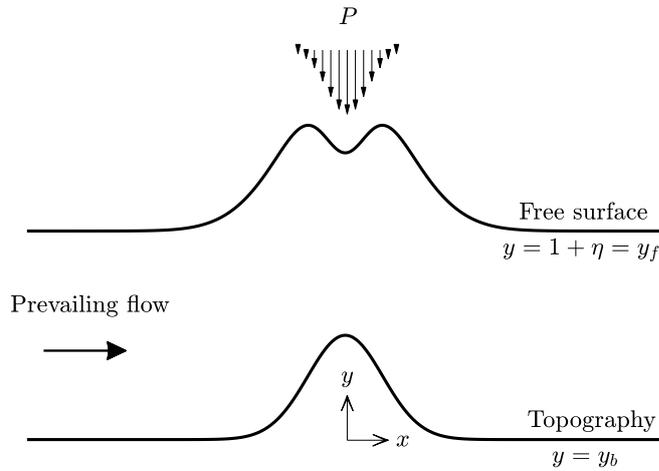


Figure 1.0.1: As the flow moves from left-to-right over uneven topography y_b or past a pressure distribution P the shape of the free surface y_f will be affected. Seeking solutions for one of these quantities, with the other two known simultaneously, defines a family of problems. When the free surface is one of the known quantities we are considering an inverse problem. The focus of this thesis is on the inverse problem for the topography.

amplitude of the forcing and the Froude number. The Froude number, F , is the ratio of the typical fluid speed to the propagation speed of linear gravity waves, $F = \frac{U}{\sqrt{gH}}$, where U is the typical flow speed, g the acceleration due to gravity and H the typical flow depth. The KdV permits nonuniqueness in the forward problem by way of the well known solitary wave when $F > 1$ and this extends to the fKdV for which there are solutions that can be considered as perturbations to the uniform stream and those that can be considered as perturbations to the solitary wave occurring for the same Froude number and amplitude of forcing. The fKdV acts as a useful guide in both the forward and inverse problems to inform exploration of the fully nonlinear solution space. The solution to the inverse problem using the fKdV can be written as an exact formula.

In looking to calculate fully nonlinear forward solutions over a semi-circular topography Forbes and Schwartz, 1982 cast the problem as a boundary integral equation via complex mappings, a technique used by Vanden-Broeck, 1987 to

calculate perturbation to solitary wave type solutions in the fully nonlinear problem. This method has been extended to consider topographies more complicated than a semi-circle, see for example the work of Binder, Dias, and Vanden-Broeck (2008), Binder, Vanden-Broeck, and Dias (2005), Keeler, Binder, and Blyth (2018), and Lustri, McCue, and Binder (2012), and limiting configurations of the free surface (Hunter and Vanden-Broeck, 1983; Vanden-Broeck, 1997; Wade, Binder, Mattner, and Denier, 2017), where waves approach what is known as the Stoke's limiting configuration at the point beyond which the wave would break if it were any steeper (Stokes, 1880). The boundary integral method can also be used for purely pressure forced flows (Binder and Vanden-Broeck, 2007) or for combined forcing types (Binder and Vanden-Broeck, 2011).

The boundary integral method was used by Binder, Blyth, and McCue, 2013 to compute forward and inverse solutions for problems forced by either a pressure distribution or non-uniform topography over a finite support. These results were compared to the predictions of the fKdV and it was found that, while accurate near $F = 1$, not only the fKdV does not predict so well the results of the fully nonlinear problem when F is not near one but that the topography and pressure inverse problems deviate more greatly from one another than predicted by the fKdV, especially for the Froude number not close to one. This highlights the need for a robust fully nonlinear methodology to be found in order to better study the fully nonlinear solution space. The authors found that, while at low resolution they could output good inverse results, when the resolution of the model was increased the model would not converge. The authors worked around this by computing inverse solutions at low resolution before interpolating them to higher resolution to be used as input into the forward problem, this way they ensured that the prescribed free surface could be recovered in the forward sense from their inverse solutions.

This same sensitivity to the model resolution was reported by Tam, Yu, Kelso, and Binder, 2015. The authors here established channel flows over Gaussian

and semi-circular topographies of known form producing free surface featuring a hydraulic fall and then tested the boundary integral method in both the forward and inverse sense against the experimental free surface data collected. The authors found good agreement in the forward problem and also, when the model resolution was taken low enough to not be problematic, in the inverse problem. They noted that the inverse method was less accurate at points of slope discontinuity, for example the boundary of the semi-circular regions. The model was not sensitive to small perturbations in input data at low resolution, it only began to act in an ill-posed sense at higher resolution.

The sensitivity of the inverse model to mesh size is not unique to steady flows and was reported for time dependent periodic flows by Vasan and Deconinck, 2013. The authors here sought solutions in the form of truncated Fourier series and solved the resultant system by way of the Levenberg-Marquardt algorithm, which is a least squares solver. The authors state the ill-posed nature of the inverse problem and relate this to the appearance of hyperbolic functions in their formulation and subsequent issues with machine precision. They avoided this by truncating their Fourier modes to avoid high wave numbers and allow for smooth solutions to be converged upon, however these solutions struggle to capture finer detail in the topography at the lowered resolution.

Keller, 1976 gives an introduction to inverse problems, with some simple examples highlighting issues like non-uniqueness that may come about from trying to solve them. Kabanikhin, 2008 addresses the subjects of inverse problems and ill-posed problems more generally and rigorously. In this paper we find a statement of the three conditions required for "Hadamard well-posedness" of a problem, these are that the system: has a solution; the solution is unique; small changes to input data result in only small changes to the output solution (stability). For problems that aren't well-posed the author details regularisation methods that seek approximate solutions to these systems. One such method is the Levenberg-Marquardt algorithm previously mentioned, details of which can be found in Croeze, Pittman, and Reynolds, 2012. While we do not employ this algorithm

in the present work it was useful to gain a broader understanding of the need for regularisation methods and what they achieve.

In Chapter 2 of this thesis we lay out the required background for understanding the boundary integral method and some general theory of ill-posed problems. In Chapter 3 we discretise the boundary integral equations and show how the inverse problem can be viewed as solving a Fredholm integral equation of the first kind, a type of problem known to be ill-posed (Hansen, 1990b, 1992b). The method of regularisation we choose to employ is truncated singular value decomposition (TSVD) as the discretisation of the boundary integral equations lead to an ill-posed system of linear equations, naturally represented in matrix form. This method is explained and examples given in Chapter 2 before it is applied to the full problem in Chapter 3. It is found to be much faster than the Newton method and with appropriate regularisation can handle resolutions for which the inverse Newton method would not converge.

Having developed our tools, in Chapter 4 we test the methodology by prescribing a topography, consisting of either one or two Gaussians, calculating solutions to the forward problem and then using these solutions as input into the inverse problem in order to recover the originally prescribed topography. We do this for different types of forward solution in both supercritical and subcritical flows and find that good regularised results can be obtained for flows that decay in the far-field.

In Chapter 5 we test the limits of the regularisation of our model by adding large amounts of white noise to the forward solutions before attempting to again recover the topography used as input into the forward problem. We show how our solutions to noisy problems could be improved by repeating ‘experiments’ and averaging the results by simulating multiple sets of noisy data for the same surface.

Finally in Chapter 6 we prescribe free surfaces directly, using the developed inverse method to explore the solution space of the inverse fully nonlinear

problem, making use of the speed of the method to explore the response of the topography to particular features being prescribed on the surface, like trapped waves, while parameters like the wave length of the wave are varied.

2

Background

2.1 Notation

Cauchy Principle Value Integral

Throughout this document we make use of the Cauchy principle value integral and distinguish it from the usual integral by use of the dashed integral symbol $\rlap{-}\int$. For a function $f(x)$, containing a singularity at $x = b$, the principle value integral of the function over the interval $[a, c]$ with $a < b < c$ is defined as

$$\rlap{-}\int_a^c f(x) dx = \lim_{\epsilon \rightarrow 0^+} \left(\int_a^{b-\epsilon} f(x) dx + \int_{b+\epsilon}^c f(x) dx \right).$$

Norms

We use the notation $\| \cdot \|$ to denote the L2-norm when applied to continuous functions. For the function $f(x)$ this is defined to be

$$\|f\| = \sqrt{\int_{-\infty}^{\infty} (f(x))^2 dx}.$$

Most of the vectors contained in this document will represent discrete approximations of continuous functions over an interval $[a, c]$. Let the vector $\underline{\mathbf{f}}$ be such a vector, dividing the interval $[a, c]$ into N equally spaced mesh-points,

the vector $\underline{\mathbf{f}}$ approximates the scalar function f by having each of its N elements correspond to the value of function f evaluated at that mesh-point. We shall reserve the notation $\|\cdot\|$ when applied to vectors to mean

$$\|\underline{\mathbf{f}}\| = \sqrt{\frac{(c-a)}{N}(\underline{\mathbf{f}} \cdot \underline{\mathbf{f}})} \quad (2.1.1)$$

such that this vector norm approximates the L2-norm of the underlying continuous function enabling better comparison between continuous and discrete results. The absolute value notation $|\underline{\mathbf{f}}|$ will be used to refer to the usual vector L2-norm, given by

$$|\underline{\mathbf{f}}| = \sqrt{(\underline{\mathbf{f}} \cdot \underline{\mathbf{f}})}.$$

The Frobenius norm of the $(m \times n)$ matrix \mathbf{M} with entries $m_{i,j}$ is defined as

$$\|\mathbf{M}\|_F = \sqrt{\sum_i^m \sum_j^n |m_{i,j}|^2}.$$

Fourier Transforms

The Fourier transform of a function $f(x)$, denoted by the addition of a tilde, i.e. $\tilde{f}(w)$, is defined as

$$\tilde{f}(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{iwx} dx. \quad (2.1.2)$$

The inverse Fourier transform is then defined by

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(w) e^{-iwx} dw. \quad (2.1.3)$$

The Dirac Delta function, $\delta(x)$, can be expressed in terms of the Fourier transform of unity as

$$\sqrt{2\pi}\delta(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 1 \cdot e^{iwx} dx. \quad (2.1.4)$$

2.2 Fluid Mechanics

2.2.1 Potential flow model

The motion $\hat{\mathbf{u}}$ of an inviscid and incompressible fluid under the influence of a gravitational field $\underline{\mathbf{g}}$ and of constant density ρ_0 with pressure \hat{p} is modelled by use of the Euler equations and a statement of conservation of mass:

$$\frac{D\hat{\mathbf{u}}}{Dt} = -\frac{1}{\rho_0}\hat{\nabla}\hat{p} + \underline{\mathbf{g}}; \quad (2.2.1)$$

$$\hat{\nabla} \cdot \hat{\mathbf{u}} = 0, \quad (2.2.2)$$

where $\hat{\nabla} = (\frac{\partial}{\partial \hat{x}}, \frac{\partial}{\partial \hat{y}})$. We consider a 2D fluid flow of velocity $\hat{\mathbf{u}} = (\hat{u}, \hat{v})$ in the (\hat{x}, \hat{y}) plane bounded above by the free surface, $\hat{y} = H + \hat{\eta}(\hat{x})$, and below by the impermeable topography, $\hat{y} = \hat{y}_b(\hat{x})$. The gravitational acceleration is $\underline{\mathbf{g}} = (0, -g)$ and a pressure field $\hat{P} = \hat{P}(\hat{x})$ acts on the free surface. We assume that the flow far downstream approaches a horizontal uniform stream of depth H and speed U , that is $\hat{u} \rightarrow U$ and $\hat{P}, \hat{v}, \hat{\eta}, \hat{y}_b \rightarrow 0$ as $\hat{x} \rightarrow \infty$. We restrict our attention to irrotational flows, therefore $\underline{\boldsymbol{\omega}} = \hat{\nabla} \times \hat{\mathbf{u}} = \underline{\mathbf{0}}$, and seek steady solutions such that any time dependence is removed from the problem. Using this information, and the vector identity $(\hat{\mathbf{u}} \cdot \hat{\nabla})\hat{\mathbf{u}} = \frac{1}{2}\hat{\nabla}(\hat{\mathbf{u}} \cdot \hat{\mathbf{u}}) - \hat{\mathbf{u}} \times (\hat{\nabla} \times \hat{\mathbf{u}})$, (2.2.1) may now be rewritten as

$$\hat{\nabla} \left(\frac{1}{2}(\hat{u}^2 + \hat{v}^2) + \frac{\hat{P}}{\rho_0} + g\hat{y} \right) = \underline{\mathbf{0}}. \quad (2.2.3)$$

Integrate this equation then to obtain the Bernoulli condition

$$\frac{1}{2}(\hat{u}^2 + \hat{v}^2) + \frac{\hat{P}}{\rho_0} + g\hat{y} = \text{constant}. \quad (2.2.4)$$

The boundary conditions on the free surface and topography require that flow does not pass through them, i.e. the component of the velocity field normal to the boundary must be zero. Mathematically, the flow must satisfy $\hat{\mathbf{u}} \cdot \underline{\mathbf{n}} = 0$ when evaluated on either boundary, where $\underline{\mathbf{n}}$ is the normal vector to the boundary. For example on the free surface we have $\hat{y} = H + \hat{\eta}(\hat{x})$ and so $\underline{\mathbf{n}} = \hat{\nabla}(\hat{y} - H - \hat{\eta}(\hat{x})) =$

$(-\frac{d\hat{\eta}}{d\hat{x}}, 1)$. The boundary condition on the surface can then be written as

$$\hat{v} = \hat{u} \frac{d\hat{\eta}}{d\hat{x}} \quad \text{on} \quad \hat{y} = H + \hat{\eta}(\hat{x}). \quad (2.2.5)$$

The process of finding the boundary condition on the topography is identical, yielding

$$\hat{v} = \hat{u} \frac{d\hat{y}_b}{d\hat{x}} \quad \text{on} \quad \hat{y} = \hat{y}_b(\hat{x}). \quad (2.2.6)$$

Introduce into (2.2.4) the nondimensional velocities $\underline{\mathbf{u}} = \frac{1}{U}\hat{\mathbf{u}}$, lengths $(x, y, \eta, y_b) = \frac{1}{H}(\hat{x}, \hat{y}, \hat{\eta}, \hat{y}_b)$ and pressure $P = \frac{1}{p_s}\hat{P}$, with the scaling p_s to be chosen shortly, to give

$$\frac{1}{2}U^2(u^2 + v^2) + p_s \frac{P}{\rho_0} + gHy = \text{constant}. \quad (2.2.7)$$

First dividing by U^2 we see that by letting $p_s = gH\rho_0$ we balance the order of the pressure and gravity terms and obtain the nondimensionalised Bernoulli equation in the form

$$\frac{1}{2}(u^2 + v^2) + \frac{1}{F^2}(y + P) = B \quad (2.2.8)$$

where $F = \frac{U}{\sqrt{gH}}$ is the Froude number associated with the flow and B is a constant value. Applying this condition on the far downstream surface we find

$$\frac{1}{2}(1^2 + 0^2) + \frac{1}{F^2}(1 + 0) = B \quad (2.2.9)$$

and therefore at any point in the flow we must satisfy

$$\frac{1}{2}(u^2 + v^2) + \frac{1}{F^2}(y + P) = \frac{1}{2} + \frac{1}{F^2} \quad (2.2.10)$$

Now, as we have $\hat{\nabla} \times \hat{\mathbf{u}} = \mathbf{0}$ we may introduce a nondimensional velocity potential ϕ such that $\underline{\mathbf{u}} = \underline{\nabla}\phi$ and therefore $\underline{\nabla}^2\phi = 0$, where $\underline{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}) = H\hat{\nabla}$. We also introduce a stream function ψ whose relation to our velocity components

and ϕ is given by the Cauchy-Riemann equations

$$\begin{aligned} u &= \frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y}, \\ v &= \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x}. \end{aligned}$$

A consequence of introducing ψ and ϕ is that (2.2.2) is now automatically satisfied. The free surface problem can now be stated in terms of nondimensional variables as

$$\nabla^2 \phi = 0, \quad \text{in} \quad \sigma < y < 1 + \eta \quad (2.2.11)$$

$$-\infty < x < \infty,$$

$$\frac{1}{2}(u^2 + v^2) + \frac{1}{F^2}(y + P) = \frac{1}{2} + \frac{1}{F^2} \quad \text{on} \quad y = 1 + \eta, \quad (2.2.12)$$

$$\frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial x} \frac{d\eta}{dx} \quad \text{on} \quad y = 1 + \eta, \quad (2.2.13)$$

$$\frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial x} \frac{dy_b}{dx} \quad \text{on} \quad y = y_b, \quad (2.2.14)$$

$$u \rightarrow 1 \text{ and } P, v, \eta, y_b \rightarrow 0 \quad \text{as} \quad x \rightarrow \infty. \quad (2.2.15)$$

The formulation here only requires that the free surface decays in the far field in one direction, i.e. as $x \rightarrow \infty$. As we consider flow to be from left to right surfaces that have waves on their upstream portion would represent flows with energy being added to them as the incident waves arrive. Whenever a solution is calculated with waves on its upstream it will subsequently be mirrored so as to have its waves appear instead on the downstream portion of the flow, making use of the fact that potential flows are reversible. This is to ensure that the flow satisfies the Sommerfeld radiation condition (see e.g. Schot, 1992), which requires that there is no additional energy coming from infinity in order for the model to represent a physical flow.

2.2.2 The forced Korteweg-De Vries equations (fKdV)

The forced Korteweg-de Vries (fKdV) equation for a forcing $f(x)$ is given by

$$\eta_{xxx} + 9\eta\eta_x - 6\mu\eta_x = -3f_x, \quad (2.2.16)$$

where $\mu = F - 1$, first derived by Akylas, 1984 for a moving pressure field. The forcing can represent either a pressure forcing, topographical forcing, or some combination of these forcings (Cole, 1985; Shen, 1995; Grimshaw and Smyth, 1986). In the absence of forcing, i.e. $f = 0$, then we obtain from the fKdV equation the usual KdV equation

$$\eta_{xxx} + 9\eta\eta_x - 6\mu\eta_x = 0. \quad (2.2.17)$$

Integrating the fKdV once with respect to x yields

$$\eta_{xx} + \frac{9}{2}\eta^2 - 6\mu\eta = -3f + A,$$

for the constant of integration A ; the conditions $f \rightarrow 0$ and $\eta \rightarrow 0$ as $x \rightarrow \infty$ set $A = 0$. The forcing from the fKdV in the inverse problem is then easily obtained as

$$f(x) = 2\mu\eta - \frac{1}{3}\eta_{xx} - \frac{3}{2}\eta^2. \quad (2.2.18)$$

For example, prescribing surface to take the form of a typical Gaussian as $\eta = ae^{-(bx)^2}$, the forcing can be immediately written as

$$f(x) = ae^{-(bx)^2} \left(2(F - 1) + \frac{2b^2}{3}(1 - 2b^2x^2) - \frac{3a}{2}e^{-(bx)^2} \right). \quad (2.2.19)$$

2.2.3 Derivation of the governing equations for the boundary integral method

The boundary integral method we employ follows closely the work of Vandenberg, 1997 and its subsequent developments, see Binder, Blyth, and McCue,

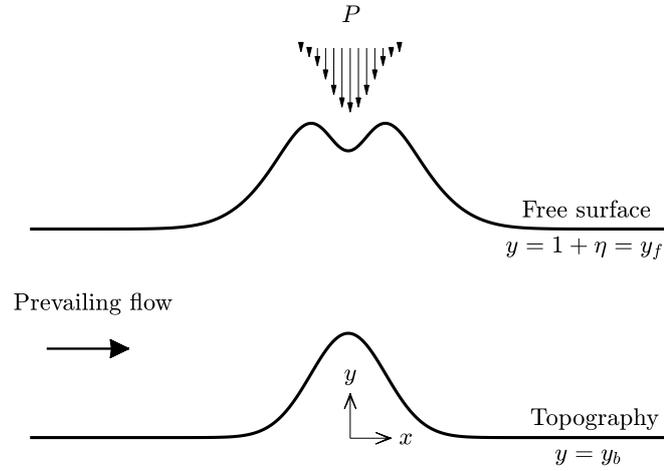


Figure 2.2.1: As the flow moves from left-to-right over uneven topography y_b or past a pressure distribution P the shape of the free surface y_f will be affected. Seeking solutions for one of these quantities, with the other two known simultaneously, defines a family of problems. When the free surface is one of the known quantities we are considering an inverse problem.

2013 and Tam, Yu, Kelso, and Binder, 2015. We begin our derivation from the potential flow model (2.2.11-2.2.15).

First we let $z = x + iy$, and then we introduce the complex potential \mathfrak{f} , an analytic function given by $\mathfrak{f} = \phi + i\psi$, and subsequently the complex velocity $w = \frac{d\mathfrak{f}}{dz}$. Applying the Cauchy-Riemann relations we then find w as

$$w = \frac{d\mathfrak{f}}{dz} = \frac{\partial\phi}{\partial x} + i\frac{\partial\psi}{\partial x} = u - iv,$$

satisfying $|\underline{\mathbf{u}}| = |w|$. The streamlines of the flow are the family of curves given by $\psi = \mathcal{C}$ with each streamline corresponding to taking different values for the constant \mathcal{C} . We let the surface streamline be defined by $\psi = 0$, it then follows the relation $u = \frac{d\psi}{dy}$ and the nondimensionalised unit uniform depth and fluid velocity $\underline{\mathbf{u}} \rightarrow (1, 0)$ in the far field as $x \rightarrow \infty$ that the streamline following the topography will be given by $\psi = -1$. By considering the problem in the (ϕ, ψ) plane the curves describing the free surface and the topography have been

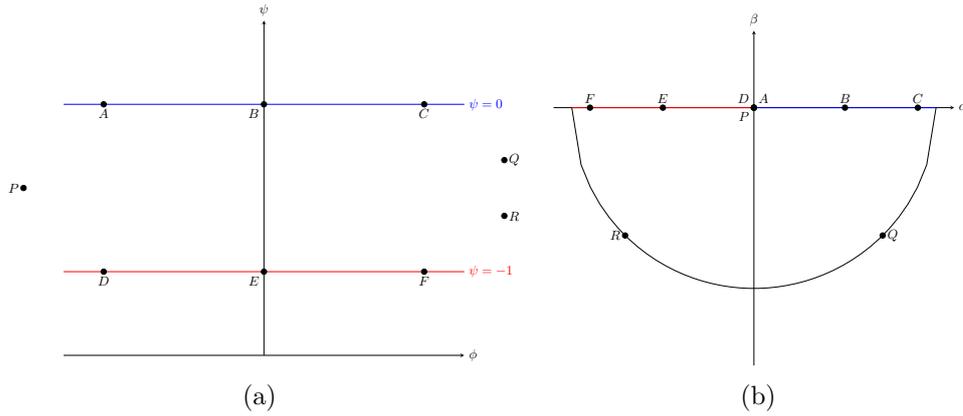


Figure 2.2.2: How the applied complex mappings transform the fluid domain, bounded by the free surface (blue) and the topography (red). Examples of points in the fluid domain are marked and tracked through the mapping: A, B and C are points on the free surface with A representing a point far upstream and C one far downstream; D, E and F are points on the topography analogous to those on the free surface; P represents a point inside flow far upstream, whereas Q and R are points far downstream. (a) The fluid domain in the (ϕ, ψ) plane. (b) The fluid domain in the (α, β) plane.

mapped to parallel straight lines; our domain is now an infinite rectangular strip (Figure 2.2.2(a)).

A further mapping is now introduced, described by $\xi = e^{\pi f} = e^{\pi\phi} e^{i\pi\psi}$ with $\xi = \alpha + i\beta$, transforming the domain from the rectangular strip in the complex f plane to the lower half of the complex ξ plane (Figure 2.2.2(b)). This mapping takes the free surface and maps it to the positive real axis in the ξ plane, whereas the topography is flipped before being mapped to the negative real axis in the ξ plane. Points far downstream inside the flow are mapped to arcs in the lower half plane connecting the topography and free surface; points far upstream are collapsed onto the origin. Further, we introduce the analytic function given by $\tau - i\theta$, with $\tau = \tau(\alpha, \beta)$ and $\theta = \theta(\alpha, \beta)$ such that

$$w = u - iv = e^{\tau - i\theta} = e^{\tau} e^{-i\theta}.$$

These introduced variables can be related back to physical features of the flow, indeed we find that the speed of the flow is given by e^{τ} as $|\underline{u}| = |w| = |e^{\tau - i\theta}| = |e^{\tau}|$

and by equating the real and imaginary parts of $u - iv = e^{\tau - i\theta}$ we find:

$$\begin{aligned} u &= e^{\tau} \cos(\theta); \\ v &= e^{\tau} \sin(\theta), \end{aligned}$$

and as such we can consider θ as the angle between the streamlines and the x axis. In this sense, we are looking to parameterise the fluid flow in terms of the speed and angle of the fluid's motion. A direct result of the above equations is that we may now write

$$u^2 + v^2 = e^{2\tau}, \quad (2.2.20)$$

substitution of which into Bernoulli's equation gives

$$\frac{1}{2}e^{2\tau} + \frac{1}{F^2}(y + P) = \frac{1}{2} + \frac{1}{F^2} \quad \text{on} \quad y = y_f = 1 + \eta. \quad (2.2.21)$$

We will now apply Cauchy's integral theorem to the analytic function $\tau - i\theta$ in the ξ plane, selecting the contour Γ formed by the real axis and the arc in the lower half-plane of a semi-circle of radius R centred at the origin. To avoid a pole on the real axis at $\xi = \alpha_0$ we take a semicircular indentation connecting the points $(\alpha_0 - \epsilon, 0)$ and $(\alpha_0 + \epsilon, 0)$ for a small value ϵ (Figure 2.2.3). As such, with no poles contained within the contour Γ , application of Cauchy's integral theorem results in

$$\oint_{\Gamma} \frac{\tau - i\theta}{\xi - \alpha_0} d\xi = 0.$$

In order to evaluate this integral the contour Γ is split into four subsections (Figure 2.2.3), allowing for the integral to be expressed, with a slight abuse of integral notation, as

$$\left[\int_{\Gamma_1} + \int_{\Gamma_2} + \int_{\Gamma_3} + \int_{\Gamma_4} \right] \left(\frac{\tau(\alpha, \beta) - i\theta(\alpha, \beta)}{\xi - \alpha_0} \right) d\xi = 0. \quad (2.2.22)$$

Both Γ_1 and Γ_3 lie on the real axis ($\xi = \alpha$) and so on these segments we have

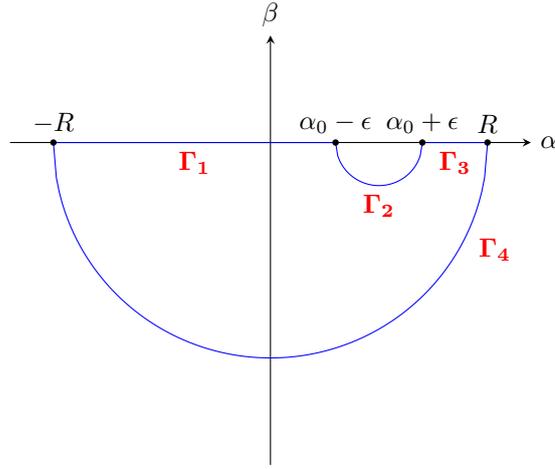


Figure 2.2.3: A diagram of the contour Γ and how this contour is then subdivided in order to compute the contour integral around Γ .

$d\xi = d\alpha$, as such the integrals over these segments may be rewritten as:

$$\int_{\Gamma_1} \frac{\tau(\alpha, \beta) - i\theta(\alpha, \beta)}{\xi - \alpha_0} d\xi = \int_{-R}^{\alpha_0 - \epsilon} \frac{\tau(\alpha, 0) - i\theta(\alpha, 0)}{\alpha - \alpha_0} d\alpha; \quad (2.2.23)$$

$$\int_{\Gamma_3} \frac{\tau(\alpha, \beta) - i\theta(\alpha, \beta)}{\xi - \alpha_0} d\xi = \int_{\alpha_0 + \epsilon}^R \frac{\tau(\alpha, 0) - i\theta(\alpha, 0)}{\alpha - \alpha_0} d\alpha. \quad (2.2.24)$$

Now consider the sum of (2.2.23) and (2.2.24) and take the limit of the sum as $\epsilon \rightarrow 0$ and $R \rightarrow \infty$. Noting the definition of a Cauchy principle value integral then the resulting expression can be written as

$$\lim_{R \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \left\{ \left[\int_{\Gamma_1} + \int_{\Gamma_3} \right] \left(\frac{\tau(\alpha, \beta) - i\theta(\alpha, \beta)}{\xi - \alpha_0} \right) d\xi \right\} = \int_{-\infty}^{\infty} \frac{\tau(\alpha, 0) - i\theta(\alpha, 0)}{\alpha - \alpha_0} d\alpha. \quad (2.2.25)$$

In order to evaluate the integral along Γ_2 we introduce the substitution $\xi = \alpha_0 + \epsilon e^{it} \implies d\xi = i\epsilon e^{it} dt$. Converting the limits of integration, $\xi = \alpha_0 - \epsilon$ and $\xi = \alpha_0 + \epsilon$, to expressions in terms of t we find:

$$\xi = \alpha_0 - \epsilon = \alpha_0 + \epsilon e^{it} \implies t = -\pi;$$

$$\xi = \alpha_0 + \epsilon = \alpha_0 + \epsilon e^{it} \implies t = 0.$$

Following this, α and β are also found in terms of t by equating the real and imaginary parts of

$$\alpha + i\beta = \xi = \alpha_0 + \epsilon e^{it} = \alpha_0 + \epsilon(\cos(t) + i \sin(t)),$$

which gives:

$$\alpha = \alpha_0 + \epsilon \cos(t), \quad \beta = \epsilon \sin(t).$$

We are now ready to rewrite the integral along Γ_2 in terms of t as

$$\begin{aligned} & \int_{\Gamma_2} \frac{\tau(\alpha, \beta) - i\theta(\alpha, \beta)}{\xi - \alpha_0} d\xi \\ &= \int_{-\pi}^0 \frac{\tau(\alpha_0 + \epsilon \cos(t), \epsilon \sin(t)) - i\theta(\alpha_0 + \epsilon \cos(t), \epsilon \sin(t))}{\alpha_0 + \epsilon e^{it} - \alpha_0} i\epsilon e^{it} dt \\ &= i \int_{-\pi}^0 \tau(\alpha_0 + \epsilon \cos(t), \epsilon \sin(t)) - i\theta(\alpha_0 + \epsilon \cos(t), \epsilon \sin(t)) dt. \end{aligned}$$

Taking the limit $\epsilon \rightarrow 0$ of the above expression we find it simplifies to

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \int_{\Gamma_2} \frac{\tau(\alpha, \beta) - i\theta(\alpha, \beta)}{\xi - \alpha_0} d\xi &= i \int_{-\pi}^0 \tau(\alpha_0, 0) - i\theta(\alpha_0, 0) dt \\ &= i \left(\tau(\alpha_0, 0) - i\theta(\alpha_0, 0) \right) \int_{-\pi}^0 dt = i\pi \left(\tau(\alpha_0, 0) - i\theta(\alpha_0, 0) \right). \end{aligned} \quad (2.2.26)$$

Finally we turn to evaluating the integral along the remaining contour segment Γ_4 . Consulting again Figure 2.2.2 recall that, in the limit $R \rightarrow \infty$, points on the arc Γ_4 correspond to points that are in the far-field downstream of the physical flow. In the far-field we have $1 = u - iv = w = e^{\tau - i\theta}$, this implies that we have $\tau = \theta = 0$. Hence,

$$\lim_{R \rightarrow \infty} \int_{\Gamma_4} \frac{\tau(\alpha, \beta) - i\theta(\alpha, \beta)}{\xi - \alpha_0} d\xi = 0. \quad (2.2.27)$$

We are now prepared to take the limit of (2.2.22) as $R \rightarrow \infty$ and $\epsilon \rightarrow 0$. Taking

these limits and substituting in the results of (2.2.25 - 2.2.27) we obtain

$$i\pi(\tau(\alpha_0, 0) - i\theta(\alpha_0, 0)) + \int \frac{\tau(\alpha, 0) - i\theta(\alpha, 0)}{\alpha - \alpha_0} d\alpha = 0. \quad (2.2.28)$$

Now noting that we have not specified the value of α_0 , which is nothing but a real number and can be thought to act as a variable in the above equation, we swap the symbols α and α_0 . Equating real and imaginary parts of (2.2.28) we obtain:

$$\theta(\alpha, 0) = -\frac{1}{\pi} \int \frac{\tau(\alpha_0, 0)}{\alpha_0 - \alpha} d\alpha_0; \quad (2.2.29)$$

$$\tau(\alpha, 0) = \frac{1}{\pi} \int \frac{\theta(\alpha_0, 0)}{\alpha_0 - \alpha} d\alpha_0. \quad (2.2.30)$$

We will now denote the values of τ and θ on the free surface, where $\alpha > 0$, as τ_f and θ_f respectively. Similarly we denote the values of τ and θ along the topography, where we have $\alpha < 0$, as τ_b and θ_b .

Consider (2.2.29) and (2.2.30) for values of $\alpha > 0$; splitting the integrals and writing in terms of surface and topography variables we find

$$\begin{aligned} \theta_f(\alpha) &= -\frac{1}{\pi} \int_{-\infty}^0 \frac{\tau_b(\alpha_0)}{\alpha_0 - \alpha} d\alpha_0 - \frac{1}{\pi} \int_0^{\infty} \frac{\tau_f(\alpha_0)}{\alpha_0 - \alpha} d\alpha_0, \\ \tau_f(\alpha) &= \frac{1}{\pi} \int_{-\infty}^0 \frac{\theta_b(\alpha_0)}{\alpha_0 - \alpha} d\alpha_0 + \frac{1}{\pi} \int_0^{\infty} \frac{\theta_f(\alpha_0)}{\alpha_0 - \alpha} d\alpha_0. \end{aligned}$$

Introduce into the first integral terms of each of the above equations the change of variables $\alpha = e^{\pi\phi}$ and $\alpha_0 = -e^{\pi\phi_0}$, and then into the second integral terms the change of variables $\alpha = e^{\pi\phi}$, $\alpha_0 = e^{\pi\phi_0}$. Rearranging the result, this yields

$$\theta_f(\phi) = \int_{-\infty}^{\infty} \frac{\tau_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\tau_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0, \quad (2.2.31)$$

$$\tau_f(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0. \quad (2.2.32)$$

By a similar process we take (2.2.29) and (2.2.30) with $\alpha < 0$ and split the

integrals to obtain

$$\theta_b(\alpha) = -\frac{1}{\pi} \int_{-\infty}^0 \frac{\tau_b(\alpha_0)}{\alpha_0 - \alpha} d\alpha_0 - \frac{1}{\pi} \int_0^{\infty} \frac{\tau_f(\alpha_0)}{\alpha_0 - \alpha} d\alpha_0,$$

$$\tau_b(\alpha) = \frac{1}{\pi} \int_{-\infty}^0 \frac{\theta_b(\alpha_0)}{\alpha_0 - \alpha} d\alpha_0 + \frac{1}{\pi} \int_0^{\infty} \frac{\theta_f(\alpha_0)}{\alpha_0 - \alpha} d\alpha_0.$$

Now we make the change of variables $\alpha = -e^{\pi\phi}$ and $\alpha_0 = -e^{\pi\phi_0}$ to the first integral terms of each of the above equations. To the second integral terms we make the change of variables $\alpha = -e^{\pi\phi}$ and $\alpha_0 = e^{\pi\phi_0}$. As before these equations are simplified to read

$$\theta_b(\phi) = \int_{-\infty}^{\infty} \frac{\tau_b(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\tau_f(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0, \quad (2.2.33)$$

$$\tau_b(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0. \quad (2.2.34)$$

Because the function $\frac{df}{dz} = w = e^{\tau - i\theta}$ is an analytic function we have also that its reciprocal $\frac{dz}{df} = \frac{1}{w} = e^{-\tau + i\theta}$ is analytic. Hence, we can write

$$x_\phi + iy_\phi = \frac{dz}{df} = e^{-\tau + i\theta}.$$

Finally, applying this relation to the free surface and topography and then equating the real and imaginary components gives:

$$\begin{aligned} \frac{\partial x_b}{\partial \phi} &= e^{-\tau_b} \cos(\theta_b); & \frac{\partial x_f}{\partial \phi} &= e^{-\tau_f} \cos(\theta_f); \\ \frac{\partial y_b}{\partial \phi} &= e^{-\tau_b} \sin(\theta_b); & \frac{\partial y_f}{\partial \phi} &= e^{-\tau_f} \sin(\theta_f). \end{aligned}$$

Collecting together the equations derived above the boundary integral method gives the system of equations to be solved as:

$$\frac{1}{2}e^{2\tau} + \frac{1}{F^2}(y_f + P) = \frac{1}{2} + \frac{1}{F^2} \quad (2.2.35)$$

$$\theta_f(\phi) = \int_{-\infty}^{\infty} \frac{\tau_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\tau_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0; \quad (2.2.36)$$

$$\theta_b(\phi) = \int_{-\infty}^{\infty} \frac{\tau_b(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\tau_f(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0; \quad (2.2.37)$$

$$\tau_f(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0; \quad (2.2.38)$$

$$\tau_b(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0; \quad (2.2.39)$$

$$\frac{\partial x_b}{\partial \phi} = e^{-\tau_b} \cos(\theta_b); \quad (2.2.40) \quad \frac{\partial x_f}{\partial \phi} = e^{-\tau_f} \cos(\theta_f); \quad (2.2.41)$$

$$\frac{\partial y_b}{\partial \phi} = e^{-\tau_b} \sin(\theta_b); \quad (2.2.42) \quad \frac{\partial y_f}{\partial \phi} = e^{-\tau_f} \sin(\theta_f). \quad (2.2.43)$$

2.3 Linear algebra

2.3.1 Solutions to linear equations

Consider the matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (2.3.1)$$

and the general problem of seeking all possible solutions for the unknown $(n \times 1)$ vector $\mathbf{x} = (x_1, \dots, x_n)^T$ given the known quantities \mathbf{A} , an $(m \times n)$ matrix, and $\mathbf{b} = (b_1, \dots, b_m)^T$, an $(m \times 1)$ vector, such that (2.3.1) is satisfied. This is equivalent to solving a system of m simultaneous linear equations in n variables.

If there exists at least one vector \mathbf{x} such that (2.3.1) is satisfied, i.e. all m equations may be simultaneously satisfied, then the system is referred to as consistent. If no such solutions exist then the system is inconsistent. A consistent system may have either exactly one solution or instead an infinite number of them, in which case it is called an indeterminate system.

A set of N vectors $V = \{\underline{\mathbf{v}}_i \mid i = 1, \dots, N\}$ is called linearly independent if no one element can be expressed as a linear combination of the others, that is to say that the only solution to

$$\sum_i c_i \underline{\mathbf{v}}_i = \underline{\mathbf{0}}$$

for the coefficients c_i is given by $c_i = 0 \forall i$. If the set is not linearly independent then it is called linearly dependent.

The span of the set of vectors V is given by the set of all possible linear combinations of the elements $\underline{\mathbf{v}}_i$ or

$$\text{span}(V) = \left\{ \sum_i c_i \underline{\mathbf{v}}_i \mid c_i \in \mathbb{R}, \underline{\mathbf{v}}_i \in V \right\}.$$

By denoting the i^{th} column of \mathbf{A} by $\underline{\mathbf{a}}_i$ define the column space of \mathbf{A} as $\text{colsp}(\mathbf{A}) = \text{span}(\{\underline{\mathbf{a}}_i \mid i = 1, \dots, n\})$. The row space of \mathbf{A} may be defined similarly. Now define the rank of \mathbf{A} to be equal to the dimension of its column space which may be expressed as

$$\text{rank}(\mathbf{A}) = \dim(\text{colsp}(\mathbf{A}))$$

and can be shown to be also equal to the dimension of the row space. The rank is bounded above by $\text{rank}(\mathbf{A}) \leq \min(m, n)$ with the case of equality being referred to as a full-rank matrix. If this is not the case it is rank deficient.

Expanding the left-hand side of (2.3.1) as

$$\mathbf{A}\underline{\mathbf{x}} = x_1 \underline{\mathbf{a}}_1 + \dots + x_n \underline{\mathbf{a}}_n = \underline{\mathbf{b}}$$

it becomes clear that the system is consistent only if $\underline{\mathbf{b}} \in \text{colsp}(\mathbf{A})$.

The $(m \times n + 1)$ matrix formed by appending the vector $\underline{\mathbf{b}}$ as an extra column to the matrix \mathbf{A} is known as an augmented matrix and is written $[\mathbf{A}|\underline{\mathbf{b}}]$. For a consistent system we can see that $\text{rank}(\mathbf{A}) = \text{rank}([\mathbf{A}|\underline{\mathbf{b}}])$; because $\underline{\mathbf{b}} \in \text{colsp}(\mathbf{A})$ the addition of this column leaves the number of linearly independent columns

unchanged.

The Rouché-Capelli theorem (Capelli, 1892) states that the linear system given by (2.3.1) is consistent if and only if $\text{rank}(\mathbf{A}) = \text{rank}([\mathbf{A}|\mathbf{b}])$. Further, if we also have $\text{rank}(\mathbf{A}) = n$ then the solution is unique, otherwise an infinite number of solutions may be found.

Let us now consider the different possibilities for systems of m equations in n unknowns for various m, n .

Underdetermined, $m < n$: A system with fewer equations than unknowns is referred to as an underdetermined system and has $m < n$. Therefore, in an underdetermined system, we have $\text{rank}(\mathbf{A}) \leq \min(m, n) = m < n$ giving $\text{rank}(\mathbf{A}) \neq n$. Applying the Rouché-Capelli theorem the system will either: have an infinite number of solutions if $\text{rank}(\mathbf{A}) = \text{rank}([\mathbf{A}|\mathbf{b}])$ i.e., $\mathbf{b} \in \text{colsp}(\mathbf{A})$; or else be inconsistent with no solutions.

Square, $m = n$: A system with an equal number of equations and unknowns is referred to as a square system and has $m = n$. Hence, a square system has $\text{rank}(\mathbf{A}) \leq n$. Let us consider the two separate cases of $\text{rank}(\mathbf{A}) = n$ and $\text{rank}(\mathbf{A}) < n$:

- **Square full rank, $\text{rank}(\mathbf{A}) = n$:** In the case of a full rank square matrix there always exists a unique solution. This can be seen by either: simply multiplying (2.3.1) by the inverse of \mathbf{A} , which will exist as \mathbf{A} is of full rank; or combining the relations that the rank of a matrix is both bounded above by the minimum of the number of rows and columns and bounded below by the rank of any contained sub-matrix (namely \mathbf{A} for the augmented matrix $[\mathbf{A}|\mathbf{b}]$) to obtain the relation $n = \text{rank}(\mathbf{A}) \leq \text{rank}([\mathbf{A}|\mathbf{b}]) \leq \min(m, n + 1) = m = n$ which requires that $\text{rank}(\mathbf{A}) = \text{rank}([\mathbf{A}|\mathbf{b}]) = n$ and so by application of the Rouché-Capelli theorem the system has a unique solution.

• **Square rank deficient, $\text{rank}(\mathbf{A}) < n$:** In the case of a rank deficient square matrix our system acts effectively as an underdetermined system; if $\text{rank}(\mathbf{A}) = \text{rank}([\mathbf{A}|\mathbf{b}])$ then an infinite number of solutions exist, otherwise no solutions exist.

Overdetermined, $m > n$: A system with more equations than unknowns is referred to as an overdetermined system and has $m > n$. As for square and underdetermined systems we treat this with the Rouché-Capelli theorem: if $\text{rank}(\mathbf{A}) \neq \text{rank}([\mathbf{A}|\mathbf{b}])$ then no solutions exist; if $\text{rank}(\mathbf{A}) = \text{rank}([\mathbf{A}|\mathbf{b}])$ and the matrix \mathbf{A} has full rank then we also have $\text{rank}(\mathbf{A}) = n$ so there is a unique solution, whereas if \mathbf{A} is rank deficient but we still have $\text{rank}(\mathbf{A}) = \text{rank}([\mathbf{A}|\mathbf{b}])$ then there are an infinite number of solutions.

A powerful and commonly employed tool in the analysis of matrix equations is the study of the eigenvalues, λ , and corresponding eigenvectors, $\underline{\mathbf{v}}$, of the matrix \mathbf{A} which satisfy

$$\mathbf{A}\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}},$$

however, these are only defined when \mathbf{A} is a square matrix. A related and more general notion is that of the singular value decomposition (SVD) (e.g. Griffel, 1989), which is defined for all matrices regardless of size. We limit ourselves here to the consideration only of real matrices, however the SVD analysis holds for matrices with complex entries if the transpose operation is replaced with that of the conjugate transpose. The SVD of any $(m \times n)$ matrix \mathbf{A} may be written in the form

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

where \mathbf{U} , \mathbf{V} are respectively $(m \times m)$ and $(n \times n)$ unitary matrices whose columns consist of the eigenvectors of the matrices $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$. The $(m \times n)$ matrix $\mathbf{\Sigma}$ has along its leading diagonal the singular values σ_n , $n = 1, \dots, r \leq \min(m, n)$, corresponding to the positive square roots of the eigenvalues of $\mathbf{A}\mathbf{A}^T$, and all other elements zero. The number of singular values that are not zero, r , is the

same as the rank of the matrix \mathbf{A} . The singular values are ordered by size starting with the largest in the (1, 1) position.

While the inverse \mathbf{A}^{-1} of \mathbf{A} is defined only if \mathbf{A} is a full rank square matrix the idea can be generalised to all $(m \times n)$ matrices regardless of rank. The Moore-Penrose inverse (MP-inverse), \mathbf{A}^+ , is uniquely defined for all matrices \mathbf{A} (James, 1978; Penrose, 1955) as the matrix that satisfies the properties:

$$\begin{aligned} \mathbf{A}\mathbf{A}^+\mathbf{A} &= \mathbf{A}; & \mathbf{A}^+\mathbf{A}\mathbf{A}^+ &= \mathbf{A}^+; \\ (\mathbf{A}\mathbf{A}^+)^T &= \mathbf{A}\mathbf{A}^+; & (\mathbf{A}^+\mathbf{A})^T &= \mathbf{A}^+\mathbf{A}. \end{aligned}$$

If \mathbf{A} is invertible then we have $\mathbf{A}^+ = \mathbf{A}^{-1}$. The MP-inverse can be written in terms of the SVD of \mathbf{A} as

$$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T, \quad (2.3.2)$$

where $\mathbf{\Sigma}^+$ is formed by taking the reciprocal of all non-zero entries of $\mathbf{\Sigma}$ and then transposing (e.g. Ben-Israel and Greville, 2003, p.207).

The solvability condition for (2.3.1) gives that solutions exist if

$$\mathbf{A}^+\mathbf{A}\underline{\mathbf{b}} = \underline{\mathbf{b}} \quad (2.3.3)$$

is satisfied. Where solutions exist they can all be constructed by use of the MP-inverse as

$$\underline{\mathbf{x}} = \mathbf{A}^+\underline{\mathbf{b}} + (\mathbf{I}_n - \mathbf{A}^+\mathbf{A})\underline{\mathbf{w}}, \quad (2.3.4)$$

where $\underline{\mathbf{w}}$ is an arbitrary column vector of length n . The second term on the RHS of in (2.3.4) represents solutions in the null space of \mathbf{A} . If \mathbf{A} is of full column rank then the solution

$$\underline{\mathbf{x}}_p = \mathbf{A}^+\underline{\mathbf{b}} \quad (2.3.5)$$

will be unique; in this case we have $\mathbf{A}^+\mathbf{A} = \mathbf{I}_n$ and so there is a cancellation of the second term of (2.3.4), i.e., the null space of M contains only the zero vector. This corresponds to the cases of a unique solution existing for consistent

overdetermined and square systems. When \mathbf{A} is not of full rank an infinite number of solutions can be constructed, this is the case of a consistent underdetermined system.

If the solvability condition fails then $\underline{\mathbf{b}} \notin \text{Im}\mathbf{M}$, the system (2.3.1) is inconsistent. Numerically the solvability condition will often fail due to rounding errors. It can be shown both that (2.3.4) now provides the solutions to the related least squares minimisation problem of

$$\min_{\underline{\mathbf{x}}} \|\mathbf{A}\underline{\mathbf{x}} - \underline{\mathbf{b}}\|_2. \quad (2.3.6)$$

and that from these minimising solutions there is a unique vector of minimum Euclidean norm, given by $\underline{\mathbf{x}}_p = \mathbf{A}^+\underline{\mathbf{b}}$ (Planitz, 1979).

The condition number of the matrix \mathbf{A} is defined as the ratio of the largest and smallest singular values of \mathbf{A} as $\text{cond}(\mathbf{A}) = \frac{\max \sigma_i}{\min \sigma_i}$. If \mathbf{A} is square and not of full rank then it will necessarily have at least one singular value be zero and so have an infinite condition number. While a square matrix with finite condition number is technically invertible, the more poorly conditioned (the larger the condition number) a matrix is the more difficult it is to invert and poorly conditioned matrices often are problematic in a numerical setting as the errors inherent to working in finite precision will be magnified when trying to invert the matrix. This makes trying to solve poorly conditioned matrix problematic and an example of an ill-posed problem in the sense that a small change in input data can lead to a disproportionate change in the output.

The method of truncated singular value decomposition circumvents the issue of ill-conditioning by seeking to obtain instead only approximate solutions to (2.3.1) by solving a related system formed by replacing the matrix \mathbf{A} with a lower rank approximation \mathbf{A}_κ . At a chosen rank κ the $(m \times n)$ matrix \mathbf{A}_κ can be written in terms of the SVD of \mathbf{A} as

$$\mathbf{A}_\kappa = \mathbf{U}\mathbf{\Sigma}_\kappa^+\mathbf{V}^T, \quad (2.3.7)$$

where $\kappa \leq \text{rank}(\mathbf{A})$ and $\mathbf{\Sigma}_\kappa$ is formed by retaining the largest k singular values in $\mathbf{\Sigma}$ and setting the others to zero. By the Eckart–Young–Mirsky theorem the matrix \mathbf{A}_κ is the matrix that minimises $\|\mathbf{A} - \mathbf{A}_\kappa\|_F$ (Eckart and Young, 1936). The resulting rank k approximate problem has least squares solutions of the form as that of (2.3.4), from which the unique solution $\underline{\mathbf{x}}_\kappa$ that also minimises $\|\underline{\mathbf{x}}_\kappa\|$ is chosen as

$$\underline{\mathbf{x}}_\kappa = \mathbf{A}_\kappa^+ \underline{\mathbf{b}}. \quad (2.3.8)$$

2.4 Newton Method

2.4.1 One equation in one variable

The Newton method is an iterative method for solving nonlinear, allowing us to numerically seek a root, x_s , of a function, $F(x)$ such that

$$F(x_s) = 0. \quad (2.4.1)$$

Given an estimate X_i to the root x_s we can write $x_s = X_i + \epsilon_i$, or $\epsilon_i = x_s - X_i$, so that ϵ_i represents the error between our current estimate and the true solution. A linear approximation is made by taking the Taylor expansion of (2.4.1), as

$$F(x_s) = F(X_i + \epsilon_i) = F(X_i) + \epsilon_i F'(X_i) + \mathcal{O}(\epsilon_i^2) = 0, \quad (2.4.2)$$

and omitting the $\mathcal{O}(\epsilon_i^2)$ terms to yield

$$F(X_i) + \epsilon_i F'(X_i) \approx 0. \quad (2.4.3)$$

We use (2.4.3) to obtain an approximation for ϵ_i , which is then used to form the improved estimate, $X_{i+1} \approx X_i + \epsilon_i = x_s$, defined as

$$X_{i+1} = X_i - \frac{F(X_i)}{F'(X_i)}. \quad (2.4.4)$$

From an initial estimate X_0 the process is iterated until the magnitude of the residual $|F(X_i)|$ falls below some defined tolerance.

2.4.2 Systems of equations

Newton's method can be generalised to solve a system of equations n equations in n unknowns. Consider instead the system

$$\underline{\mathbf{F}}(\underline{\mathbf{x}}) = \underline{\mathbf{0}} \quad (2.4.5)$$

where $\underline{\mathbf{F}}(\underline{\mathbf{x}}) = (F_1(\underline{\mathbf{x}}), \dots, F_n(\underline{\mathbf{x}}))^T$ and $\underline{\mathbf{x}} = (x_1, \dots, x_n)^T$. For a given iteration we have an estimate $\underline{\mathbf{X}}_i = (X_1, \dots, X_n)^T$ to the true solution $\underline{\mathbf{x}}_s = (x_{[s,1]}, \dots, x_{[s,n]})^T$ and a corresponding error $\underline{\boldsymbol{\epsilon}}_i = \underline{\mathbf{x}}_s - \underline{\mathbf{X}}_i$. Proceeding as in the single variable case we take the first order Taylor expansion of $\underline{\mathbf{F}}(\underline{\mathbf{x}})$ around $\underline{\mathbf{x}} = \underline{\mathbf{x}}_s$, now written in matrix form as

$$\underline{\mathbf{F}}(\underline{\mathbf{X}}_i) + J_F \underline{\boldsymbol{\epsilon}}_i \approx \underline{\mathbf{0}}, \quad (2.4.6)$$

where J_F , the Jacobian matrix of $\underline{\mathbf{F}}$ evaluated at $\underline{\mathbf{x}} = \underline{\mathbf{X}}_i$, is given by

$$J_F = \left(\begin{array}{ccc} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{array} \right) \bigg|_{\underline{\mathbf{x}}=\underline{\mathbf{X}}_i} .$$

Rearranging to find an approximation for $\underline{\boldsymbol{\epsilon}}_i$, we then define the improved estimate $\underline{\mathbf{X}}_{i+1} \approx \underline{\mathbf{X}}_i + \underline{\boldsymbol{\epsilon}}_i = \underline{\mathbf{x}}_s$ as

$$\underline{\mathbf{X}}_{i+1} = \underline{\mathbf{X}}_i - J_F^{-1} \underline{\mathbf{F}}(\underline{\mathbf{X}}_i) \quad (2.4.7)$$

and iterate until $|\underline{\mathbf{F}}(\underline{\mathbf{X}}_i)|$ falls below a defined tolerance. A consequence of (2.4.7) is the requirement that we consider only square systems, with an equal number of equations as there is variables, else the inverse of J_F will not exist.

2.5 Inverse problems

While the term inverse problem is somewhat loosely defined it can be helpful to think of this type of problem as determining a question given an answer or finding a set of causes given their effects. A good review on the topic is found in Keller, 1976. A typical example is X-ray tomography where it is attempted to construct a 3D image of the inside of object from information obtained at its surface about the transmission of X-rays through the object along radial slices. Further examples can be found in Kabanikhin, 2008 and in Groetsch, 2007. By their nature they are often ill-posed problems in the sense of Hadamard well-posedness in that they permit multiple solutions or depend discontinuously on their input data. Relevant to the topic at hand are matrix inverse problems and Fredholm equations of the first kind.

2.5.1 Matrix equations and examples

Consider the matrix equation given by $\mathbf{Ax} = \mathbf{b}$ where

$$\mathbf{A} = \begin{pmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 19.25 \\ 6.84 \end{pmatrix}.$$

The solution to this equation can be calculated as $\mathbf{x} = (2.9, 2)^T$. However, the matrix \mathbf{A} is poorly conditioned with $\text{cond}(\mathbf{A}) = \frac{\sigma_1}{\sigma_2} \approx \frac{5.799}{1.72 \times 10^{-3}} \approx 3363$. As such we expect that small perturbation to the entries of \mathbf{A} or \mathbf{b} will cause a larger change in the solution \mathbf{x} . For example, introducing the perturbed quantities

$$\bar{\mathbf{A}} = \begin{pmatrix} 4.51 & 3.1 \\ 1.6 & 1.1 \end{pmatrix} \quad \text{and} \quad \bar{\mathbf{b}} = \begin{pmatrix} 19.25 \\ 6.83 \end{pmatrix},$$

allows for three possible perturbed systems: a perturbation to only \mathbf{b} giving the system $\mathbf{Ax} = \bar{\mathbf{b}}$; a perturbation to only \mathbf{A} giving the system $\bar{\mathbf{A}}\mathbf{x} = \mathbf{b}$; a perturbation to both \mathbf{A} and \mathbf{b} giving the system $\bar{\mathbf{A}}\mathbf{x} = \bar{\mathbf{b}}$. The solutions to these

System:	$\mathbf{Ax} = \underline{\mathbf{b}}$	$\mathbf{Ax} = \bar{\mathbf{b}}$	$\bar{\mathbf{A}}\mathbf{x} = \underline{\mathbf{b}}$	$\bar{\mathbf{A}}\mathbf{x} = \bar{\mathbf{b}}$
Solution by direct inversion:	$\begin{pmatrix} 2.9 \\ 2 \end{pmatrix}$	$\begin{pmatrix} -0.2 \\ 6.5 \end{pmatrix}$	$\begin{pmatrix} -29 \\ 48.4 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 3.3 \end{pmatrix}$
Truncated system:	$\mathbf{A}_k\mathbf{x} = \underline{\mathbf{b}}$	$\mathbf{A}_k\mathbf{x} = \bar{\mathbf{b}}$	$\bar{\mathbf{A}}_k\mathbf{x} = \underline{\mathbf{b}}$	$\bar{\mathbf{A}}_k\mathbf{x} = \bar{\mathbf{b}}$
Approximate solution by pseudoinverse:	$\begin{pmatrix} 2.901 \\ 1.999 \end{pmatrix}$	$\begin{pmatrix} 2.901 \\ 1.998 \end{pmatrix}$	$\begin{pmatrix} 2.899 \\ 1.993 \end{pmatrix}$	$\begin{pmatrix} 2.899 \\ 1.993 \end{pmatrix}$

Table 2.1: A comparison of the direct solutions to a perturbed matrix equation and the regularised solutions computed by way truncated singular value decomposition (TSVD).

systems are shown in the second row of Table 2.1 and it can be seen that small variations in the input data \mathbf{A} and $\underline{\mathbf{b}}$ cause a much larger variation in the output $\underline{\mathbf{x}}$.

Now regularise the problem by applying the method of truncated singular value decomposition. Form the matrices \mathbf{A}_1 and $\bar{\mathbf{A}}_1$ by calculating their respective SVDs and replacing their σ_2 singular values with zero. By then taking the Moore-Penrose pseudoinverse of these matrices the approximate solutions can be written in the form $\underline{\mathbf{x}}_{\kappa} = \mathbf{A}_{\kappa}^+\underline{\mathbf{b}}$. The regularised solutions to these perturbed systems are shown in the fourth row of Table 2.1, it is readily apparent that there is significantly less variation in the regularised approximate solutions than in those solutions found by direct inversion of the perturbed systems, and that the regularised solutions all closely approximate the solution to the true unperturbed problem.

2.5.2 Fredholm equations of the first kind and examples

A Fredholm equation of the first kind is of the form

$$\int_a^b f(y)K(x, y) dy = g(x) \quad (2.5.1)$$

where the solution $f(x)$ is sought and the functions $K(x, y)$ and $g(x)$ are known.

We first consider the example problem with $K(x, y) = \sin(x - y)$ and $g(x) = \left(\sin(L) \cos(L) - L \right) \cos(x)$, that is

$$\int_{-L}^L f(y) \sin(x - y) dy = \left(\sin(L) \cos(L) - L \right) \cos(x), \quad (2.5.2)$$

which has been constructed so as to allow the solution $f(x) = \sin(x)$. Note however that, depending on the choice for L , this solution can be non-unique; consider substituting the trial solution $f_c(x) = \sin(x) + c$, for some constant c , into the left-hand side of (2.5.2) to give

$$\begin{aligned} \int_{-L}^L \left(\sin(y) + c \right) \sin(x - y) dy &= \int_{-L}^L \sin(y) \sin(x - y) dy + c \int_{-L}^L \sin(x - y) dy \\ &= \left(\sin(L) \cos(L) - L \right) \cos(x) + 2 \sin(L) \cos(x). \end{aligned}$$

From the above it can be seen that a choice of $L = n\pi$ for $n \in \mathbb{N}^+$ ensures that the second term vanishes and as such f_c is also a solution to (2.5.2) for any value c when $L = n\pi$. Similarly, taking $m \in \mathbb{Z} \setminus \{-1, 1\}$ and defining $f_m(x) = \sin(mx)$ then the integral

$$\begin{aligned} \int_{-L}^L f_m(y) \sin(x - y) dy &= \sin(x) \int_{-L}^L \sin(my) \cos(y) dy \\ &\quad - \cos(x) \int_{-L}^L \sin(my) \sin(y) dy \end{aligned}$$

will vanish for values $L = n\pi$ due to the orthogonality properties of the sin and cos functions. Accordingly, $f_{c,m} = \sin(x) + c_m \sin(mx) + c$ will also solve (2.5.2) when $L = n\pi$. It can be shown that $f(x) = \sin(x)$ is the minimum norm solution to (2.5.2).

In order to attempt to solve numerically (2.5.1) the N equally-spaced

mesh-points $x_i = -L + \Delta(i - 1)$ are introduced, where $\Delta = \frac{2L}{N-1}$, y_i is defined analogously, before approximating the integrals by the trapezoidal rule. Writing $g_i = g(x_i)$, $f_i = f(y_i)$ and $K_{i,j} = K(x_i, y_j)$ this methodology yields the matrix system

$$\mathbf{M}\underline{\mathbf{f}} = \underline{\mathbf{g}} \quad (2.5.3)$$

where $\underline{\mathbf{f}} = (f_1, \dots, f_N)^T$, $\underline{\mathbf{g}} = (g_1, \dots, g_N)^T$ and the matrix \mathbf{M} has entries $m_{i,j}$ given by:

$$m_{i,1} = \frac{\Delta}{2} K_{i,1}; \quad m_{i,N} = \frac{\Delta}{2} K_{i,N}; \quad m_{i,j} = \Delta K_{i,j}.$$

Considering the system corresponding to (2.5.2) with $K(x, y) = \sin(x - y)$ it can be found that, for $N \geq 2$, $\text{rank}(\mathbf{M}) = 2$ and so the system cannot be solved by direct inversion of \mathbf{M} . The rank is reflected in a plot of the computed singular values of \mathbf{M} ; the first two singular values are $\mathcal{O}(1)$ whereas all subsequent singular values are on the order of machine precision (Figure 2.5.1), an artifact of their computation as they should be exactly zero. For this kernel there is a clear separation in the singular values between the non-zero and numerically zero singular values and so the choice of truncation rank $\kappa = 2$ is readily made (e.g. Hansen, 1987). TSVD solutions for $\kappa = 2$ and $\kappa = 3$ are shown alongside the true solution in Figure 2.5.2. Truncation of M with $\kappa = 2$ allows for recovery of the original solution to Fredholm equation whereas the output for $\kappa = 3$ displays oscillations around the true solution. Further increasing κ would lead to error quickly dominating the output with rapid oscillations occurring on the mesh-spacing. Note that the output solutions are approximately $\sin(x)$, although adding a constant or multiples of $\sin(mx)$ to his solution would solve the system the TSVD method selects the solution of minimum norm.

Often, inverse problems require solving an equation like (2.5.3) where $\underline{\mathbf{g}}$ is not known exactly but determined experimentally and so contains a certain level of error. To simulate this the right-hand side of (2.5.3) was replaced with $\underline{\mathbf{g}}_\epsilon$ where $\underline{\mathbf{g}}_\epsilon = \underline{\mathbf{g}} + \underline{\boldsymbol{\epsilon}}$ and $\underline{\boldsymbol{\epsilon}}$ is a column vector whose entries are randomly drawn from a

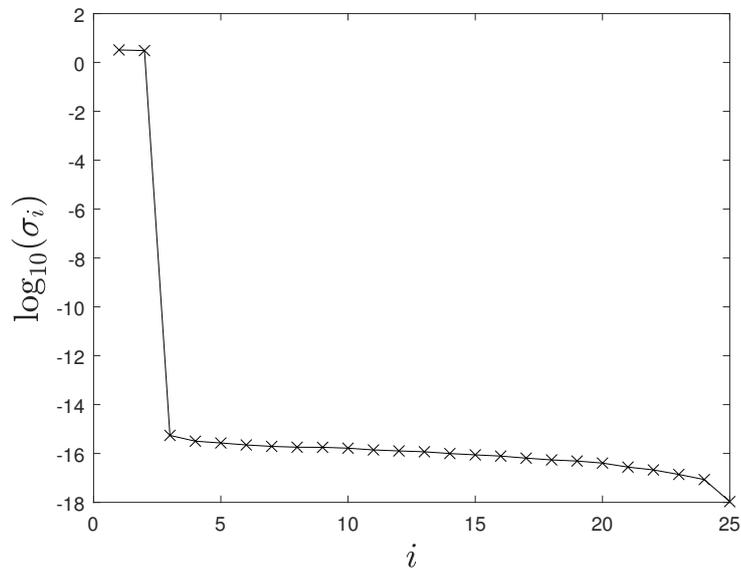


Figure 2.5.1: A plot of the singular values of the matrix corresponding to the discretised kernel $K(x, y) = \sin(x - y)$ with $L = \pi$, $N = 25$.

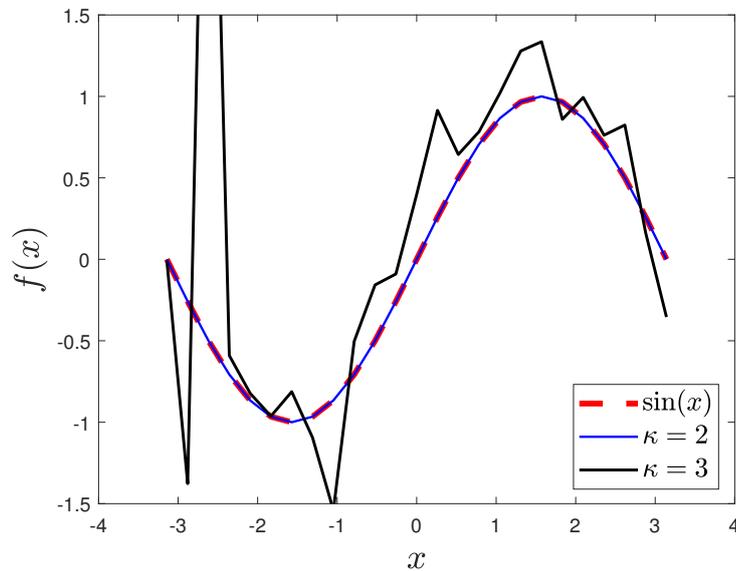


Figure 2.5.2: $L = \pi$, $N = 25$ A comparison of the true solution (red dashed) to the truncated solutions \underline{f}_{κ} for truncation ranks $\kappa = 2$ (blue) and $\kappa = 3$ (black).

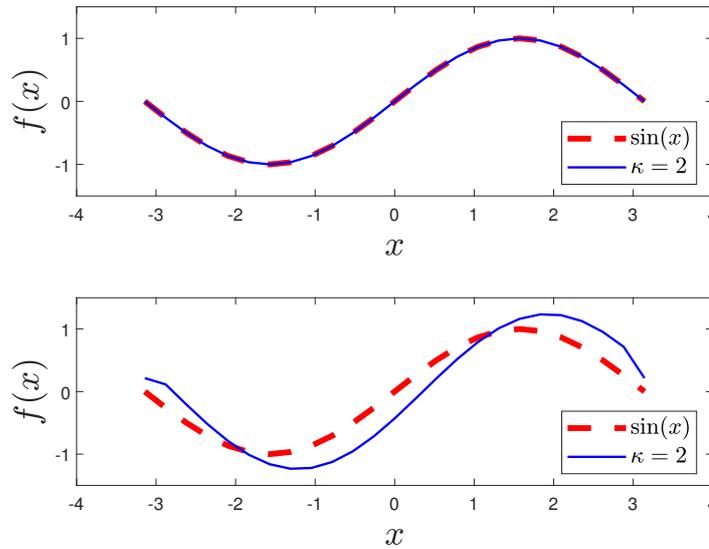


Figure 2.5.3: $L = \pi, N = 25$ A comparison of the true solution (red dashed) to (2.5.2) to the truncated solution \mathbf{f}_κ to the perturbed system for truncation rank $\kappa = 2$ (blue). Top: Standard deviation of added noise $\epsilon = 2 \times 10^{-4}$. Bottom: Standard deviation of added noise $\epsilon = 2$.

normal distribution of zero mean and fixed standard deviation ϵ .

Figure 2.5.3 shows the results of the solving by TSVD the perturbed system $\mathbf{M}\mathbf{f} = \mathbf{g}_\epsilon$ for two different values of ϵ . The clear separation in the singular values means this system is numerically well-conditioned and the addition of noise does not drastically alter the output as it would for an ill-conditioned system. By truncating a matrix we are effectively improving its numerical conditioning to improve the stability of output solutions with respect to perturbations at the expense replacing the true system with an approximating one.

Now consider a Fredholm integral equation of the form (2.5.1) with the kernel $K(x, y) = \frac{1}{2} \left(1 + \tanh \left(\frac{\pi(x-y)}{2} \right) \right)$. It will be shown in Section 3.3.2 that the inverse problem for the topography involves a Fredholm equation of the first kind with this exact kernel but a different right-hand side. As before we shall construct the problem so as to have the inverse solution $f(x) = \sin(x)$, for this kernel however we are unable to perform the integration to obtain an expression for the right-hand side, $g(x)$, and instead these values must be computed first in

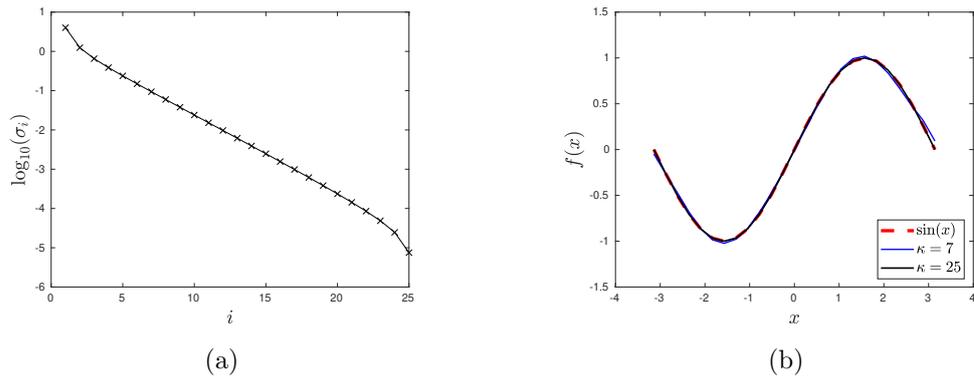


Figure 2.5.4: (a) A plot of the singular values of the matrix corresponding to the discretised kernel $K(x, y) = \frac{1}{2} \left(1 + \tanh \left(\frac{\pi(x-y)}{2} \right) \right)$ with $L = \pi, N = 25$. (b) Inverse solutions for the truncation rank $\kappa = 7$ (blue) and $\kappa = 25$ (black) compared to the true solution (red)

the forward sense before being used in the inverse problem.

For the discretised system resulting from this kernel there is no clear separation in the singular values (Figure 2.5.4(a)), which now decay smoothly down to the level of machine precision for a sufficiently large value of N (Figure 2.5.5(a)). For lower values of N this system can be accurately solved by direct inversion (Figure 2.5.4(b)), i.e. retaining all singular values, however for higher values of N the retention of the smallest singular values leads to the appearance of large oscillatory errors in the solution (Figure 2.5.5(b)) as the matrix becomes more poorly conditioned. However, for these higher values of N , accurate solutions can still be obtained given the correct choice of the truncation rank κ .

Hansen, 1990a develop on the concept of the Discrete Picard Condition (DPC) for discrete ill-posed problems, first discussed in Varah (1979, 1983). The DPC concerns the relative decay of what are referred to as the Fourier coefficients, $|\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}|$, and the singular values, σ_i . For the system $\mathbf{M}\underline{\mathbf{f}} = \underline{\mathbf{g}}$ it is said that $\underline{\mathbf{g}}$ satisfies the DPC if for all the numerically nonzero singular values σ_i the corresponding Fourier coefficients $|\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}|$ decay to zero faster than σ_i . The authors bound the regularisation error in terms of the ratio $\frac{\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}}{\sigma_i}$ such that satisfaction of the DPC means that useful regularised approximate solutions can be obtained. While the unperturbed underlying problem may satisfy the DPC,

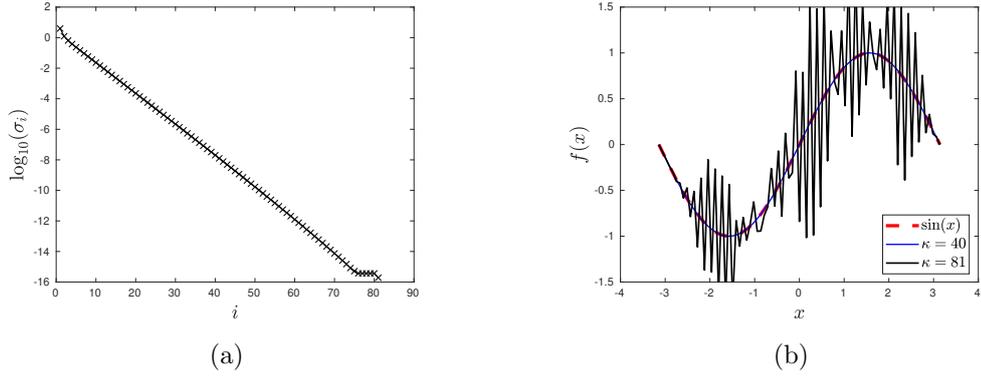


Figure 2.5.5: $L = \pi, N = 81$ (a) A plot of the singular values of the matrix corresponding to the discretised kernel $K(x, y) = \frac{1}{2} \left(1 + \tanh \left(\frac{\pi(x-y)}{2} \right) \right)$ with $L = \pi, N = 81$. (b) Inverse solutions for the truncation rank $\kappa = 7$ (blue) and $\kappa = 25$ (black) compared to the true solution (red).

the authors note that in practice the errors in the entries of \mathbf{M} and $\underline{\mathbf{g}}$ often mean that the DPC is found numerically to not be satisfied by all the Fourier coefficients. However, if the DPC is satisfied by the underlying problem then a proper selection of the truncation rank k yields a regularised system that also satisfies the DPC. Conversely, if the underlying problem does not satisfy the DPC then in general it can not be said if useful solutions can be obtained by this method.

In order to numerically check satisfaction of the DPC Hansen, 1990a suggest looking at the decay of the moving geometric mean ρ_i given by

$$\rho_i = \frac{1}{\sigma_i} \left(\prod_{j=i-q}^{i+q} |\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}| \right)^{\frac{1}{2q+1}}, \quad (2.5.4)$$

for $i = q + 1, \dots, N - q$ and where q is a small integer, which should be computed only for numerically nonzero σ_i and $|\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}|$. When considering a perturbed problem the perturbed Fourier coefficients will decay until they reflect the level of noise present, for this reason the perturbed Fourier coefficients for which $|\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}| < \epsilon$ should be considered numerically zero. The authors recommend taking q to be equal to 1, 2 or 3 and for the course of the present work we have taken $q = 3$. The authors then take the DPC to be

satisfied when ρ_i decays ‘on average’ monotonically to zero. They suggest not basing the choice for the truncation rank only on the ρ_i curve but using this plot in conjunction with other strategies.

Another such tool useful in selecting a truncation rank is known as the L-curve, which found its first use in least squares problems (Lawson and Hanson, 1974; Miller, 1970). The L-curve is a plot of the norm of the solution $|\underline{\mathbf{f}}_\kappa|$ against the residual norm $|\mathbf{M}\underline{\mathbf{f}}_\kappa - \underline{\mathbf{g}}|$, parameterised by the truncation rank κ . In Hansen, 1992a and Hansen and O’Leary, 1993 the authors suggest using this curve for helping to select the regularisation parameter for ill-posed problems. They show the L-curve will feature a corner if it is assumed that the DPC is satisfied, the perturbations to the right-hand side are normally distributed with zero mean (white noise) and that the noise to signal ratio is not too large. The authors suggest viewing this plot on a log - log scale. Figure 2.5.6 shows a highly idealised L-curve, made for illustrative purposes, based on the numerical examples found in Hansen, 1990b. As κ is increased the curve is traversed upwards from the $|\mathbf{M}\underline{\mathbf{f}}_\kappa - \underline{\mathbf{g}}|$ axis moving from right to left. The horizontal flat region of the curve is reached once the truncation rank is taken high enough that there is a enough information retained to approximate the solution, however as the truncation rank is further increased the inclusion of smaller singular values eventually leads to magnification of errors on the right-hand side until these begin to dominate the solution and so the norm of the solution rapidly grows with κ , leading to a steep increase in $|\underline{\mathbf{f}}_\kappa|$. It is where the curve transitions from near horizontal to near vertical that is referred to as the L-curve’s corner. Hansen, 1992a and Hansen and O’Leary, 1993 suggest choosing the regularisation parameter such that the solution sits as close to the corner of the L-curve as possible while still being able to satisfy the DPC. This solution balances the need to minimise both the residual norm, such that the regularised solution is a good approximate solution to the matrix equation (2.5.3), and the solution norm, as we seek a smooth solution without large oscillations. The authors go on to show numerical examples of how this method produces good regularised solutions. The use of the L-curve is

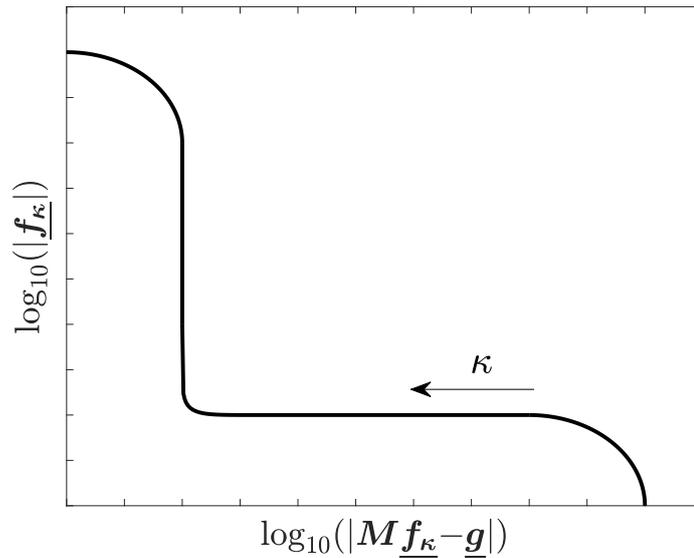


Figure 2.5.6: An example of an idealised L-curve. Selection of κ corresponding to a solution close to the corner on the horizontal leg of the L-curve attempts to obtain a smooth regularised solution \underline{f}_κ while also ensuring that this regularised solution does actually perform well as an approximate solution to (2.5.3).

further suggested for use in solving Fredholm integral equations of the first kind in Hansen, 1992b.

In Figure 2.5.7(a) we consider the DPC for our given example problem with the tanh kernel. Examining this plot it can be seen that if the truncation rank κ is chosen such that $\kappa < 55$ then the regularised problem will satisfy the DPC. For $i \geq 55$ the Fourier coefficients are on the order of machine precision and so these values should not be used to consider the DPC. Figure 2.5.7(b) shows the L-curve calculated for this problem, the red cross marks the truncation rank $\kappa = 35$ used to calculate the solution shown in Figure 2.5.7(d). The L-curve presented here looks quite different to the illustrative example shown previously, we will see shortly that once we have added noise to the problem's input data that this curve will display a clearer corner. Nonetheless, it can be seen that the selection of a truncation rank, for which the DPC is satisfied, close to the end of the horizontal part of the L-curve produces a good solution. Another aide we have used in selecting the truncation rank κ is looking at the relation between the norm of the solution and κ , in Figure 2.5.7(c) these quantities are plotted against

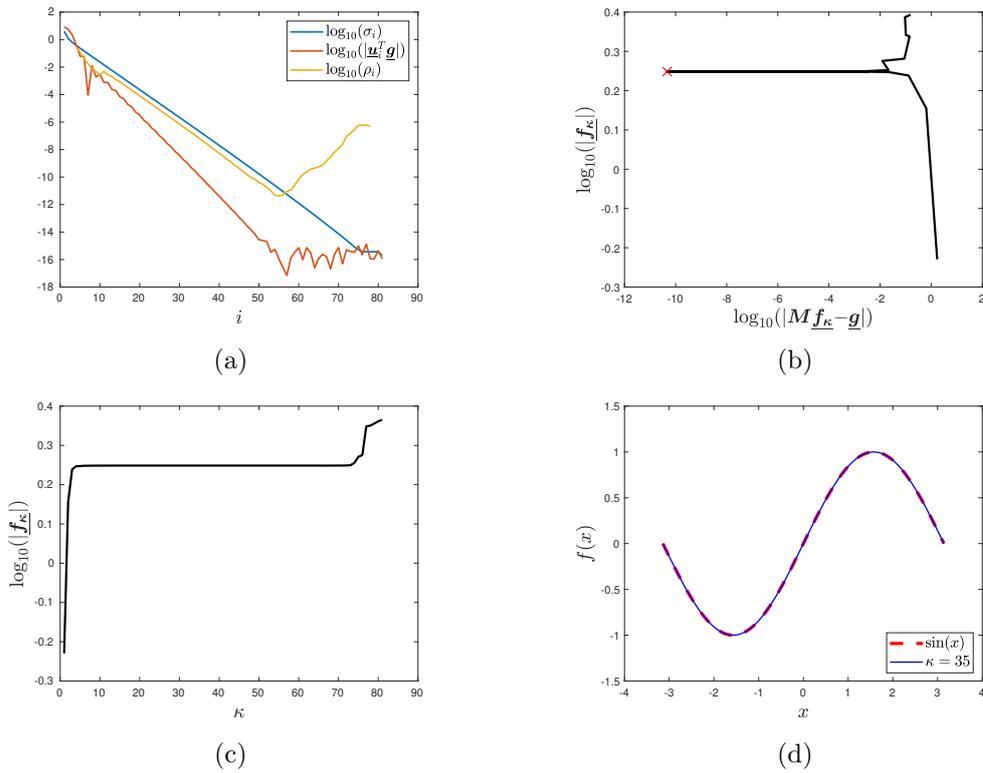


Figure 2.5.7: $L = \pi$, $N = 81$ (a) Picard plot to help consider the satisfaction of the DPC (b) The L-curve for the tanh kernel, the red cross indicates the solution with $\kappa = 35$. (c) A plot of the norm of the solution to the inverse problem against the truncation rank κ . (d) The inverse solution with $\kappa = 35$ (blue) compared to the true solution (red).

one another. Initially, the norm of the solution rapidly rises with the truncation rank until stabilising for an interval of κ , in this region a sufficient number of singular vectors have been included in the solution to well approximate the true solution. After this region the norm then rapidly rises again with κ , this is the point at which numerical errors from the inclusion of the smaller singular values begins to dominate the solution and oscillations on the grid scale begin to appear. For this problem a choice of $\kappa \in (20, 65)$ produces solutions identical to graphical accuracy, a property that will not persist with the addition of perturbation.

We illustrate the use of these tools in selecting the truncation rank κ by considering again the replacement of \mathbf{g} with $\mathbf{g}_\epsilon = \mathbf{g} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon}$ is a vector whose entries are drawn randomly from a normal distribution with zero mean and standard deviation ϵ . Unlike the problem involving the sin kernel this

problem is much more sensitive to this noise and as noted there is no clear truncation rank based on any separation in their values.

In Figure 2.5.8 noise has been added to $\underline{\mathbf{g}}$ with a standard deviation $\epsilon = 1 \times 10^{-8}$. Plotting the norm of $\underline{\mathbf{f}}_\kappa$ against κ the norm appears to be stable in the interval $\kappa \in (10, 35)$ (Figure 2.5.8(c)). The DPC is now only satisfied numerically for $i \leq 27$; we cannot comment on the DPC for greater values of i because the perturbed Fourier coefficients $|\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}_\epsilon|$ have decayed to the order of the added noise ϵ , subsequently they cannot be distinguished numerically from zero and should not be included in our calculations (Figure 2.5.8(a)). The position on the L-curve of the solution corresponding to $\kappa = 27$ was calculated and found to lie close to the L-curve's corner (Figure 2.5.8(b)). Combining the information gleaned from these three plots suggests that choosing $\kappa = 27$ as the truncation rank would allow for a good approximate solution to the underlying problem to be obtained, indeed the results at this rank are shown in Figure 2.5.8(d) with the $\kappa = 27$ indistinguishable from the true solution at graphical accuracy. Note that while in the unperturbed problem a truncation rank of $\kappa = 40$ provided a smooth and accurate solution (see 2.5.5(b)) this is no longer the case in the perturbed problem.

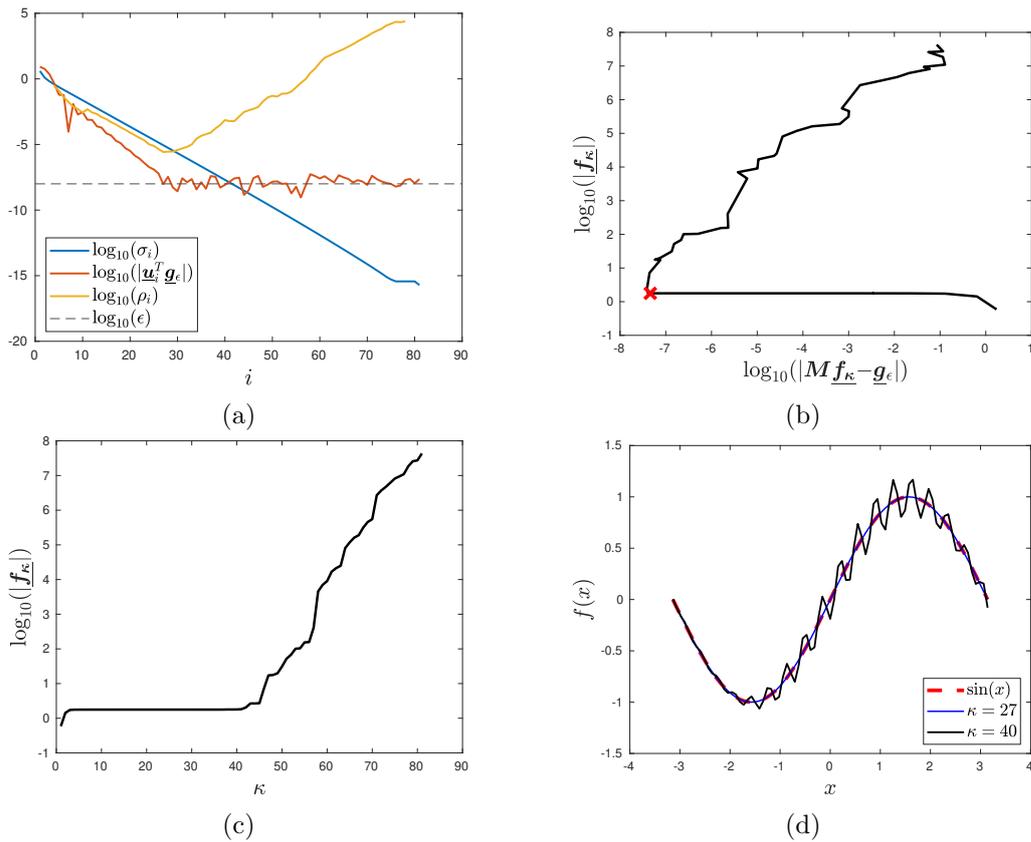


Figure 2.5.8: $L = \pi$, $N = 81$, $\epsilon = 1 \times 10^{-8}$ (a) Picard plot for the noisy problem showing that the DPC is not satisfied for $i > 27$. (b) The L-curve with the red cross indicating the solution with $\kappa = 27$. (c) A plot of the norm of the inverse solution against the truncation rank κ . (d) An inverse solution with a truncation rank chosen to satisfy the DPC $\kappa = 27$ (blue) and one with the truncation rank chosen too high $\kappa = 40$ (black) compared to the true solution (red).

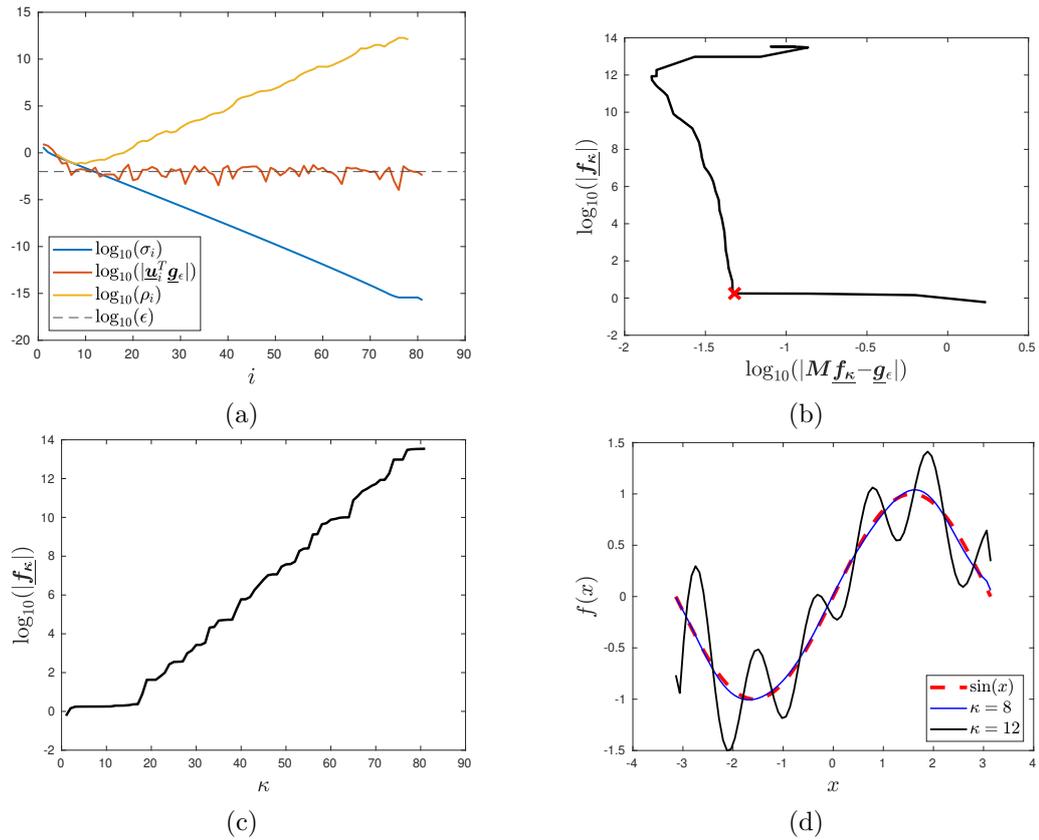


Figure 2.5.9: $L = \pi, N = 81, \epsilon = 1 \times 10^{-2}$ (a) Picard plot for the noisy problem showing that the DPC is satisfied for $i < 8$. (b) The L-curve for this problem with the red cross indicating the solution with $\kappa = 8$. (c) A plot of the norm of the inverse solution against the truncation rank κ . (d) Inverse solutions with $\kappa = 8$ (blue) and $\kappa = 12$ (black) compared to the true solution (red).

3

Forward and Inverse Methods

3.1 Discretisation of Boundary Integral equations

The boundary integral formulation gives the system of equations

$$e^{2\tau_f(\phi)} + \frac{2}{F^2}(y_f(\phi) + P(\phi)) - 1 - \frac{2}{F^2} = 0; \quad (3.1.1)$$

$$\tau_f(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0; \quad (3.1.2)$$

$$\tau_b(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0; \quad (3.1.3)$$

$$\frac{\partial y_f}{\partial \phi} = e^{-\tau_f} \sin(\theta_f); \quad (3.1.4) \quad \frac{\partial y_b}{\partial \phi} = e^{-\tau_b} \sin(\theta_b); \quad (3.1.5)$$

$$\frac{\partial x_f}{\partial \phi} = e^{-\tau_f} \cos(\theta_f); \quad (3.1.6) \quad \frac{\partial x_b}{\partial \phi} = e^{-\tau_b} \cos(\theta_b); \quad (3.1.7)$$

subject to $y_f \rightarrow 1$ and $y_b, \theta_f, \theta_b, P \rightarrow 0$ as $\phi \rightarrow \infty$.

To lay the ground for both approaches we introduce the N_f, N_b equally-spaced mesh-points Φ_i, ϕ_j over the truncated domain $[-L, L]$ for the surface and

topography respectively:

$$\Phi_i = -L + (i - 1)\Delta\Phi, \quad \phi_j = -L + (j - 1)\Delta\phi,$$

for $i = 1, 2, \dots, N_f$ and $j = 1, 2, \dots, N_b$ with $\Delta\Phi = 2L/(N_f - 1)$ and $\Delta\phi = 2L/(N_b - 1)$. We define the discretised variables at the mesh-points:

$$\begin{aligned} \Theta_i &= \theta_f(\Phi_i); & Y_i &= y_f(\Phi_i); & X_i &= x_f(\Phi_i); & P_i &= P(\Phi_i); \\ \theta_j &= \theta_b(\phi_j); & y_j &= y_b(\phi_j); & x_j &= x_b(\phi_j), \end{aligned}$$

for $i = 1, 2, \dots, N_f$ and $j = 1, 2, \dots, N_b$, and the mid-point values:

$$\begin{aligned} \Phi_i^M &= \frac{\Phi_{i+1} + \Phi_i}{2}; & Y_i^M &= \frac{Y_{i+1} + Y_i}{2}; & P_i^M &= \frac{P_{i+1} + P_i}{2}; \\ \Theta_i^M &= \frac{\Theta_{i+1} + \Theta_i}{2}; & \phi_j^M &= \frac{\phi_{j+1} + \phi_j}{2}; & \theta_j^M &= \frac{\theta_{j+1} + \theta_j}{2}, \end{aligned}$$

for $i = 1, 2, \dots, N_f - 1$ and $j = 1, 2, \dots, N_b - 1$.

The discrete forms of (3.1.2) and (3.1.3), evaluated at the midpoints and approximating the integrals by the trapezoidal rule, are used these to define $\mathbb{T}_i^M = \tau_f(\Phi_i^M)$ and $\tau_j^M = \tau_b(\phi_j^M)$ respectively as

$$\begin{aligned} \mathbb{T}_i^M \equiv \tau_f(\Phi_i^M) &= \sum_{k=2}^{N_f} \left(\frac{\Delta\Phi}{2} (G_{f[k-1,i]} \Theta_{k-1} + G_{f[k,i]} \Theta_k) \right) \\ &\quad - \sum_{k=2}^{N_b} \left(\frac{\Delta\phi}{2} (G_{b[k-1,i]} \theta_{k-1} + G_{b[k,i]} \theta_k) \right); \quad (3.1.8) \end{aligned}$$

$$\begin{aligned} \tau_j^M \equiv \tau_b(\phi_j^M) &= \sum_{k=2}^{N_f} \left(\frac{\Delta\Phi}{2} (g_{f[k-1,j]} \Theta_{k-1} + g_{f[k,j]} \Theta_k) \right) \\ &\quad - \sum_{k=2}^{N_b} \left(\frac{\Delta\phi}{2} (g_{b[k-1,j]} \theta_{k-1} + g_{b[k,j]} \theta_k) \right), \quad (3.1.9) \end{aligned}$$

for $i = 1, 2, \dots, N_f - 1$ and $j = 1, 2, \dots, N_b - 1$, where:

$$\begin{aligned} G_{f[k,i]} &= \left(1 - e^{\pi(\Phi_i^M - \Phi_k)}\right)^{-1}; & G_{b[k,i]} &= \left(1 + e^{\pi(\Phi_i^M - \phi_k)}\right)^{-1}; \\ g_{f[k,j]} &= \left(1 + e^{\pi(\phi_j^M - \Phi_k)}\right)^{-1}; & g_{b[k,j]} &= \left(1 - e^{\pi(\phi_j^M - \phi_k)}\right)^{-1}. \end{aligned}$$

Evaluating (3.1.4) and (3.1.5) at the mid-points, and using a central difference approximation for the derivatives, we obtain:

$$Y_i = Y_{i+1} - \Delta\Phi e^{-\tau_i^M} \sin(\Theta_i^M); \quad (3.1.10)$$

$$y_j = y_{j+1} - \Delta\phi e^{-\tau_j^M} \sin(\theta_j^M), \quad (3.1.11)$$

for $i = 1, 2, \dots, N_f - 1$ and $j = 1, 2, \dots, N_b - 1$. Consistent with the assumption of a uniform stream downstream we set $Y_{N_f} = 1$ and $y_{N_b} = 0$. Analogously, from (3.1.6) and (3.1.7) we obtain:

$$X_i = X_{i+1} - \Delta\Phi e^{-\tau_i^M} \cos(\Theta_i^M); \quad (3.1.12)$$

$$x_j = x_{j+1} - \Delta\phi e^{-\tau_j^M} \cos(\theta_j^M), \quad (3.1.13)$$

where we now set $X_{N_f} = x_{N_b} = L$. Finally, (3.1.1) is evaluated at the midpoints yielding

$$e^{2\tau_i^M} + \frac{2}{F^2}(Y_i^M + P_i^M) - 1 - \frac{2}{F^2} = 0, \quad (3.1.14)$$

for $i = 1, 2, \dots, N_f - 1$.

3.2 Forward Problem Methods

Previous work on the forward problem has computed solutions via applying Newton's method to the set of boundary integral equations, for example see Binder, Vanden-Broeck, and Dias, 2005; Binder, Blyth, and McCue, 2013; Vanden-Broeck, 1997. In the present work when considering the forward problem we prescribe the pressure, $P(\phi)$, and the topography $y_b(\phi) = y_T(\phi)$.

We look to solve for $\theta_f(\phi)$ and $\theta_b(\phi)$. During this process the surface $y_f(\phi)$ will be calculated and, once a solution is found, the quantities P, y_b, y_f can be converted to be in terms of the physical variable $x(\phi)$.

Given the $N_f + N_b$ unknowns Θ_i and θ_j we require a system of $N_f + N_b$ equations in order to apply Newton's method. We obtain $N_f - 1$ equations from Bernoulli's equation applied to the midpoints on the surface (3.1.14). Writing the known values on the topography as $y_{T,j} = y_T(\phi_j)$ a further $N_b - 1$ equations are obtained by comparing these quantities to the values for y_j calculated by (3.1.11) as

$$y_{T,j} - y_j = 0 \quad (3.2.1)$$

for $j = 1, \dots, N_b - 1$. Two further equations come from applying boundary conditions to θ_f and θ_N . If the flow is supercritical, with $F > 1$, then the condition applied is $\theta_1 = \Theta_1 = 0$. When the flow is subcritical, with $F < 1$, we instead take $\theta_{N_b} = \Theta_{N_f} = 0$. We do not compute forward solutions to critical flow ($F = 1$) in the present work, this will be discussed in Section 3.2.5.

Any further discussion of the forward problem will assume an equal number of points to be taken on the surface and topography with $N_f = N_b = N$. The standard forward problem in this work then is taken to mean solving by Newton's method the $2N$ equations:

$$e^{2T_i^M} + \frac{2}{F^2}(Y_i^M + P_i^M) - 1 - \frac{2}{F^2} = 0; \quad (3.2.2)$$

$$y_{T,i} - y_i = 0; \quad (3.2.3)$$

$$\begin{cases} \theta_1 = \Theta_1 = 0 & \text{if } F > 1 \\ \theta_N = \Theta_N = 0 & \text{if } F < 1 \end{cases} \quad (3.2.4)$$

with $i = 1, \dots, N - 1$, for the unknowns θ_j and Θ_j with $j = 1, \dots, N$, where we have the known values $y_{T,j}$ and P_j . The quantities Y_j , y_j and T_i^M are updated during each iteration of Newton's method by way of (3.1.8 - 3.1.11).

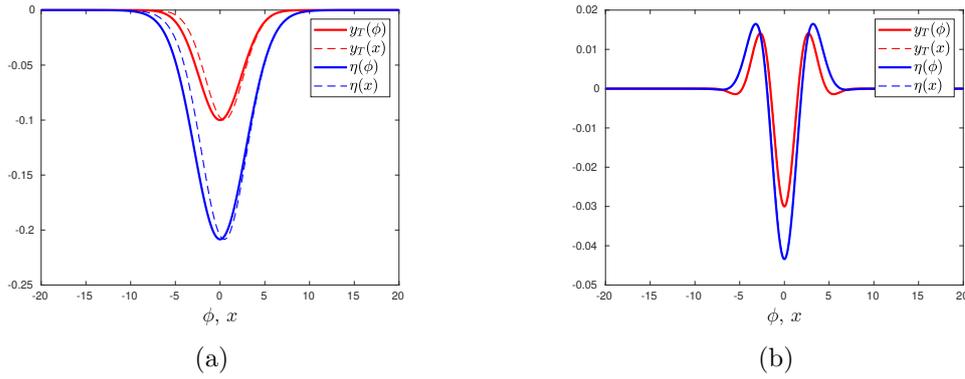


Figure 3.2.1: $L = 20, N = 641, F = 1.2, b = 0.3$. Uniform stream perturbation solutions to the forward problem for supercritical flow. These solutions $\eta = y_f - 1$ follow approximately the shape of the topography y_T . The solid lines are plotted as a function of the potential ϕ and the dashed lines are plotted as a function of the physical variable x . (a) Gaussian topography of form (3.2.5) with $a = -0.1$. (b) Decaying cosine topography of form (3.2.6) with $a = -0.03$.

3.2.1 Solutions as a perturbation to the uniform stream

The most basic, and easiest to calculate, solutions to the forward problem are those corresponding to a perturbation to the uniform stream and in this section we present some examples of these types of solution for illustrative purposes before turning to other types of solution. For these examples we shall take either a Gaussian topography,

$$y_T = ae^{-(b\phi)^2}, \quad (3.2.5)$$

or a topography that results from taking a Gaussian and then multiplying it by a cosine,

$$y_T = ae^{-(b\phi)^2} \cos(\phi). \quad (3.2.6)$$

For supercritical flow, with $F > 1$, the free surface typically follows the rough form of the prescribed topography y_T , see Figure 3.2.1. This is not the case for subcritical flow where generally waves are found to be on the upstream side of the surface, see Figure 3.2.2. In order to have the radiation condition satisfied we must consider the flow in this example to be from right to left.

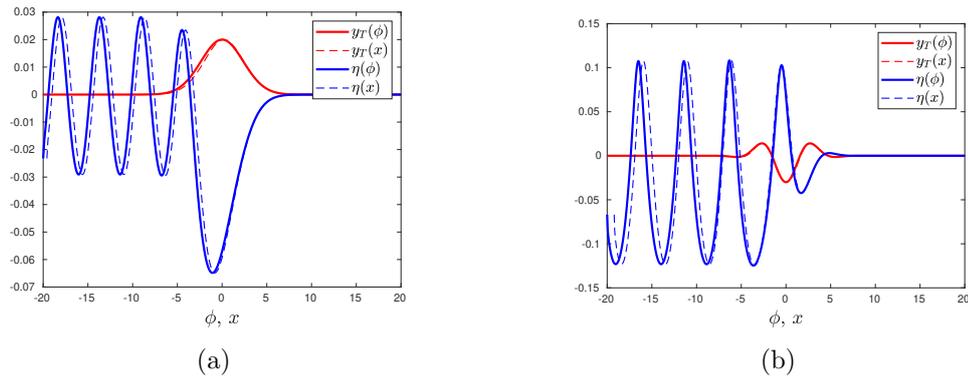


Figure 3.2.2: $L = 20$, $N = 641$, $F = 1.2$, $b = 0.3$. Uniform stream perturbation solutions to the forward problem for subcritical flow. The solid lines are plotted as a function of the potential ϕ and the dashed lines are plotted as a function of the physical variable x . (a) Gaussian topography of form (3.2.5) with $a = -0.1$. (b) Decaying cosine topography of form (3.2.6) with $a = -0.03$.

3.2.2 Solutions as a perturbation to the solitary wave

Vanden-Broeck, 1987 found that in the fully nonlinear problem with a fixed non-zero forcing there is a value F^* below which no steady solutions exist. For values above F^* (but below the value for which the steepest limiting wave configuration is reached) two solutions exist, one an analogue to the uniform stream and one the solitary wave in the unforced problem (see Figure 3.2.3). Figure 3.2.4(a) shows an example of the uniform stream solution and the unforced solitary wave in the fully nonlinear problem, Figure 3.2.4(b) shows their respective analogues in the forced problem. In order to compute solitary wave type solutions on the upper branch of solutions we first computed a solution on the lower branch with the desired forcing before using numerical continuation to travel along the lower branch and around the corner until the desired point on the upper branch was reached.

3.2.3 Hydraulic falls and jumps

Hydraulic falls occur when the flow transitions from being subcritical upstream to supercritical downstream. If this flow is reversed it is instead referred to as a hydraulic jump. A generalised hydraulic fall has waves on its upstream surface

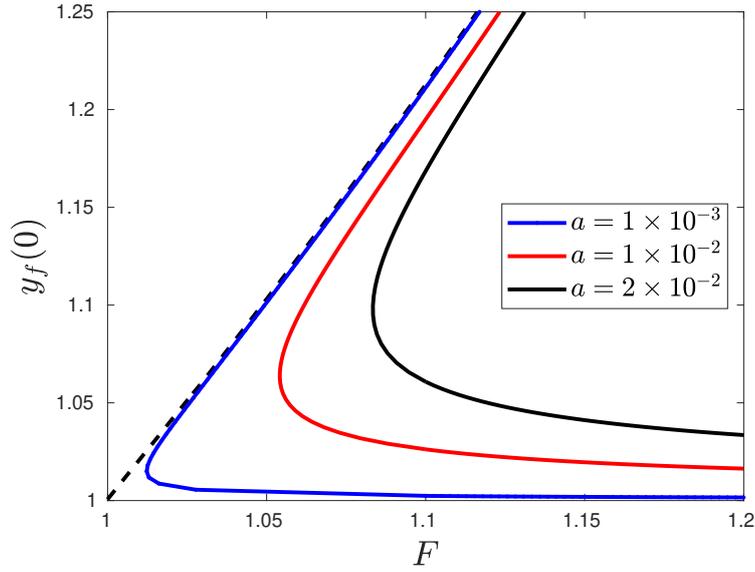


Figure 3.2.3: Bifurcation curve for forward solutions to forced flow showing the solitary wave and uniform stream branches for multiple amplitudes of forcing with $L = 20$, $N = 641$, $b = 1$.

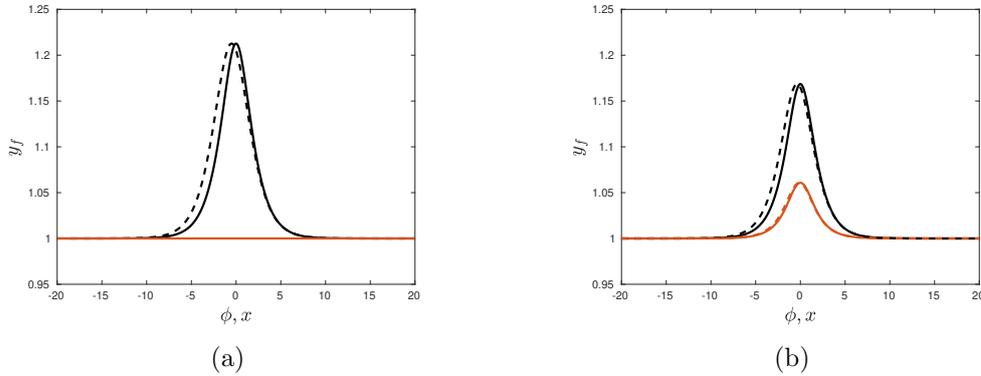


Figure 3.2.4: $L = 20$, $N = 641$, $F = 1.1$, $b = 1$. Nonuniqueness in the forward problem over a topography of the form (3.2.5) results in free surface solutions that are like a perturbation to the uniform stream (orange) or a perturbation to the solitary wave (black). Solutions are plotted both as $y_f(\phi)$ (solid line) and as $y_f(x)$ (dashed line). (a) Unforced flow with $a = 0$. (b) Forced flow with $a = 2 \times 10^{-2}$.

while a hydraulic fall is wave-free. We shall restrict our attention to hydraulic falls; in order to eliminate the upstream waves we must satisfy

$$F^2 = \frac{2(y_f(-\infty))^2}{1 + y_f(-\infty)}.$$

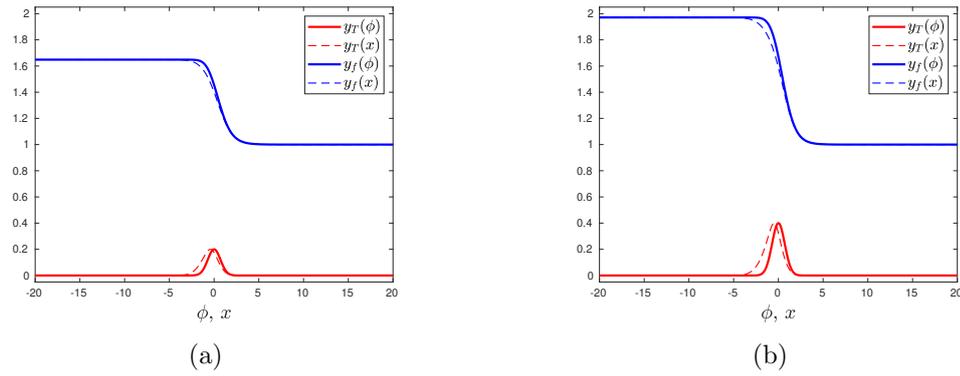


Figure 3.2.5: $L = 20, N = 721, b = 1$. Examples of hydraulic fall solutions to flow past a Gaussian topography of the form (3.2.5). The solid lines are plotted as a function of the potential ϕ and the dashed lines are plotted as a function of the physical variable x . (a) $a = 0.2$. The Froude number was calculated as $F \approx 1.4326$. (b) $a = 0.4$. The Froude number was calculated as $F \approx 1.4326$.

The above condition is derived from applying the conservation of mass and Bernoulli's equation to the uniform upstream and downstream flow. More information on this condition and generalised hydraulic falls can be found in Binder, Vanden-Broeck, and Dias, 2005; Dias and Vanden-Broeck, 2004; Dias and Vanden-Broeck, 1989, 2002.

To compute these solutions by Newton's method we allow the Froude number F to come as part of the solution. We will have for this problem $2N + 1$ unknowns; θ_i and Θ_i for $i = 1, \dots, N$ and now F , so we will need $2N + 1$ equations. Tweaking the standard forward problem equations for Newton's method we obtain $2N - 2$ equations from (3.2.2) and (3.2.3), then supplement these with the two boundary conditions $\theta_1 = \Theta_1 = 0$ and the discrete analogue to the above relation between the Froude number and the upstream uniform depth Y_1 given by

$$F^2 = \frac{2(y_f(-\infty))^2}{1 + y_f(-\infty)}$$

3.2.4 Trapped wave solutions

As seen in Figure 3.2.2(a) flow past a Gaussian topography produced a solution with a train of waves on its upstream surface. One natural question to ask is can this train of waves be terminated at some point such that there is a uniform surface both upstream and downstream. Forbes and Schwartz, 1982 found that this could occur for flow past singular semi-elliptical topographical disturbances of specific dimensions. Following their work on calculating generalised hydraulic falls (Dias and Vanden-Broeck, 2002) it was shown by Dias and Vanden-Broeck, 2004 that by positioning a second topographical feature the waves appearing on top of the hydraulic fall could be confined to a finite interval with the far field surface being of uniform flow at each end. Binder, Vanden-Broeck, and Dias, 2005 subsequently calculated trapped wave solutions with supercritical flow at each end where a hydraulic jump was matched to a hydraulic fall, each induced by a triangular feature on the topography, where waves appeared only atop the locally subcritical region between the obstacles. Trapped wave solutions were also calculated by the authors in subcritical flow past two triangles and then, in a later paper, Binder, Dias, and Vanden-Broeck, 2008 calculated solutions with trapped waves past an inclined gate by using two pressure distributions to terminate the wave train appearing on each side of the gate. Holmes, Hocking, Forbes, and Baillard, 2013 then studied subcritical flow with trapped waves over a symmetric topography featuring two Gaussians, showing that solutions could be found for both elevations and depressions in the topography and that a discrete set of these solutions could be found for different values of the separation between the two topographical features (example surfaces shown in Figure 3.2.6). It is solutions of this type that we chose to compute to later use as input in the inverse problem; in order to calculate these solutions we prescribe the topography

$$y_T = a \left(e^{-b(\phi-c)^2} + e^{-b(\phi+c)^2} \right), \quad (3.2.7)$$

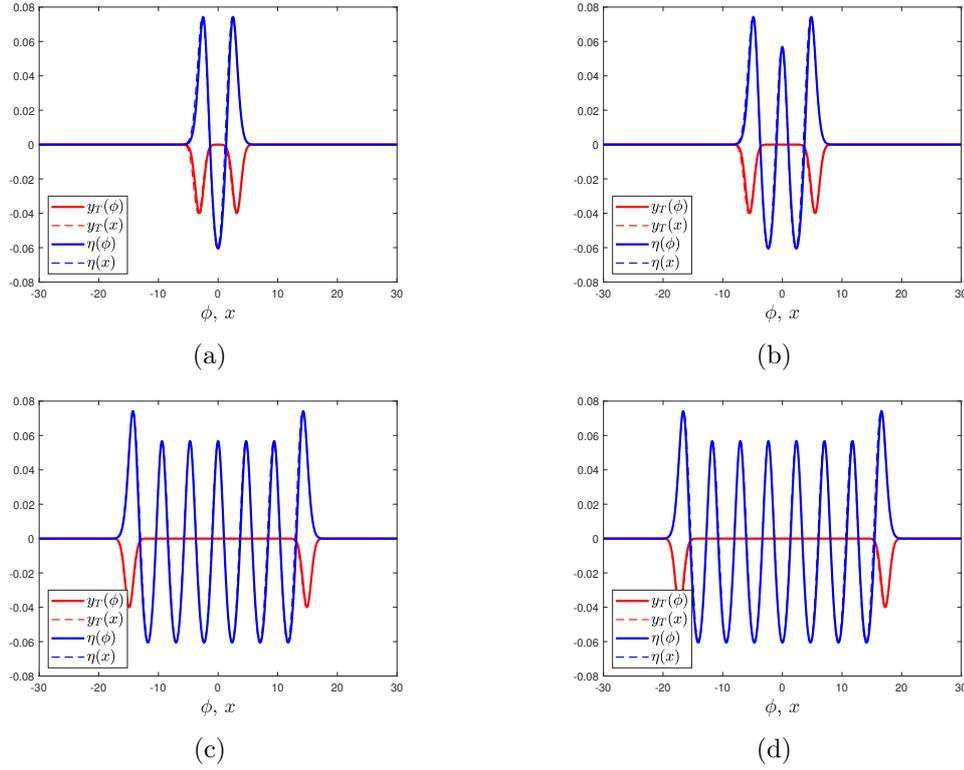


Figure 3.2.6: $L = 30, N = 641, F = 0.8, b = 1, a = -0.04$. Examples of trapped wave free surface solutions (blue) in subcritical flow over a two bumped topography (red) given by (3.2.7) where the separation comes as part of the solution. The solid lines are plotted as a function of the potential ϕ and the dashed lines are plotted as a function of the physical variable x . Different profiles are obtained for different initial guesses for c . (a) Solution with two wave peaks with $c \approx 3.1334$. (b) Solution with three wave peaks with $c \approx 5.4871$. (c) Solution with seven wave peaks with $c \approx 14.9021$. (d) Solution with eight wave peaks with $c \approx 17.2559$.

where c , the parameter used to control the separation of the two Gaussians, is left unknown and allowed to come as part of the solution. With $2N + 1$ unknowns we obtain $2N - 2$ equations from (3.2.2) and (3.2.3), keeping in mind that after each Newton iteration we must use the new approximation for c to re-evaluate y_T in (3.2.7). The final three equations needed come from applying the boundary conditions $\theta_1 = \Theta_1 = \Theta_N = 0$ to correspond to uniform flow both upstream and downstream. By using different initial guesses for c solutions can be found with a different number of trapped waves on the surface; in Figure 3.2.6 examples are shown of forward solutions with varying numbers of trapped waves and their associated underlying topographies.

3.2.5 Forward solutions in critical flow

The steady fKdV, integrated once, is given by

$$\eta_{xx} + \frac{9}{2}\eta^2 - 6(F-1)\eta = -3f(x). \quad (3.2.8)$$

Keeler, Binder, and Blyth, 2017 studied the decay of the free surface for forced critical flows by asymptotic analysis of a scaled fKdV equation. They considered only even free surface solutions to flow over even an forcing term. We shall begin with a modified form of this scaled equation

$$u_{\xi\xi} + u^2 - \hat{F}u = \alpha s(\xi) \quad (3.2.9)$$

subject to $u, u_\xi, u_{\xi\xi} \rightarrow 0$ as $|\xi| \rightarrow \infty$, where u represents the free surface; ξ the horizontal spatial variable; \hat{F} is a scaled measured of the Froude number, equal exactly to zero for critical flow; $\alpha \geq 0$ is a constant derived from the scaling and $s(\xi)$ is the forcing term. This equation differs from that used in Keeler, Binder, and Blyth, 2017 only in that we have retained the third term, which vanishes in the case of critical flow. As only even solutions are sought the problem is solved on $\xi \in [0, \infty)$ before being reflected about $\xi = 0$.

We consider the far field decay of the surface for flow over the Gaussian topography $s(\xi) = e^{-\xi^2}$. For a flow that is not critical as we approach the far field the first and third terms of (3.2.9) will balance one another as the the second and fourth terms rapidly become of negligible size. This leaves the approximate equation

$$u_{\xi\xi} - \hat{F}u = 0. \quad (3.2.10)$$

This is a linear O.D.E. with general solution

$$u = Ae^{\sqrt{\hat{F}}\xi} + Be^{-\sqrt{\hat{F}}\xi}$$

where A, B are arbitrary constants. Applying the condition $u \rightarrow 0$ as $\xi \rightarrow \infty$

requires that $A = 0$, giving the solution

$$u = Be^{-\sqrt{F}\xi} \quad (3.2.11)$$

from which it can be clearly seen that the solution in the far field decays exponentially.

Considering instead a critical flow the governing equation reduces to

$$u_{\xi\xi} + u^2 = \alpha e^{-\xi^2}. \quad (3.2.12)$$

There are two possible balances between the three possible combinations of the terms in (3.2.12); assuming that the first term, that containing a second derivative, is of negligible size compared to the other two terms leads to a contradiction. Assuming as such means

$$u^2 \sim \alpha e^{-\xi^2} \implies u \sim \pm\sqrt{\alpha}e^{-\frac{1}{2}\xi^2} \quad (3.2.13)$$

which, upon differentiating twice, gives

$$u_{\xi\xi} \sim \pm\xi^2\sqrt{\alpha}e^{-\frac{1}{2}\xi^2} \quad (3.2.14)$$

contradicting the original assumption that the second derivative was negligible compared to the other terms.

If the balance in (3.2.12) occurs between the second derivative and the forcing on the right-hand side we have instead

$$u_{\xi\xi} \sim \alpha e^{-\xi^2}. \quad (3.2.15)$$

Applying the result of (B.0.13) and then (B.0.14) we can then state immediately

that

$$u_\xi \sim \frac{\alpha}{2\xi} e^{-\xi^2} \quad \text{and} \quad u \sim \frac{\alpha}{4\xi^2} e^{-\xi^2}, \quad (3.2.16)$$

showing that decay of the free surface in the far field will be super-exponential.

Finally, if the balance in (3.2.12) occurs instead between the two terms on the left-hand side we are left with the approximate equation

$$u_{\xi\xi} + u^2 = 0. \quad (3.2.17)$$

By first multiplying by u_ξ this can be integrated to

$$\frac{1}{2}u_\xi^2 = -\frac{1}{3}u^3 \quad (3.2.18)$$

where the constant has been set to zero to satisfy the far field conditions. Taking the square root, with $u < 0$, we are left with a separable first order ODE which can be integrated and rearranged for u to find the large ξ behaviour as

$$u(\xi) \sim \frac{-4}{\left(A \pm \xi \sqrt{\frac{2}{3}}\right)^2} \quad (3.2.19)$$

where A is a constant of integration. It can be seen that as $\xi \rightarrow \infty$ we have $u \sim \frac{-6}{\xi^2}$ so u is in this case decaying algebraically like $\frac{1}{\xi^2}$. Keeler, Binder, and Blyth, 2017 observed numerically that all calculated solutions in the (α, u) solution space displayed algebraic decay except for those solutions corresponding exactly to the termination point of a solution branch, these solutions instead displayed super-exponential decay. The authors found numerically for this forcing a sequence of nested solution branches and characterised these by the nature of the turning point occurring at $\xi = 0$ and the number of local maxima/minima on the free surface. The nature of the turning point can be investigated by evaluating (3.2.12) at $\xi = 0$, giving

$$u_{\xi\xi}(0) = \alpha - [u(0)]^2; \quad (3.2.20)$$

the sign of $u_{\xi\xi}(0)$ will depend on the relative sizes of u_0 and α meaning that $u(0)$ could be either a local maximum or minimum. The authors found, that in terms of the increasing value of α below which solutions branches cease to exist, solution branches alternated between having maxima or minima at $\xi = 0$.

The (scaled) fKdV allows the inverse problem for the forcing to quickly be solved. For critical flow the governing equation is

$$u_{\xi\xi} + u^2 = \alpha s(\xi). \quad (3.2.21)$$

We will now fix a free surface and value of α before calculating the forcing via the inverse problem. This obtained inverse forcing will then be used as input to the forward problem as α is allowed to vary in order to study the effects on the decay of the resulting free surface. Setting $\alpha = 1$ and prescribing the free surface as $u(\xi) = e^{-\xi^2}$ the forcing is given immediately by substitution as $s(\xi) = (4\xi^2 - 2)e^{-\xi^2} + e^{-2\xi^2}$. Inserting this solution back into the forward problem, allowing α to vary, yields the problem to be solved as

$$u_{\xi\xi} + u^2 = \alpha \left((4\xi^2 - 2)e^{-\xi^2} + e^{-2\xi^2} \right), \quad (3.2.22)$$

subject to $u, u_\xi, u_{\xi\xi} \rightarrow 0$ as $|\xi| \rightarrow \infty$. An immediate difference in solutions to this problem and the problem of solving (3.2.12) can be seen by considering the behaviour of the free surface at $\xi = 0$. Evaluating we find

$$u_{\xi\xi}(0) = -\alpha - [u(0)]^2 \leq 0; \quad (3.2.23)$$

for this forcing it is not possible for $\xi = 0$ to be a local minimum. While we have not attempted to locate any other solution branches for this forcing they are expected to exist but they will only have maxima at $\xi = 0$.

There are two possible balances between the terms (3.2.22): one between $u_{\xi\xi}$ and u^2 ; and one between $u_{\xi\xi}$ and $\alpha s(\xi)$, as was the case for the Gaussian forcing.

If the balance in the far field in (3.2.22) occurs between the second derivative and the forcing term then we have that

$$u_{\xi\xi} \sim \alpha \left((4\xi^2 - 2)e^{-\xi^2} + e^{-2\xi^2} \right), \quad (3.2.24)$$

which can be integrated to

$$u_{\xi} \sim \alpha \left(-2\xi e^{-\xi^2} + \int e^{-2\xi^2} d\xi \right). \quad (3.2.25)$$

Making the substitution $t = \xi\sqrt{2}$ into the second term of the above integral and applying the result of (B.0.13) we have that

$$\int e^{-2\xi^2} d\xi = \frac{1}{\sqrt{2}} \int e^{-t^2} dt \sim \frac{-1}{4\xi} e^{-2\xi^2}. \quad (3.2.26)$$

The dominant behaviour then of (3.2.25) is

$$u_{\xi} \sim \alpha(-2\xi)e^{-\xi^2}, \quad (3.2.27)$$

which is easily integrated to give

$$u \sim \alpha e^{-\xi^2}. \quad (3.2.28)$$

The decay of the free surface will be super-exponential when the balance occurs between these terms.

If the balance in the far field in (3.2.22) occurs instead between the second derivative and the squared term then the problem to be solved is identical to that of the Gaussian forcing, see (3.2.17-3.2.19) for which solutions were found to decay algebraically like $\frac{1}{\xi^2}$.

Following the method of Keeler, Binder, and Blyth, 2017 solutions to (3.2.22) were calculated first on $\xi \in [0, \infty)$, applying $u, u_{\xi}, u_{\xi\xi} \rightarrow 0$ as $\xi \rightarrow \infty$, before reflecting the solution about the u -axis. The authors then recast the problem as

a system of first order ODES, for the present forcing this system is:

$$u_\xi = v; \quad (3.2.29)$$

$$v_\xi = \alpha \left((4\xi^2 - 2)e^{-\xi^2} + e^{-2\xi^2} \right) - u^2. \quad (3.2.30)$$

For a given u_0 the system is treated as an initial value problem and integrated forwards from $\xi = 0$ with initial values:

$$u(0) = u_0; \quad (3.2.31) \quad v(0) = u_\xi(0) = 0. \quad (3.2.32)$$

The value u_0 is then refined such that the trajectory in the phase space (u, u_ξ) approaches the origin as $\xi \rightarrow \infty$ as the trajectory connecting $(u_0, 0)$ to $(0, 0)$ corresponds to the solution $u(\xi)$ that satisfies the original boundary value problem. As noted in Keeler, Binder, and Blyth, 2017 the trajectory only enters the origin as $\xi \rightarrow \infty$ and so in computational practice u_0 is refined such that the trajectory approaching the origin would require integrating to larger and larger values of ξ . Further details can be found in Keeler, Blyth, and J. R. King, 2021. Some examples of the trajectories in the (u, u_ξ) plane of solutions to (3.2.22) are shown for various values of α in Figure 3.2.7(c) with the corresponding surface profiles and forcing terms plotted in Figure 3.2.7(a) and Figure 3.2.7(b) respectively. The red dotted lines in Figure 3.2.7(c) are the curves

$$u_\xi = \pm \sqrt{-\frac{2}{3}u^3}. \quad (3.2.33)$$

It can be seen that as $\xi \rightarrow \infty$ the trajectories for $\alpha > 1$ approach the origin, latching to the curve $u_\xi = \sqrt{-\frac{2}{3}u^3}$ along which we know that u decays algebraically. However, by construction the solution corresponding to $\alpha = 1$ experiences super-exponential decay. This behaviour suggests that $\alpha = 1$ may be a termination point for this solution branch as this matches the observations of Keeler, Binder, and Blyth, 2017 on the termination of solution branches for

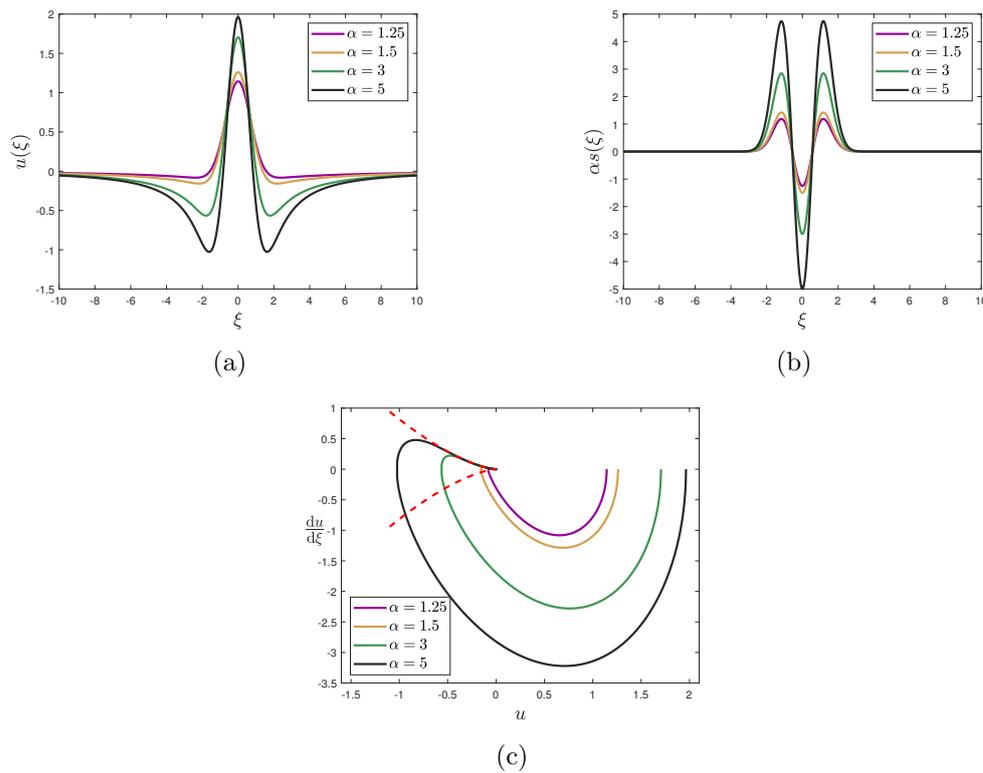


Figure 3.2.7: (a) Solutions for the free surface for different amplitudes of forcing. (b) Profiles of the forcing for different amplitudes. (c) Phase portrait for the solutions for u showing how they approach the termination point with algebraic decay for all solutions except that with $\alpha = 1$.

the forcing they considered. This is further evidenced by the fact that we have been unable to calculate any solutions with $\alpha < 1$. Solutions corresponding to branch termination points are difficult to find numerically, the construction of the present problem has circumvented this issue as the exact solution for $\alpha = 1$ is, having been prescribed originally, known in advance. Keeler, Binder, and Blyth, 2017 showed that solutions to the fully nonlinear problem over Gaussian forcing experience the same form of algebraic decay as $\xi \rightarrow \infty$ as the solutions to the weakly nonlinear problem. The authors write that they expect the same shift in decay behaviour to represent the termination point of the fully nonlinear solution branches.

For the purposes of validating solutions found to the fully nonlinear inverse problem we have used the calculated topography as input to the forward problem. This has not been possible for the inverse solutions found at $F = 1$

and the results of this section help to illuminate the difficulties experienced; forward solutions corresponding to the termination point of a solution branch are hard to find numerically, any other solutions experience algebraic decay and as such require increasing the size of the truncated domain or careful tweaking of the boundary conditions.

3.3 Inverse Problem Methods

Sections of the remainder of this chapter have been submitted as part of a manuscript, which is under review, to the Journal of Fluid Mechanics (Robbins, Blyth, Maclean, and Binder, 2023).

Binder, Blyth, and McCue (2013) and Tam, Yu, Kelso, and Binder (2015) found that using Newton's method for the inverse topography problem leads to serious convergence issues. While a sufficiently coarse grid yields numerical output that appears smooth, as the grid resolution is increased a numerical instability occurs that manifests as irregular grid-scale sawtooth oscillations on the topography profile. In Figure 3.3.1 we show results for both the forward and the inverse problem for the prescribed Gaussian topography,

$$y_T(\phi) = 0.05e^{-\phi^2}, \quad (3.3.1)$$

and in the absence of a surface pressure. The calculations are performed for three different grid resolutions with $N = \{181, 359, 363\}$ grid points. The surface profiles for the forward problem shown in Figure 3.3.1(a) are in good agreement for the three chosen resolutions. However, if these same surface profiles are used as input for the inverse topography problem, we see in Figure 3.3.1b that y_b is not recovered in the higher resolution calculations, which exhibit sawtooth oscillations. We note that this failure is not an artefact of using the numerical output from the forward problem as the input to the inverse problem; the same issue occurs when $\theta_f(\phi)$ is prescribed directly in the inverse problem. The inverse

problem therefore deserves special attention.

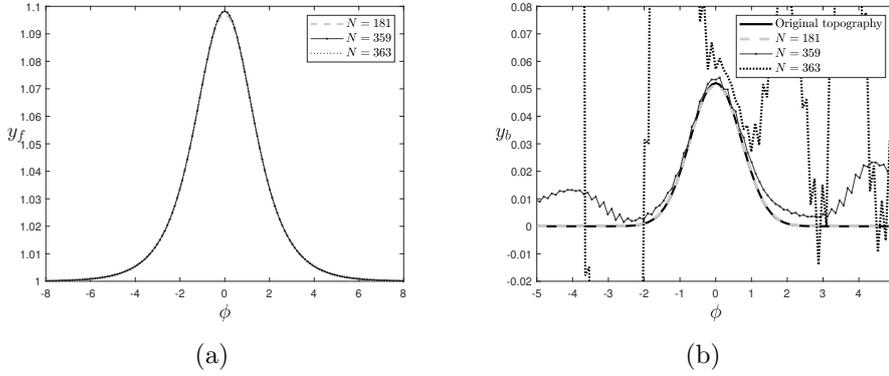


Figure 3.3.1: Forward and inverse numerical solutions obtained by Newton's method, $F = 1.2$. (a) Solutions to the forward problem for the prescribed topography y_T given (3.3.1). (b) Prescribed topography, y_T , compared to Tam, Yu, Kelso, and Binder (2015) inverse Newton's method, y_b .

3.3.1 Inverse Pressure Problem

The inverse pressure problem is to solve for the pressure forcing given a fixed free surface and topography, provided by prescribing θ_f and θ_b respectively. In our results we will consider problems with $\theta_b = 0$, i.e. with uniform topography, however we retain θ_b in our formulae for generality of the result.

First, we note that, with θ_f and θ_b known, (3.1.2) allows for us to write an exact expression for τ_f in terms of known variables,

$$\tau_f(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0.$$

This expression for τ_f may then be substituted into (3.1.4), which can then be integrated to obtain

$$y_f(\phi) = \int_{\phi}^{\infty} e^{-\tau_f(\phi')} \sin(\theta_f(\phi')) d(\phi'),$$

an exact expression for y_f . By now rearranging (3.1.1) to make P the subject we

obtain

$$P(\phi) = 1 - y_f(\phi) - \frac{F^2}{2} (e^{2\tau_f(\phi)} - 1), \quad (3.3.2)$$

an expression for $P(\phi)$ given exactly and entirely in terms of known variables.

Numerical results for the pressure are obtained analogously. We first calculate T_i^M by (3.1.8), then using (3.1.10) to find Y_i , before calculating the midpoint values Y_i^M and substituting these values into the rearranged form of (3.1.14) given by

$$P_i^M = 1 - Y_i^M - \frac{F^2}{2} (e^{2T_i^M} - 1). \quad (3.3.3)$$

to give the pressure at the midpoints. Finally the pressure P_i is found by combining the rearranged expression for midpoint values, $P_{i+1} = 2P_i^M - P_i$ for $i = 1, 2, \dots, N_f - 1$, with the decay condition in the far-field $P_{N_f} = 0$.

3.3.2 Inverse topography problem as a linear problem

We consider the inverse problem for topography with the surface fixed by prescribing y_f , rather than θ_f , and a fixed pressure forcing P . We begin by rearranging (3.1.1) for τ_f to obtain an expression in terms of known quantities as

$$\tau_f(\phi) = \frac{1}{2} \ln \left(1 - \frac{2}{F^2} (y_f(\phi) + P(\phi) - 1) \right). \quad (3.3.4)$$

With τ_f known we can now take the derivative of y_f and rearrange (3.1.4) for θ_f as

$$\theta_f = \arcsin \left(e^{\tau_f} \frac{\partial y_f}{\partial \phi} \right). \quad (3.3.5)$$

As such by prescribing y_f we can simultaneously know θ_f , this was not possible when instead prescribing θ_f and allows for us now to cast the problem in linear terms. By equating (3.3.4) with (3.1.2) and rearranging we can now state the inverse problem as solving

$$\int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0 = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \frac{1}{2} \ln \left(1 - \frac{2}{F^2} (y_f(\phi) + P(\phi) - 1) \right)$$

for the unknown $\theta_b(\phi)$, where everything on the right-hand side is known. By introducing the notation

$$K(\phi, \phi_0) = \frac{1}{1 + e^{\pi(\phi - \phi_0)}} = \frac{1}{2} \left(1 - \tanh \frac{1}{2} \pi (\phi - \phi_0) \right), \quad (3.3.6)$$

and

$$b(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \frac{1}{2} \ln \left(1 - \frac{2}{F^2} (y_f(\phi) + P(\phi) - 1) \right) \quad (3.3.7)$$

the inverse problem is written more concisely as solving

$$\int_{-\infty}^{\infty} \theta_b(\phi_0) K(\phi, \phi_0) d\phi_0 = b(\phi). \quad (3.3.8)$$

Now, as stated in (3.3.8), the inverse problem for topography requires solving a linear singular Fredholm equation of the first kind. Fredholm equations of the first kind are ill-posed problems (e.g. Groetsch, 2007; Phillips, 1962).

Now, let us consider explicitly the discrete system formed by discretisation of the inverse problem. The discrete form of (3.3.4), evaluated on Φ_i , is used to define

$$T_i \equiv \tau_f(\Phi_i) = \frac{1}{2} \ln \left(1 - \frac{2}{F^2} (Y_i + P_i - 1) \right), \quad (3.3.9)$$

for $i = 1, 2, \dots, N_f$, which is calculated from the known surface values Y_i and P_i . We then obtain the values Θ_i , inserting the values Y_i and T_i into (3.3.5) and using a central difference for the derivative, yielding

$$\Theta_i = \arcsin \left(e^{T_i} \left(\frac{Y_{i+1} - Y_{i-1}}{2\Delta\Phi} \right) \right) \quad (3.3.10)$$

for $i = 2, 3, \dots, N_f - 1$ and we set $\Theta_1 = \Theta_{N_f} = 0$. Now evaluating at the midpoints, and approximating the integrals by the trapezium rule, the discretised form of

the integral equation (3.3.8) is given by

$$\begin{aligned} \sum_{k=2}^{N_b} \left(\frac{\Delta\phi}{2} (G_{b[k-1,i]} \theta_{k-1} + G_{b[k,i]} \theta_k) \right) \\ = \sum_{k=2}^{N_f} \left(\frac{\Delta\Phi}{2} (G_{f[k-1,i]} \Theta_{k-1} + G_{f[k,i]} \Theta_k) \right) \\ - \frac{1}{2} \ln \left(1 - \frac{2}{F^2} (Y_i^M + P_i^M - 1) \right) \end{aligned} \quad (3.3.11)$$

for $i = 1, 2, \dots, N_f - 1$, where:

$$G_{f[k,i]} = \left(1 - e^{\pi(\Phi_i^M - \Phi_k)} \right)^{-1}; \quad G_{b[k,i]} = \left(1 + e^{\pi(\Phi_i^M - \phi_k)} \right)^{-1}.$$

Now the inverse problem for the unknowns θ_j may now be written more succinctly as the linear matrix equation

$$\mathbf{M}\boldsymbol{\theta} = \mathbf{b}, \quad (3.3.12)$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_{N_b})^T$ is the vector of unknowns, $\mathbf{b} = (b_1, b_2, \dots, b_{N_f})^T$ is a vector of the known free-surface values

$$b_i = \sum_{k=2}^{N_f} \left(\frac{\Delta\Phi}{2} (G_{f[k-1,i]} \Theta_{k-1} + G_{f[k,i]} \Theta_k) \right) - \frac{1}{2} \ln \left(1 - \frac{2}{F^2} (Y_i^M + P_i^M - 1) \right) \quad (3.3.13)$$

for $i = 1, 2, \dots, N_f - 1$, and \mathbf{M} is an $(N_f \times N_b)$ matrix with the known elements $m_{i,j}$ given by:

$$m_{i,1} = \frac{\Delta\phi}{2} G_{b[1,i]}; \quad m_{i,N_b} = \frac{\Delta\phi}{2} G_{b[N_b,i]}; \quad m_{i,j} = \Delta\phi G_{b[j,i]},$$

for $i = 1, 2, \dots, N_f - 1$ and $j = 2, 3, \dots, N_b - 1$. The final row of the matrix equation is set to enforce a boundary condition. An obvious choice might be to set $\theta_{N_b} = 0$; however, this condition is already accounted for by the reliance of the conformal mapping on uniform flow far downstream. As such, while it is not in practice necessary to do so, we instead set $m_{N_f,1} = 1$ and $b_{N_f} = 0$ to enforce the condition that $\theta_1 = 0$. We found that this greatly improves the conditioning of the system, as shown in Figure 3.3.2.

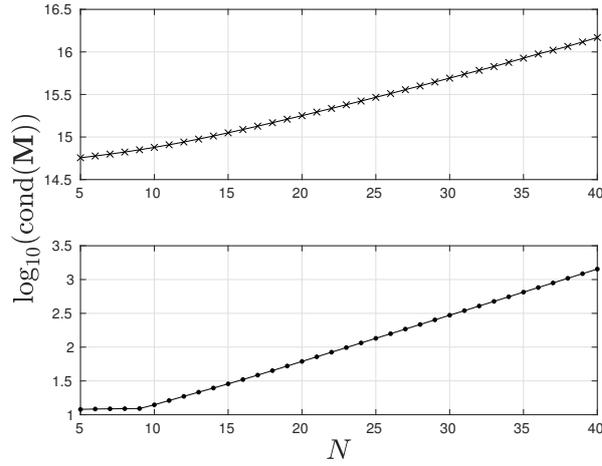


Figure 3.3.2: Dependence of the condition number of \mathbf{M} on N when $L = 10$. Top: \mathbf{M} with the final row set by the boundary condition $\theta_N = 0$. Bottom: \mathbf{M} with the final row set instead by the boundary condition $\theta_1 = 0$.

To calculate the topography one would take the solution $\underline{\theta}$ to (3.3.12) and use these values to evaluate (3.1.3) at its midpoints, yielding

$$\tau_j^M \equiv \tau_b(\phi_j^M) = \sum_{k=2}^{N_f} \left(\frac{\Delta\Phi}{2} (g_{f[k-1,j]}\Theta_{k-1} + g_{f[k,j]}\Theta_k) \right) - \sum_{k=2}^{N_b} \left(\frac{\Delta\phi}{2} (g_{b[k-1,j]}\theta_{k-1} + g_{b[k,j]}\theta_k) \right), \quad (3.3.14)$$

for $j = 1, 2, \dots, N_b - 1$, with:

$$g_{f[k,j]} = \left(1 + e^{\pi(\phi_j^M - \Phi_k)} \right)^{-1}; \quad g_{b[k,j]} = \left(1 - e^{\pi(\phi_j^M - \phi_k)} \right)^{-1}.$$

Finally, by evaluating (3.1.5) at its mid-points and using a central difference for the derivative, the equation

$$y_j = y_{j+1} - \Delta\phi e^{-\tau_j^M} \sin(\theta_j^M) \quad (3.3.15)$$

can be used to recover the profile of the topography y_b , working backwards from $y_{N_b} = 0$ which is set to be consistent with the assumption of a uniform stream downstream.

The matrix \mathbf{M} is expected to be of full rank, see the appendix of Robbins, Blyth, Maclean, and Binder, 2023. Calculating the rank of the matrix \mathbf{M} via MATLAB it was found to be full rank for $N \leq 367$. MATLAB returns the numerical rank of a matrix as the number of non-zero singular values satisfying $\sigma_i > \max(\text{size}(\mathbf{M})) * \text{eps}(\text{norm}(\mathbf{M}))$. Since the matrix \mathbf{M} is expected to be non-singular, we might follow a direct approach and simply pre-multiply both sides of (3.3.12) by the matrix inverse to obtain the solution $\underline{\boldsymbol{\theta}} = \mathbf{M}^{-1}\underline{\mathbf{b}}$. However, the matrix \mathbf{M} is very poorly conditioned. This is illustrated in Figure 3.3.2, where we set $N_f = N_b = N$ and plot the condition number $\text{cond}(\mathbf{M})$ against N for the truncation length $L = 10$. The ill-conditioning is considerably worse for the choice of boundary condition $\theta_N = 0$: even for $N = 15$ its condition number is on the order of 10^{15} ! The ill-conditioning means that in computational practice, for large enough N the matrices are effectively singular and, in the case of non-square systems, they are effectively rank deficient. Accordingly, standard approaches to solving (3.3.12) will be swamped with numerical error as the grid resolution is increased. In particular, the norm of the inverse $\|\mathbf{M}^{-1}\|$ will be large and the solution will tend to align itself with the eigenfunction corresponding to the smallest magnitude eigenvalue. Figure 3.3.3 shows the eigenfunction of \mathbf{M} associated with the eigenvalue 9.2×10^{-7} , when $N = 101$ and $L = 10$. The eigenfunction has a non-smooth sawtooth appearance.

To mitigate the difficulties with the ill-conditioning, we employ the truncated singular value decomposition (TSVD) method (e.g. Hansen, 1990b; Varah, 1973) to be discussed below. In the standard implementation of the singular value decomposition (SVD) method e.g. Griffel, 1989, the $(N_f \times N_b)$ matrix \mathbf{M} is expressed in the form $\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, where \mathbf{U} and \mathbf{V} are, respectively, $(N_f \times N_f)$ and $(N_b \times N_b)$ unitary matrices whose columns are the eigenvectors of $\mathbf{M}\mathbf{M}^T$ and $\mathbf{M}^T\mathbf{M}$ respectively. The $(N_f \times N_b)$ matrix $\boldsymbol{\Sigma}$ has along its leading diagonal the singular values σ_n , $n = 1, \dots, r \leq \min(N_f, N_b)$, corresponding to the positive square roots of the non-zero eigenvalues of $\mathbf{M}\mathbf{M}^T$, and all other elements zero. The singular values are ordered by size starting with the largest in the (1,1)

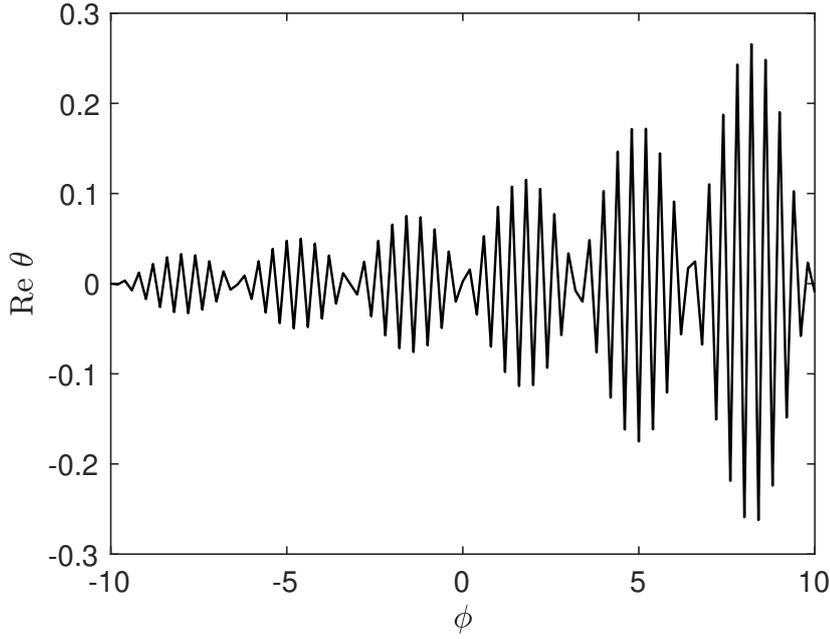


Figure 3.3.3: The eigenfunction of \mathbf{M} corresponding to the eigenvalue with magnitude 9.2×10^{-7} , when $L = 10$ and $N = 101$.

position.

Armed with the SVD we compute the Moore-Penrose inverse (Penrose, 1955) $\mathbf{M}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T$, where $\mathbf{\Sigma}^+$ is formed by replacing the non-zero entries of $\mathbf{\Sigma}$ with their reciprocals and then transposing the matrix (e.g. Ben-Israel and Greville, 2003, p.207). (Note that if \mathbf{M}^{-1} exists then $\mathbf{M}^+ = \mathbf{M}^{-1}$.) A solution to (3.3.12) exists iff the solvability condition $\mathbf{M}(\mathbf{M}^+\mathbf{b}) = \mathbf{b}$ holds; this ensures that $\mathbf{b} \in \text{Im}\mathbf{M}$. Then the complete set of solutions to (3.3.12) is given by e.g. James, 1978

$$\underline{\boldsymbol{\theta}} = \mathbf{M}^+\mathbf{b} + \underline{\mathbf{z}}, \quad (3.3.16)$$

where $\underline{\mathbf{z}} = (\mathbf{I} - \mathbf{M}^+\mathbf{M})\underline{\mathbf{w}}$, with $\underline{\mathbf{w}}$ an arbitrary $(N_b \times 1)$ vector, and $\underline{\mathbf{z}} \in \ker\mathbf{M}$. If the solvability condition fails then $\mathbf{b} \notin \text{Im}\mathbf{M}$, the system (3.3.12) is inconsistent, and (3.3.16) provides the linear least squares approximation of minimum norm e.g. Planitz, 1979.

We will discuss results for the inverse topography problem allowing for different numbers of grid points on the free surface and on the bottom, taking $N_f > N_b$ (overdetermined system), or $N_f < N_b$ (underdetermined system), or else $N_f = N_b$

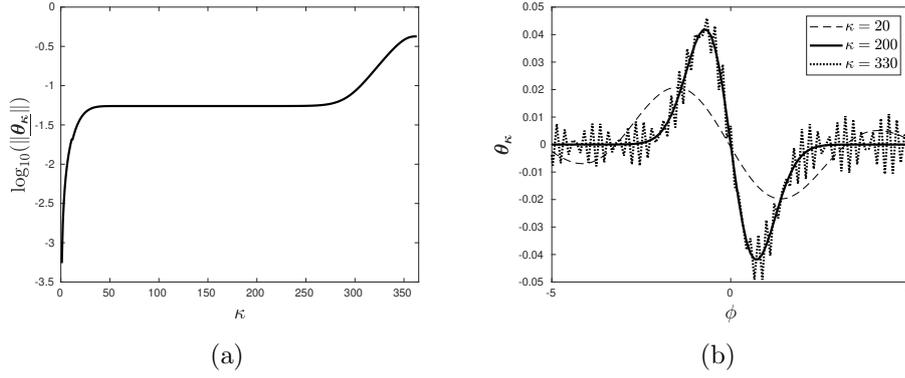


Figure 3.3.4: Results for the topography (3.3.1) at $F = 1.2$ when $L = 20$ and $N = 363$. (a) $\|\underline{\theta}_\kappa\|$ rises rapidly to a stable value for a range of the truncation parameter κ before beginning to increase further as the model fails to output smooth solutions. (b) Profiles of $\underline{\theta}_\kappa$ for various κ , showing how the output varies by selecting κ to correspond to the different sections of the $(\kappa, \|\underline{\theta}_\kappa\|)$ curve.

(square system). Since if $N_f \geq N_b$ we expect \mathbf{M} to have full rank, its kernel should be trivial so that the unique solution to (3.3.12) is given by (3.3.16) with \mathbf{z} set to zero. However, in computational practice \mathbf{M} is effectively rank deficient, as discussed above, so that the kernel is effectively non-trivial. Accordingly we expect that artificial sawtooth irregularities, like those seen in the eigenfunction in Figure 3.3.3, will become a dominant feature of the numerical solution as the grid resolution is refined. To work around this we follow the TSVD method (e.g. Hansen, 1990b; Varah, 1973) and replace \mathbf{M} in (3.3.12) with $\mathbf{M}_\kappa = \mathbf{U}\mathbf{\Sigma}_\kappa^+\mathbf{V}^T$, where $\mathbf{\Sigma}_\kappa^+$ is the rank κ matrix obtained by retaining the first κ largest singular values and setting the others to zero. The resulting rank κ approximate problem has least squares solutions of the form given in (3.3.16) with \mathbf{M}^+ replaced by \mathbf{M}_κ^+ . We find that the smoothest solution is that which minimises $\|\underline{\theta}_\kappa\|$:

$$\underline{\theta}_\kappa = \mathbf{M}_\kappa^+ \underline{\mathbf{b}}. \quad (3.3.17)$$

To illustrate the procedure we return to the test topography (3.3.1) examined in Figure 3.3.1. The free surface profile is computed first by solving the forward problem using Newton's method. Next the inverse problem is solved with the forward solution as input using the TSVD approach just described.

Figure 3.3.4(a) shows the logarithm of the norm $\|\underline{\theta}_\kappa\|$ for the backwards problem plotted against κ on a logarithmic scale. We see that $\|\underline{\theta}_\kappa\|$ reaches a plateau that extends over a wide range of κ values; thereafter the norm increases as κ approaches N and the numerical problems discussed above become prominent. The profile of $\underline{\theta}_\kappa$ plotted against ϕ is found to be visually smooth, and to remain the same, over the plateau region. Typical profiles for various κ are shown in Figure 3.3.4(b). The characteristic sawtooth numerical instability is evident for $\kappa = 330$. Typically when solving the backwards problem we produce a graph similar to that shown in Figure 3.3.4(a) to confirm the presence of a plateau. We then produce a Picard plot and plot the L-curve for the problem to aide in choosing a truncation rank κ . Finally we compute the topographic profile $y_b(\phi)$ using (3.1.11).

Figure 3.3.5 shows a convergence study for the test topography (3.3.1) for a square system with $N_f = N_b = N$. In Figure 3.3.5(a) we see good agreement between the exact topography (3.3.1), shown with a thick solid line, and the output from the backwards problem, shown for the two different discretisation levels $N = 101$ and $N = 721$ with a thin solid and a dotted line, respectively. In Figure 3.3.5(b) we plot the norm of the difference $\|y_b - y_T\|$ over the grid, where y_T is the prescribed topography given by (3.3.1), and y_b is the inversely computed topography. The error $\|y_b - y_T\|$ decreases like N^{-2} as N increases. Unless otherwise stated, in each of the results presented in the next section, the inversely found bottom profiles were used as input to the forward problem to check that the original free surface is recovered. While it is straightforward to do this for either subcritical or supercritical Froude number, the critical case $F = 1$ is computationally more challenging (Keeler, Binder, and Blyth, 2017).

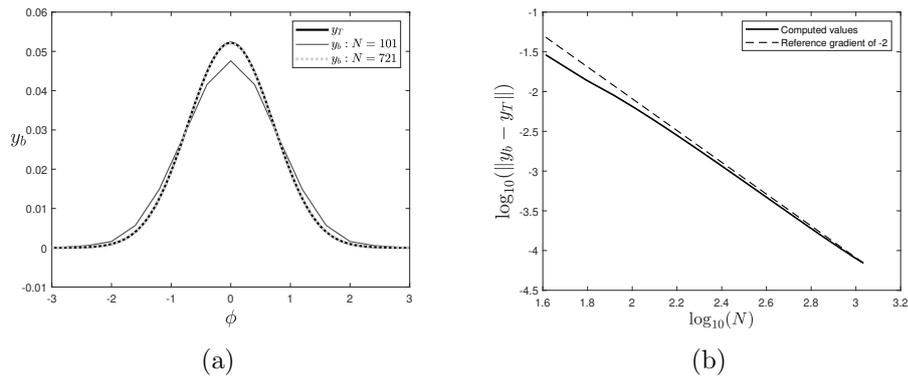


Figure 3.3.5: The case of topography (3.3.1) at $F = 1.2$ when $L = 20$ and $N = 101$. (a) The originally prescribed topography compared to those found by applying the TSVD method to the results of the forward Newton problem at two resolutions ($\kappa = 101$; $N = 101, 721$). (b) The norm of the error $\|y_b - y_T\|$, where y_T is the prescribed topography (3.3.1), and y_b is the TSVD solution.

3.4 Unforced Forward and Inverse Problems

3.4.1 Unforced forward problem

Forward solutions to the fKdV

In the absence of forcing the fKdV reduces to the usual KdV equation; namely

$$\eta_{xxx} + 9\eta\eta_x - 6\mu\eta_x = 0. \quad (3.4.1)$$

It is well known that there are two possible forms of solution for the free surface of the unforced problem (e.g. Whitham, 2011): the uniform stream $\eta = 0$, which exists as a solution for all values of F , and in supercritical flow the solitary wave solutions $\eta = 2\mu \operatorname{sech}^2\left(\sqrt{\frac{3\mu}{2}}(x - x_0)\right)$ are also permitted, where $x_0 \in \mathbb{R}$ shifts the position of the crest of the wave.

Forward solutions to the fully nonlinear BI model

The governing equations for the fully nonlinear problem are:

$$e^{2\tau_f(\phi)} + \frac{2}{F^2}(y_f(\phi) + P(\phi)) - 1 - \frac{2}{F^2} = 0; \quad (3.4.2)$$

$$\tau_f(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0; \quad (3.4.3)$$

$$\tau_b(\phi) = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0 - \int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0; \quad (3.4.4)$$

$$\frac{\partial y_f}{\partial \phi} = e^{-\tau_f} \sin(\theta_f); \quad (3.4.5) \quad \frac{\partial y_b}{\partial \phi} = e^{-\tau_b} \sin(\theta_b); \quad (3.4.6)$$

$$\frac{\partial x_f}{\partial \phi} = e^{-\tau_f} \cos(\theta_f); \quad (3.4.7) \quad \frac{\partial x_b}{\partial \phi} = e^{-\tau_b} \cos(\theta_b); \quad (3.4.8)$$

The absence of forcing means that we have $P = y_b = \theta_b = 0$. These values can be substituted into (3.4.2) which may subsequently be rearranged for τ_f as

$$\tau_f = \frac{1}{2} \ln \left(1 + \frac{2}{F^2} (1 - y_f) \right). \quad (3.4.9)$$

Similarly treating (3.4.3) another expression for τ_f is obtained as

$$\tau_f = \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0, \quad (3.4.10)$$

which, upon equating equating with (3.4.9), allows for us to find the relation

$$\int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi - \phi_0)}} d\phi_0 = \frac{1}{2} \ln \left(1 + \frac{2}{F^2} (1 - y_f) \right). \quad (3.4.11)$$

Note first that the uniform stream with $\theta_f = 0$ and $y_f = 1$ is a solution to

(3.4.11). Alternatively, one could integrate (3.4.5), keeping in mind that we integrate backwards from the boundary condition at positive infinity, to yield $y_f = \int_{\infty}^{\phi} e^{-\tau_f(\phi')} \sin(\theta_f(\phi')) d\phi'$ into which substitution of (3.4.10) gives

$$y_f = \int_{\infty}^{\phi} \left[\exp \left(- \int_{-\infty}^{\infty} \frac{\theta_f(\phi_0)}{1 - e^{\pi(\phi' - \phi_0)}} d\phi_0 \right) \sin(\theta_f(\phi')) \right] d\phi', \quad (3.4.12)$$

showing that y_f can be expressed as a nonlinear function of θ_f . As such, the right-hand side of (3.4.11) can also be considered to be a nonlinear function of θ_f and so (3.4.11) represents a nonlinear Fredholm integral equation of the second kind for θ_f . This allows for non-uniqueness in the solution, an expected result as it is known that analogues to the KdV's solitary wave solutions exist for the fully nonlinear problem (see, for example, Figure 3.2.4).

3.4.2 Unforced inverse problem

Considering the inverse problem with no forcing means that we are considering a uniform stream with no pressure acting on the surface, therefore we have $P = \theta_f = \eta = 0$ and $y_f = 1$.

Inverse solution to the fKdV

The fKdV gives the forcing in the inverse problem as

$$f(x) = 2\mu\eta - \frac{1}{3}\eta_{xx} - \frac{3}{2}\eta^2, \quad (3.4.13)$$

substituting for the flat free surface, $\eta(x) = 0$, it is immediately apparent that the forcing must also be zero, i.e. $f(x) = 0$. As there is no pressure forcing acting on the surface it follows that $y_b = 0$.

Inverse solutions to the fully nonlinear BI model

Substituting $P = 0$ and $y_f = 1$ into (3.4.2) we obtain

$$e^{2\tau_f} = 1 \implies \tau_f = 0.$$

Insertion of this result for τ_f and that $\theta_f = 0$ into (3.4.3) we find that

$$\int_{-\infty}^{\infty} \frac{\theta_b(\phi_0)}{1 + e^{\pi(\phi - \phi_0)}} d\phi_0 = 0. \quad (3.4.14)$$

Defining

$$g(\phi_0) = \frac{1}{1 + e^{\pi\phi_0}},$$

we may represent (3.4.14) as the convolution

$$(\theta_b * g)(\phi) = \int_{-\infty}^{\infty} \theta_b(\phi_0)g(\phi - \phi_0) d\phi_0 = 0.$$

Application of the convolution theorem to the above equation yields

$$\tilde{\theta}_b(w) \cdot \tilde{g}(w) = 0, \quad (3.4.15)$$

where the Fourier transform of a variable h , denoted by the addition of a tilde as \tilde{h} , is defined by

$$\tilde{h}(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(x)e^{iwx} dx.$$

If $\tilde{\theta}_b(w) = 0$ then $\theta_b = 0$ and we see that a flat topography is a solution to the inverse problem for a flat free surface. We are interested in the question of existence of other solutions and so instead we turn our attention to $\tilde{g}(w)$. First, we note that by subtracting one half from $g(\phi_0)$ we find

$$g(\phi_0) - \frac{1}{2} = \frac{1}{2} \left(\frac{1 - e^{\pi\phi_0}}{1 + e^{\pi\phi_0}} \right) = -\frac{1}{2} \tanh \left(\frac{\pi\phi_0}{2} \right)$$

and can therefore obtain the alternate expression

$$g(\phi_0) = \frac{1}{2} \left(1 - \tanh \left(\frac{\pi\phi_0}{2} \right) \right).$$

While the Fourier transform of $g(\phi_0)$ doesn't classically exist, it can still be calculated in terms of distributions (e.g. Griffel, 2002), so we shall proceed to write

$$\tilde{g}(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{2} \left(1 - \tanh \left(\frac{\pi\phi_0}{2} \right) \right) e^{iw\phi_0} d\phi_0,$$

alternatively we may write

$$2\tilde{g}(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 1 \cdot e^{iw\phi_0} d\phi_0 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tanh \left(\frac{\pi\phi_0}{2} \right) e^{iw\phi_0} d\phi_0,$$

or

$$2\tilde{g}(w) = \sqrt{2\pi}\delta(w) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tanh \left(\frac{\pi\phi_0}{2} \right) e^{iw\phi_0} d\phi_0, \quad (3.4.16)$$

where we have used (2.1.4) to express the Fourier transform of unity in terms of the Dirac delta function. In order to proceed we must obtain the Fourier transform for the $\tanh(\frac{\pi\phi_0}{2})$ term. First, consider the simpler form $T(x) = \tanh(x)$, the Fourier transform of which is then given by

$$\tilde{T}(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tanh(x) e^{iwx} dx.$$

Now define $S(x) = \operatorname{sech}^2(x)$ and note that $S(x) = \frac{d}{dx}(T(x))$. By the properties of the Fourier transform of derivatives we have the relation

$$\tilde{S}(w) = -iw \cdot \tilde{T}(w),$$

which allows for $\tilde{T}(w)$ to be found as

$$\tilde{T}(w) = \frac{-1}{iw} \tilde{S}(w) = \frac{-1}{iw} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \operatorname{sech}^2(x) e^{iwx} dx. \quad (3.4.17)$$

In order to evaluate $\tilde{S}(w)$ we denote $\Theta(z) = \operatorname{sech}^2(z)e^{iwz}$ and consider the complex contour integral

$$I = \oint_{\Gamma} \operatorname{sech}^2(z)e^{iwz} dz$$

where $z = x + iy$ and Γ is a contour lying in the complex plane. The function $\Theta(z)$ has an infinite number of poles occurring at $z = z_n$ such that

$$\cosh(z_n) = 0 \implies z_n = \frac{i\pi(2n+1)}{2} \text{ for } n \in \mathbb{Z}.$$

We construct the contour Γ as the rectangular region bounded by $y = \pi$, the

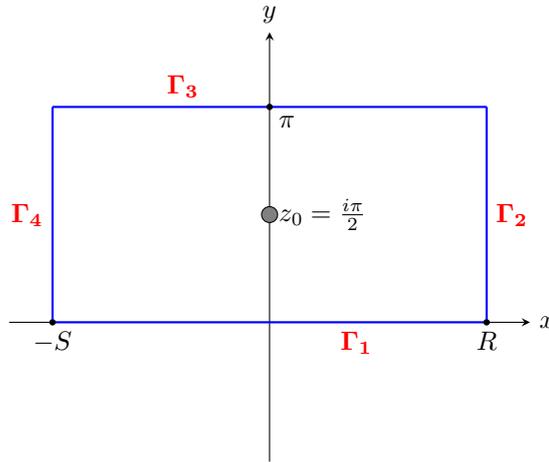


Figure 3.4.1: A diagram of how the curve Γ is subdivided for reference in computing the contour integral around Γ . This contour contains one pole at $z = z_0 = \frac{i\pi}{2}$.

real axis, and the lines $x = -S$ and $x = R$ for $R, S > 0$ and $R, S \in \mathbb{R}$ (see Figure 3.4.1). This region contains only one pole at $z_0 = \frac{i\pi}{2}$ and so by Cauchy's residue theorem we may write

$$I = \oint_{\Gamma} \operatorname{sech}^2(z)e^{iwz} dz = 2\pi i \operatorname{Res} \left[\Theta(z), z = \frac{i\pi}{2} \right]. \quad (3.4.18)$$

To calculate the residue we introduce the variable $\xi = z - \frac{i\pi}{2}$ and consider the series expansions

$$e^{iwz} = e^{iw(i\frac{\pi}{2} + \xi)} = e^{-\frac{w\pi}{2}} e^{iw\xi} = e^{-\frac{w\pi}{2}} (1 + iw\xi + \mathcal{O}(\xi^2))$$

and

$$\cosh(z) = \cosh\left(\xi + \frac{i\pi}{2}\right) = i \sinh(\xi) = i \left(\xi + \frac{\xi^3}{3!} + \mathcal{O}(\xi^5)\right).$$

By taking the reciprocal square of the $\cosh(z)$ expansion in order to obtain

$$\operatorname{sech}^2(z) = \frac{1}{i^2 \xi^2 \left(1 + \frac{\xi^2}{3!} + \dots\right)^2} = \frac{-1}{\xi^2} \left(1 - \frac{1}{3}\xi^2 + \mathcal{O}(\xi^4)\right)$$

we can then find the series expansion for $\Theta(z)$ as

$$\begin{aligned} \Theta(z) &= e^{iwz} \operatorname{sech}^2(z) = -\frac{e^{-\frac{w\pi}{2}}}{\xi^2} \left(1 - \frac{1}{3}\xi^2 + \dots\right) (1 + iw\xi + \dots) \\ &= -\frac{e^{-\frac{w\pi}{2}}}{\xi^2} (1 + iw\xi + \dots) = -e^{-\frac{w\pi}{2}} \left(\frac{1}{\xi^2} + \frac{iw}{\xi} + \mathcal{O}(\xi^0)\right). \end{aligned}$$

Now that we have written $\Theta(z)$ in the form of a Laurent expansion around the pole as

$$\Theta(z) = \sum_{n=-\infty}^{\infty} a_n \left(z - \frac{i\pi}{2}\right)^n = \sum_{n=-\infty}^{\infty} a_n \xi^n,$$

the residue is given simply as being equal to the coefficient $a_{-1} = -iwe^{-\frac{w\pi}{2}}$, substitution of which into (3.4.18) yields

$$I = \oint_{\Gamma} \operatorname{sech}^2(z) e^{iwz} dz = 2\pi w e^{-\frac{w\pi}{2}}. \quad (3.4.19)$$

Defining I_j to be

$$I_j = \int_{\Gamma_j} \Theta(z) dz,$$

where Γ_j are the path segments of Γ , as shown in Figure 3.4.1, then (3.4.19) can be expressed as

$$I = \sum_{j=1}^4 I_j = 2\pi w e^{-\frac{w\pi}{2}}. \quad (3.4.20)$$

On Γ_1 we have $z = x$, $dz = dx$ and so

$$I_1 = \int_{-S}^R \operatorname{sech}^2(x) e^{iw x} dx. \quad (3.4.21)$$

Now, on Γ_2 we have instead $z = R + iy$, $dz = i dy$ and so

$$I_2 = \int_0^\pi \operatorname{sech}^2(R + iy) e^{iw(R+iy)} i dy = \int_0^\pi \frac{4i e^{iwR} e^{-wy}}{(e^R e^{iy} + e^{-R} e^{-iy})^2} dy,$$

the absolute value of which is bounded by

$$|I_2| \leq \int_0^\pi \frac{4|e^{-wy}|}{|(e^R + e^{-R} e^{-2iy})^2|} dy. \quad (3.4.22)$$

The factor $(e^R + e^{-R} e^{-2iy})$ appearing in the denominator of the integrand can be bounded below by

$$0 < e^R - e^{-R} \leq e^R + e^{-R} e^{-2iy}$$

where the strict inequality arises as a consequence of $R > 0$. Further, by squaring and inverting these terms, we may write

$$\frac{1}{(e^R + e^{-R} e^{-2iy})^2} \leq \frac{1}{(e^R - e^{-R})^2}.$$

This result is then used alongside (3.4.22) to give a bound on the absolute value of I_2 as

$$|I_2| \leq \int_0^\pi \frac{4|e^{-wy}|}{|(e^R - e^{-R})^2|} dy = |\operatorname{cosech}^2(R)| \int_0^\pi |e^{-wy}| dy. \quad (3.4.23)$$

From the above we can then conclude that in the limit $R \rightarrow \infty$ we have

$$\lim_{R \rightarrow \infty} I_2 = 0. \quad (3.4.24)$$

On Γ_4 we have $z = -S + iy$, $dz = i dy$ and so

$$I_4 = \int_{\pi}^0 \operatorname{sech}^2(-S + iy) e^{iw(-S+iy)} i dy = \int_{\pi}^0 \frac{4i e^{-iwS} e^{-wy}}{(e^{-S} e^{iy} + e^S e^{-iy})^2} dy,$$

with which an analogous process to that applied to Γ_2 can be followed to show that

$$\lim_{S \rightarrow \infty} I_4 = 0. \quad (3.4.25)$$

On the remaining path segment, Γ_3 , we have $z = x + i\pi$, $dz = dx$ and so

$$I_3 = \int_R^{-S} \operatorname{sech}^2(x + i\pi) e^{iw(x+i\pi)} dx.$$

Making use of the result $\cosh(x + i\pi) = -\cosh(x)$ and exchanging the limits of integration we have

$$I_3 = -e^{-w\pi} \int_{-S}^R \operatorname{sech}^2(x) e^{iw x} dx = -e^{-w\pi} I_1. \quad (3.4.26)$$

We may now rewrite (3.4.20) as

$$I = \sum_{j=1}^4 I_j = (1 - e^{-w\pi}) I_1 + I_2 + I_4 = 2\pi w e^{-\frac{w\pi}{2}}.$$

Taking the limit as $R, S \rightarrow \infty$ the contributions from I_2 and I_4 vanish, with the expression simplifying to

$$\lim_{R \rightarrow \infty} \lim_{S \rightarrow \infty} I = (1 - e^{-w\pi}) \lim_{R \rightarrow \infty} \lim_{S \rightarrow \infty} I_1 = 2\pi w e^{-\frac{w\pi}{2}}$$

which can be rearranged to yield

$$\int_{-\infty}^{\infty} \operatorname{sech}^2(x) e^{iw x} dx = \frac{2\pi w e^{-\frac{w\pi}{2}}}{(1 - e^{-w\pi})}.$$

Noting the left-hand side of the above equation is $\sqrt{2\pi} \tilde{S}(w)$ then, after

multiplying the right-hand side by $e^{\frac{w\pi}{2}}/e^{\frac{w\pi}{2}}$, we arrive at

$$\tilde{S}(w) = w\sqrt{\frac{\pi}{2}} \operatorname{cosech}\left(\frac{w\pi}{2}\right),$$

which, by (3.4.17), allows us to find at last the Fourier transform of $T(x) = \tanh(x)$ to be

$$\tilde{T}(w) = i\sqrt{\frac{\pi}{2}} \operatorname{cosech}\left(\frac{w\pi}{2}\right).$$

The scaling property of the Fourier transform states that given the function $T_c(x) = T(cx)$ then its transform is given by $\tilde{T}_c = \frac{1}{|c|}\tilde{T}\left(\frac{k}{c}\right)$; as such we find

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tanh\left(\frac{\pi\phi_0}{2}\right) e^{iw\phi_0} d\phi_0 = i\sqrt{\frac{2}{\pi}} \operatorname{cosech}(w).$$

Finally, inserting this result into (3.4.16), we obtain the Fourier transform of g to be

$$\tilde{g}(w) = \sqrt{\frac{\pi}{2}}\delta(w) - \frac{i}{\sqrt{2\pi}} \operatorname{cosech}(w). \quad (3.4.27)$$

There are no values w for which $\tilde{g}(w) = 0$. We find then that, for (3.4.15) to hold we have

$$\tilde{\theta}_b \cdot \left(\sqrt{\frac{\pi}{2}}\delta(w) - \frac{i}{\sqrt{2\pi}} \operatorname{cosech}(w) \right) = 0 \quad (3.4.28)$$

and so the only allowable inverse solution for a flat free surface with no applied pressure is for the topography to also be flat, i.e.

$$\tilde{\theta}_b(w) = 0 \implies \theta_b(\phi) = 0 \implies y_b(\phi) = 0.$$

Results: Inversely Retrieving Topography from Computed Forward Solutions

In this chapter we will investigate the performance of the TSVD method for different types of free surface and the use of the L-curve and the DPC in helping to select a suitable truncation rank. Flow over a topography made of a single Gaussian will be considered first before then adding second Gaussian. By first computing solutions to the forward problem for flow over a prescribed topography y_T we obtain sets of surface data to be used in the inverse problem for which we know the solution to compare against the calculated inverse solutions. For an overview of the forward method and its solutions see Section 3.2.

4.1 Topography consisting of one Gaussian

For the purposes of this section we prescribe a topography of the form

$$y_T(\phi) = ae^{-(b\phi)^2} \tag{4.1.1}$$

to then be used to solve the forward problem, by Newton's method, providing a free surface $y_f(\phi)$. This free surface is subsequently used as the input in the inverse TSVD method to output a solution $y_b(\phi)$ for the topography. The output

y_b can then be directly compared to the prescribed y_T in order to assess the performance of the TSVD method. We will look first at solutions to supercritical flow; considering both free surface solutions that are perturbations of the uniform stream and those of solitary waves; nonlinearity in the forward problem means that both the perturbations to the uniform stream and to the solitary wave solutions can be obtained from the same topography (see Figure 3.2.3). The method will then be tested on an unforced solitary wave solution, the limit of the forced solitary wave as $a \rightarrow 0$, to investigate how well the inverse TSVD method can recover a flat topography from a non-zero input. Turning then to solutions to forward problems that are not everywhere supercritical we analyse the inverse problem for hydraulic falls before a brief consideration of subcritical flows featuring a train of surface waves.

4.1.1 Supercritical flow as a perturbation of the uniform stream

We first consider supercritical flow over a Gaussian dip in the topography (Figure 4.1.1). The output from the forward problem is displayed in Figure 4.1.1(a) alongside the prescribed topography y_T . The output topography y_b of the inverse problem using y_f as an input is then compared more closely against y_T in Figure 4.1.1(b). The truncation rank $\kappa = 89$ was chosen by consideration of Figure 4.1.1(c), which shows that the DPC will be satisfied for this choice of κ , and Figure 4.1.1(d), where it can be seen that this solution lies on the segment of the L-curve on which the norm of the solution is stable to a small change in κ . In fact, it was found that outputs y_b to this particular problem do not start to display, at graphical accuracy, the typical erroneous oscillations until the truncation rank is taken with $\kappa > 150$.

The process above was repeated for flow over a larger Gaussian, $a = 0.2$ at the higher Froude number of $F = 1.5$, which produces a greater response on the free surface (Figure 4.1.2(a)). For this problem the truncation rank $\kappa = 146$ was selected by study of the ρ_i curve as for $i \leq 146$ the problem satisfies the DPC, see

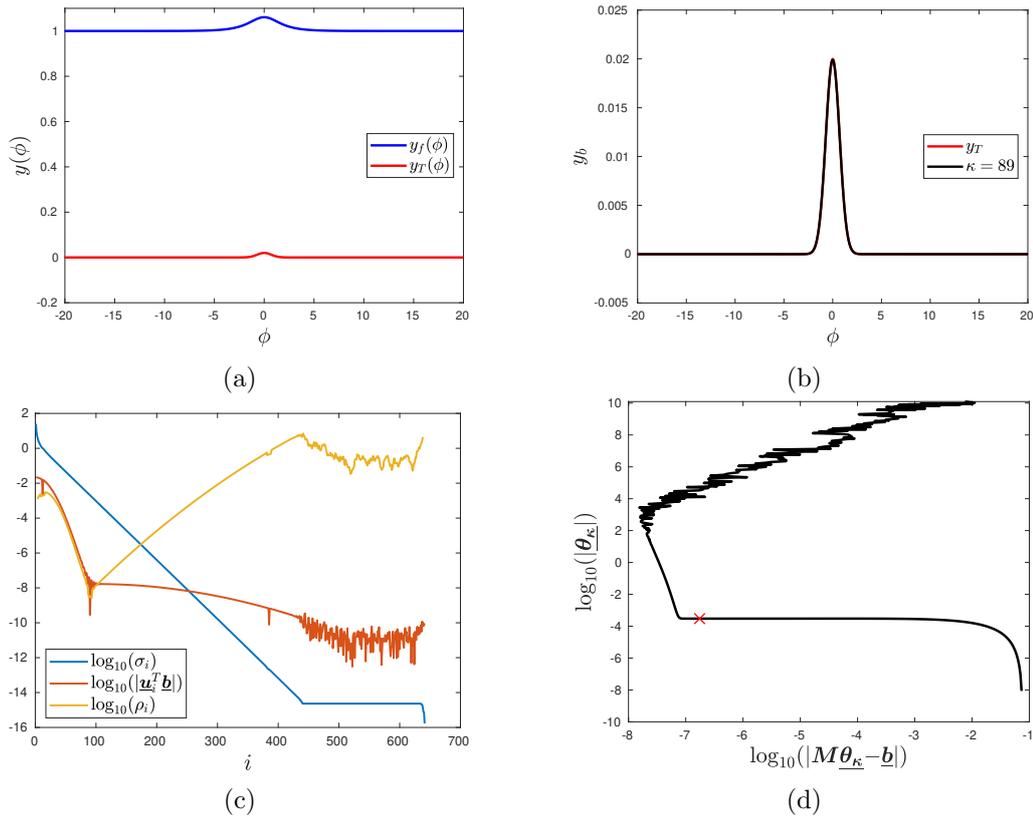


Figure 4.1.1: Inverse results for a perturbation to the uniform stream solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.02, b = 1, L = 20, N = 641, F = 1.1$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 89$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 89$.

Figure 4.1.2(c). The result for the inversely computed topography is displayed in Figure 4.1.2(b), showing that a good inverse solution can be obtained for flow past larger amplitude topographies; the agreement found in Figure 4.1.1 was not simply due to having picked such a diminutive forcing.

4.1.2 Supercritical flow as a perturbation of the solitary wave

The forward solution presented in Figure 4.1.1 is an example of a free surface that is a perturbation of the uniform stream. For the same topography it is possible to instead calculate the corresponding forced solitary wave solution (see Figure 3.2.3 and Figure 3.2.4). This solution is then used as the input to then

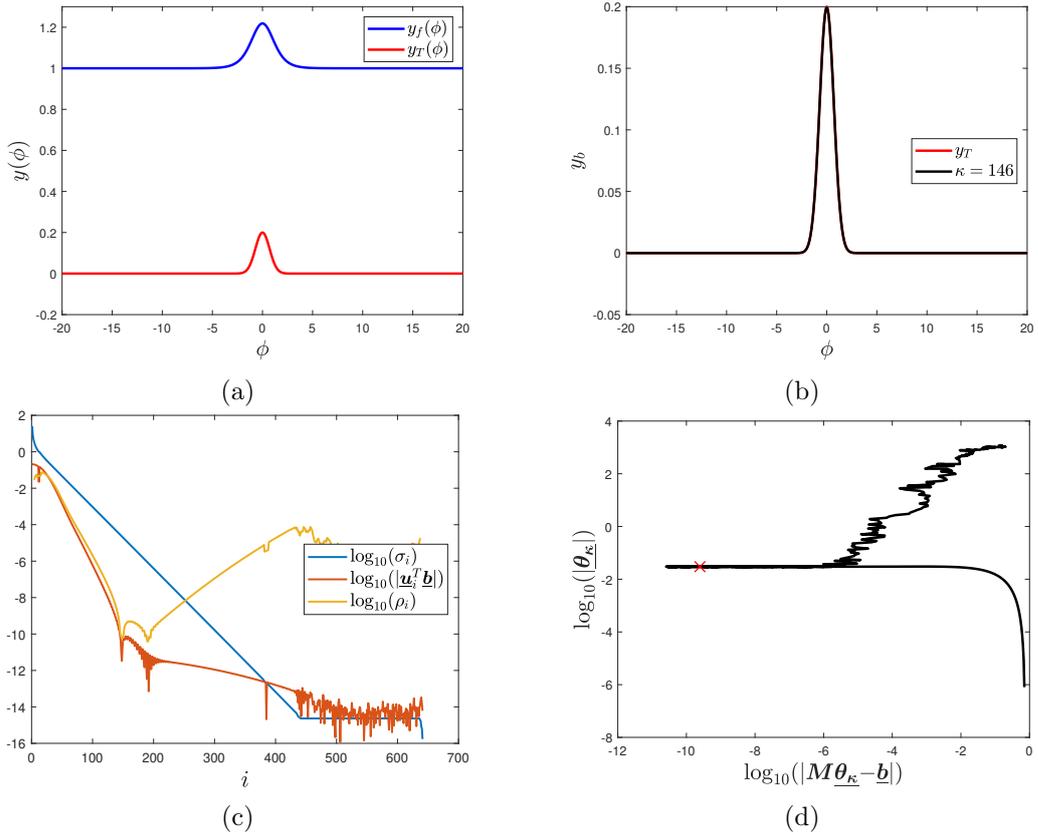


Figure 4.1.2: Inverse results for a perturbation to the uniform stream solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.2, b = 1, L = 20, N = 641, F = 1.5$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 146$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 146$.

attempt to recover the originally prescribed topography through the inverse method (Figure 4.1.3). The prescribed topography y_T and computed free surface y_f are shown in Figure 4.1.3(a). The inverse solution, y_b , for the topography is compared to y_T in Figure 4.1.3(b) showing that the TSVD method has been able to accurately recover the topography from the surface data of the solitary wave. Figure 4.1.3(c) informs our choice of the truncation rank, $\kappa = 73$, as the DPC is seemingly satisfied until this point with the monotonic decay of ρ_i for $16 \leq i \leq 73$. A choice of $\kappa \approx 90$, corresponding to the second local minimum in ρ_i , yielded results that matched, to graphical accuracy, those of truncation rank $\kappa = 73$. Solutions from this range of κ all lie before the corner of the L-curve where the norm of the solution should be stable

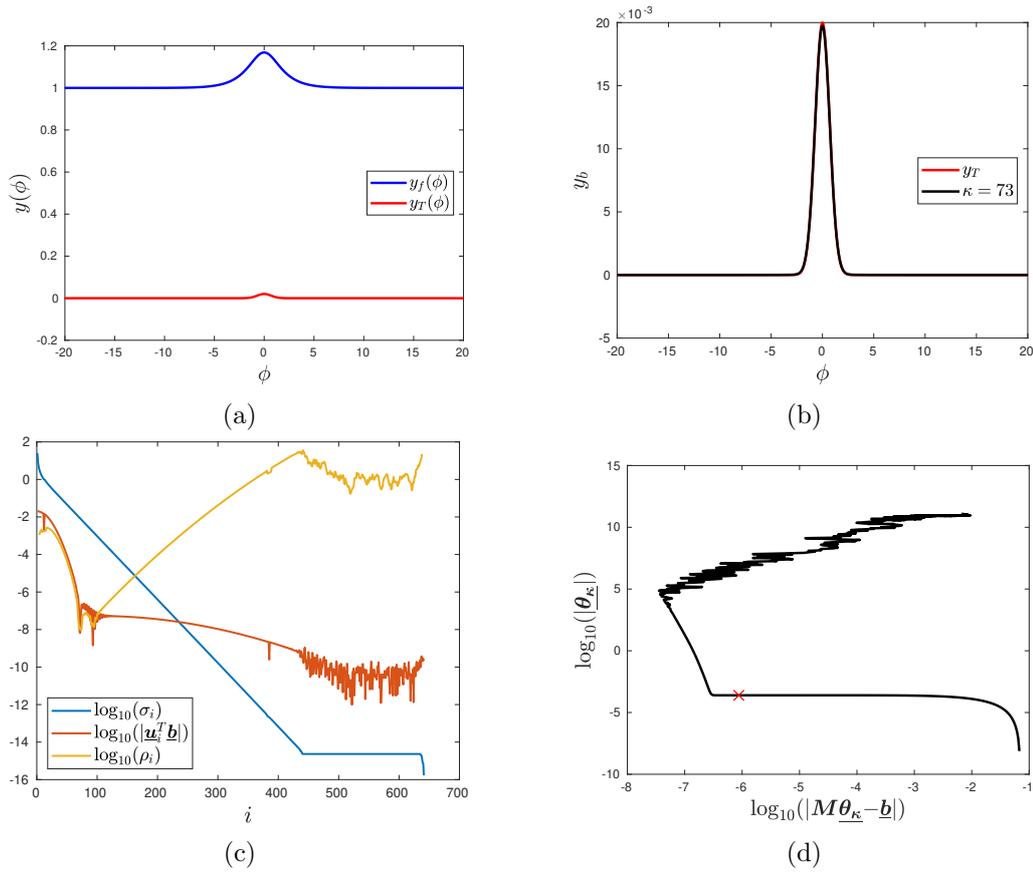


Figure 4.1.3: Inverse results for a perturbation to the solitary wave solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.02, b = 1, L = 20, N = 641, F = 1.1$ (These are the same parameters as for the uniform stream solution of Figure 4.1.1). (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 73$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 73$.

to a small change in κ (Figure 4.1.3(d)).

Keeping the topography fixed an increase in the Froude number will lead to an increase in the maximum height achieved by the solitary wave. The solitary wave profile found in Figure 4.1.3 was used as the start point from which to slowly increase F , using each iteration's solution as an initial guess for the next iteration, until the profile shown in Figure 4.1.4(a) was calculated for $F = 1.3$. There is a very clear choice of $\kappa = 41$ for the truncation rank based on the ρ_i curve shown in Figure 4.1.4(c) and the size and position of the underlying topography are well captured by the TSVD method. There is a disagreement between the

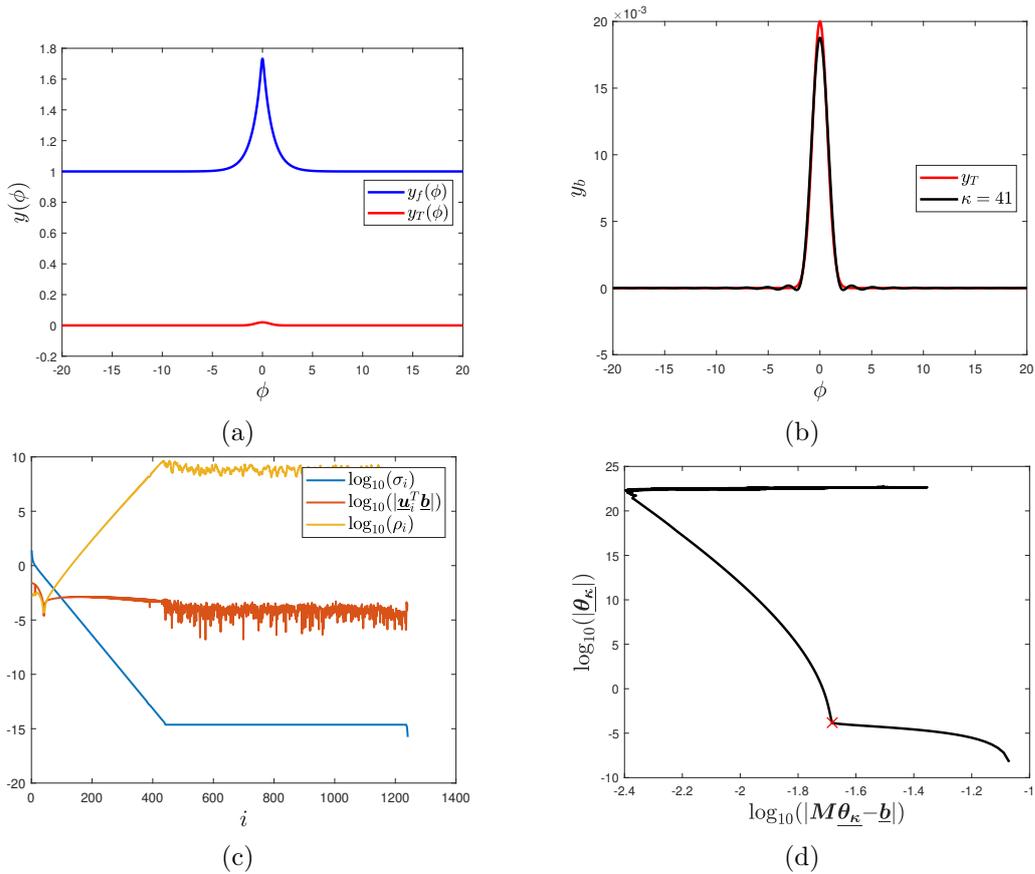


Figure 4.1.4: Inverse results for a perturbation to the solitary wave solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.02, b = 1, L = 20, N = 1241, F = 1.3$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 41$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 41$.

maximum height of the true topography and that output by the model, further some small oscillations appear at the base of the topography, however it was found that the size of these disagreements, both the height discrepancy and the oscillations, could be reduced by increasing the value of N . We did not push higher than $N = 1241$ due to the time taken for the forward problem to converge to this solution to provide input data for the inverse problem.

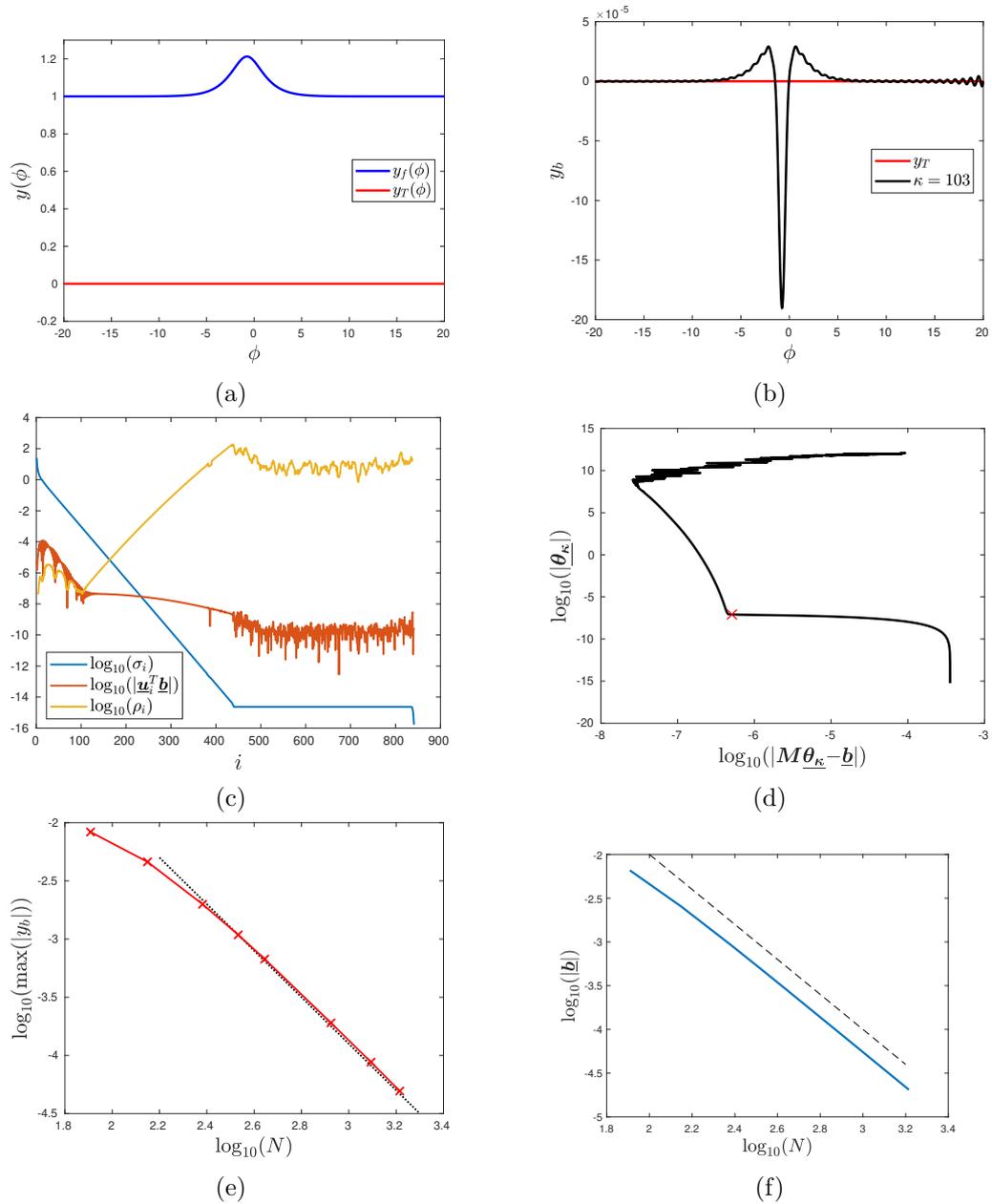


Figure 4.1.5: Inverse results for an unforced solitary wave solution with the parameter values $a = 0, L = 20, N = 841, F = 1.1$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 41$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 103$. (e) The maximum displacement of the topography from the zero level decays with N^{-2} (red) Reference line of gradient -2 (black dotted) (f) The norm of the right-hand side of the matrix equation approaches zero like N^{-2} (blue) and a reference line of gradient -2 (black dashed).

4.1.3 Unforced solitary wave

Consider instead the surface data of an unforced solitary wave (Figure 4.1.5(a)). Distinct now from the case of the forced solitary wave the decay of ρ_i for the unforced solitary wave is less smooth and does not decay purely monotonically although it is, on average, tending downwards until $\kappa = 103$ (Figure 4.1.5(c)) which is the truncation rank we have taken for the solution for the topography displayed in Figure 4.1.5(b). The sharp, but small, central feature appearing in the solution for the topography can not be eliminated in this case by a different choice of κ , persisting for all possible truncation ranks that produce solutions not dominated by error (i.e., large amplitude sawtooth oscillations). With the maximum displacement of the topography from $y = 0$ being $\mathcal{O}(10^{-4})$ these features are not visible when plotted on a scale like that of Figure 4.1.5(a). In Figure 4.1.5(e) the dependence of the maximum displacement of the topography from the zero level, $\max(|y_b|)$, on the number of input data, N , was studied by calculating the inverse solution at different resolutions and plotting the results on a log-log scale, revealing that that for the range of N tested this maximum displacement decays approximately with N^{-2} . It was shown in 3.4.2 that in the absence of forcing and no surface disturbance that the inverse method gave a flat topography with the right-hand side of the Fredholm equation becoming zero. Evaluating the norm of the right-hand side of the discretised system $\mathbf{M}\underline{\theta} = \underline{b}$ for given the input data of unforced solitary waves calculated at different values of N we find that right-hand side appears to be decaying like N^2 .

4.1.4 Hydraulic falls

When the flow transitions from subcritical to supercritical due to its interaction with the topography a hydraulic fall solution occurs, the surface profile calculated for this example and the corresponding prescribed topography are shown in Figure 4.1.6(a). Consideration of Figure 4.1.6(c) suggests that the underlying problem satisfies the DPC for $\kappa \leq 115$ and plotting the L-curve finds

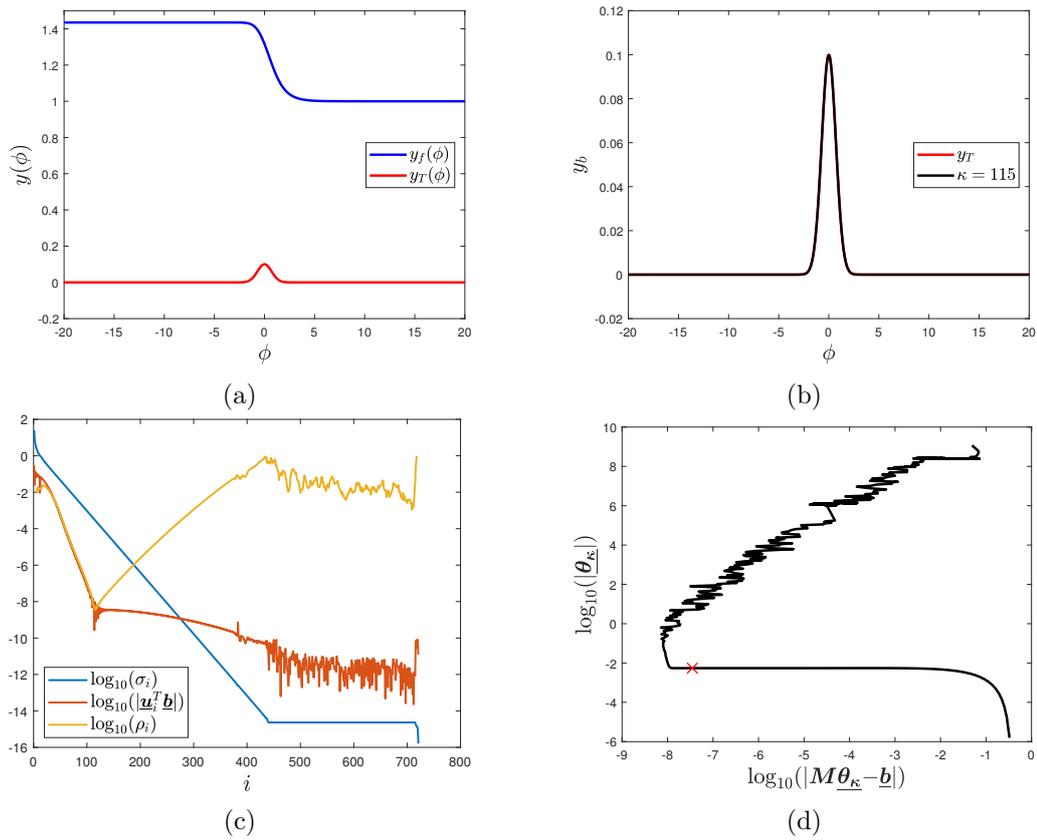


Figure 4.1.6: Inverse results for a hydraulic fall solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.1, b = 1, L = 20, N = 721$. The Froude number is calculated as part of the solution and was found to be $F \approx 1.3005$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 115$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 115$.

that $\kappa = 115$ corresponds to a solution close to the L-curve's corner (Figure 4.1.6(d)) and as such this should be a good choice of the truncation rank. The topography output for $\kappa = 115$ is shown in Figure 4.1.6(b) and is found to accurately capture the topography originally prescribed in the forward problem. This process has been repeated for a hydraulic fall over a larger topographical disturbance, leading to a greater change in uniform stream depth, in Figure 4.1.7. The model again performs well with a choice of $\kappa = 123$ based on Figure 4.1.7(c) allowing the topography to be recovered (Figure 4.1.7(b)).

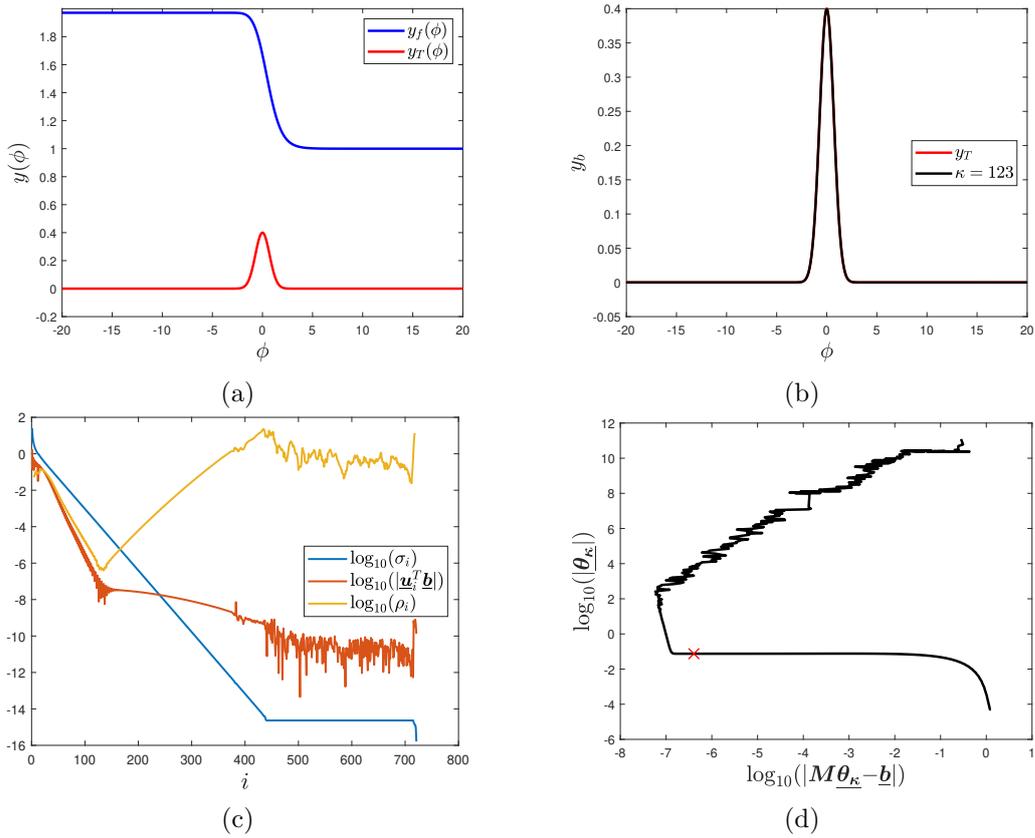


Figure 4.1.7: Inverse results for a hydraulic fall solution over a Gaussian topography (4.1.1) with the parameter values $a = 0.4, b = 1, L = 20, N = 721$. The Froude number is calculated as part of the solution and was found to be $F \approx 1.6171$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 123$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 123$.

4.1.5 Subcritical flow with a train of waves

In general, solving the forward problem with the flow being subcritical everywhere will produce a surface with an upstream train of waves created by the interaction of the flow with the channel obstruction (Figure 4.1.8(a)). In order to satisfy the radiation condition we must consider the flow to be from right to left for these solutions. For this problem the decay of the Fourier coefficients $\mathbf{u}_i^T \mathbf{b}$ occurs slowly, after an initial interval of oscillation, and as such there is not a large interval of κ for which this problem might be said to satisfy the DPC (Figure 4.1.8(c)). Although the position and approximate dimensions

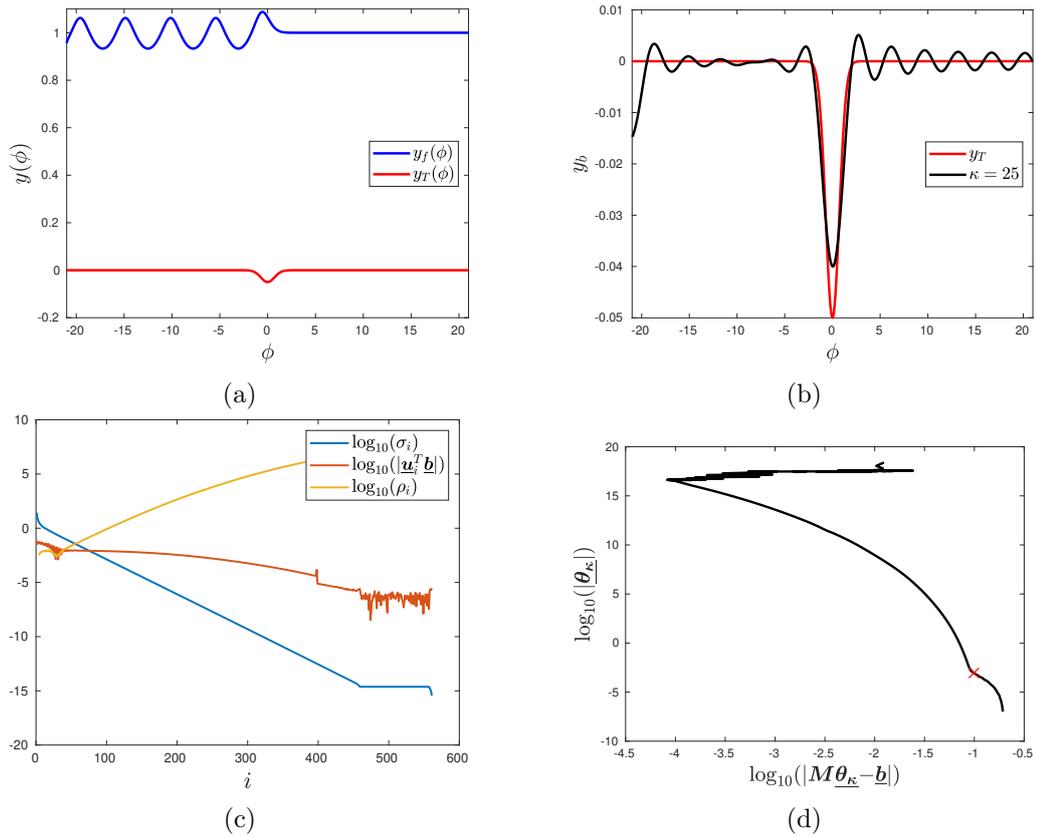


Figure 4.1.8: Inverse results for a subcritical solution with a train of waves over a Gaussian topography (4.1.1) with the parameter values $a = -0.05$, $b = 1$, $L = 21$, $N = 561$, $F = 0.8$. (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 25$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 25$.

of the topography could be obtained through application of the TSVD method for flows of this type (Figure 4.1.8(b)) it was not possible to remove the spurious oscillations occurring around the zero-level by selection of κ . In Figure 4.1.8(d) the L-curve for this problem is plotted. It can be seen that this L-curve does have a corner and that the selection of $\kappa = 25$ based on the DPC corresponds with a solution just prior to this corner. However, unlike the L-curves presented previously in this chapter, there is no flattened region before this corner corresponding to an interval of κ in which $\underline{\theta}_\kappa$ is stable with respect to κ . It was found that increasing N did not lead to a reduction in the amplitude of these oscillations about the zero-level, nor did changes to the domain truncation L .

4.2 Topography consisting of two Gaussians

Let us now consider a topography with two identical Gaussian features, described by

$$y_T = a \left(e^{-(b(\phi-c))^2} + e^{-(b(\phi+c))^2} \right), \quad (4.2.1)$$

where c is a parameter that allows for the separation of these features to be varied. We will compute a free surface both in the supercritical and subcritical flow regimes and attempt to then solve the inverse problem for the topography, the results of which large agree with those of a single Gaussian feature. We will then look at the case of trapped waves in subcritical flow, achieved by careful positioning of the topography, and how the inverse problem fares with this input.

4.2.1 Supercritical flow

The forward solution over a two bumped topography topography with the separation $c = 10$ was computed in order to be used as an input to the inverse problem (Figure 4.2.1(a)). A truncation rank of $\kappa = 158$ due to the decay of the ρ_i curve suggesting that the DPC is satisfied for $i \leq 158$. This produced an output for the topography that is in good agreement with the true topography y_T (Figure 4.2.1(b)).

4.2.2 Subcritical flow with a train of waves

As might be expected from the results of Figure 4.1.8, the case of subcritical flow over a single Gaussian, attempts to solve the inverse problem for the subcritical flow over two dips in the topography with waves on the surface as for $\phi < 5$ is substantially more challenging than for supercritical flow. It appears from Figure 4.2.2(c) that the DPC is not satisfied for any values $i \geq 14$ and the result corresponding to $\kappa = 13$ is a poor approximation to the underlying topography. Interestingly, taking larger values for κ , say $\kappa = 40$, can accurately capture the

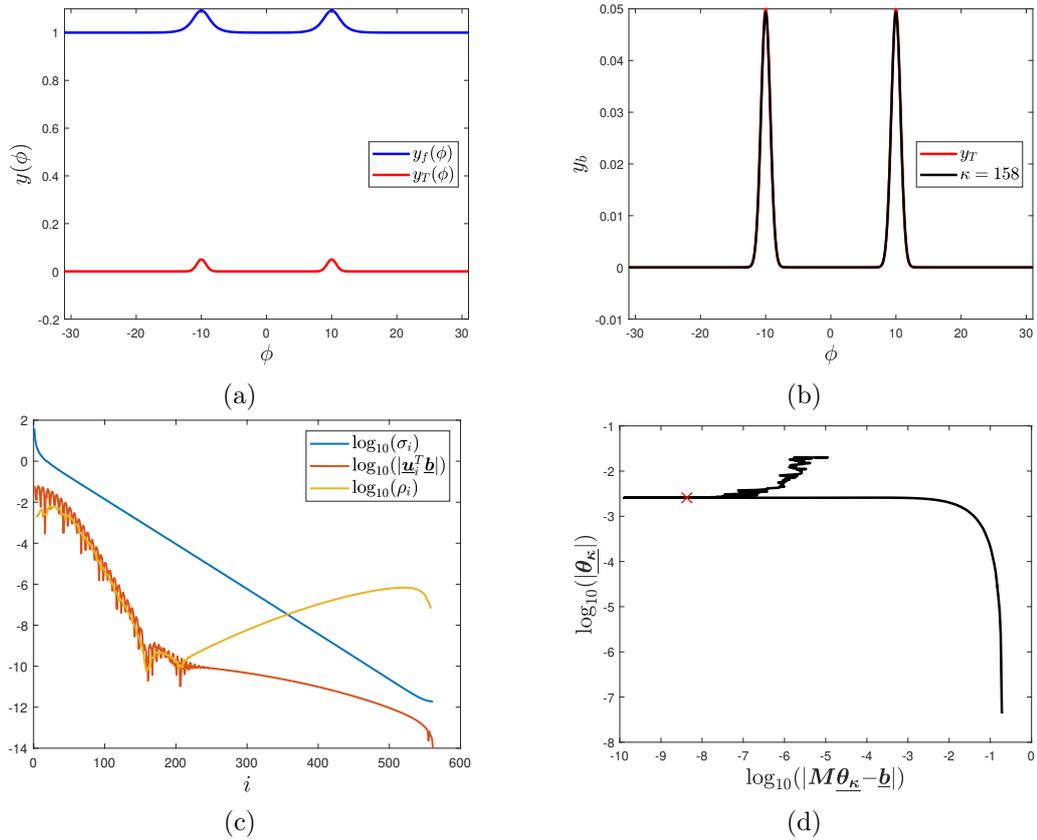


Figure 4.2.1: Inverse results for a supercritical solution with a train of waves over a Gaussian topography (4.2.1) with the parameter values $a = 0.05, b = 1, L = 20, N = 541, F = 1.2$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 158$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 158$.

size and positioning of the Gaussian dips on the topography although spurious oscillations appear in the solution.

4.2.3 Subcritical flow with trapped waves

The trapped waves on the surface of a subcritical flow studied in this section can be considered as the matching of two subcritical wave trains such that their effects on the fluid are "cancelled out" at either end of an interval on which surface waves are apparent. Looking again at Figure 4.2.2(a) one can imagine varying c in (4.2.1), changing the relative positions of the topographical features, until this cancellation of the train of waves occurs. For this purpose the separation of the

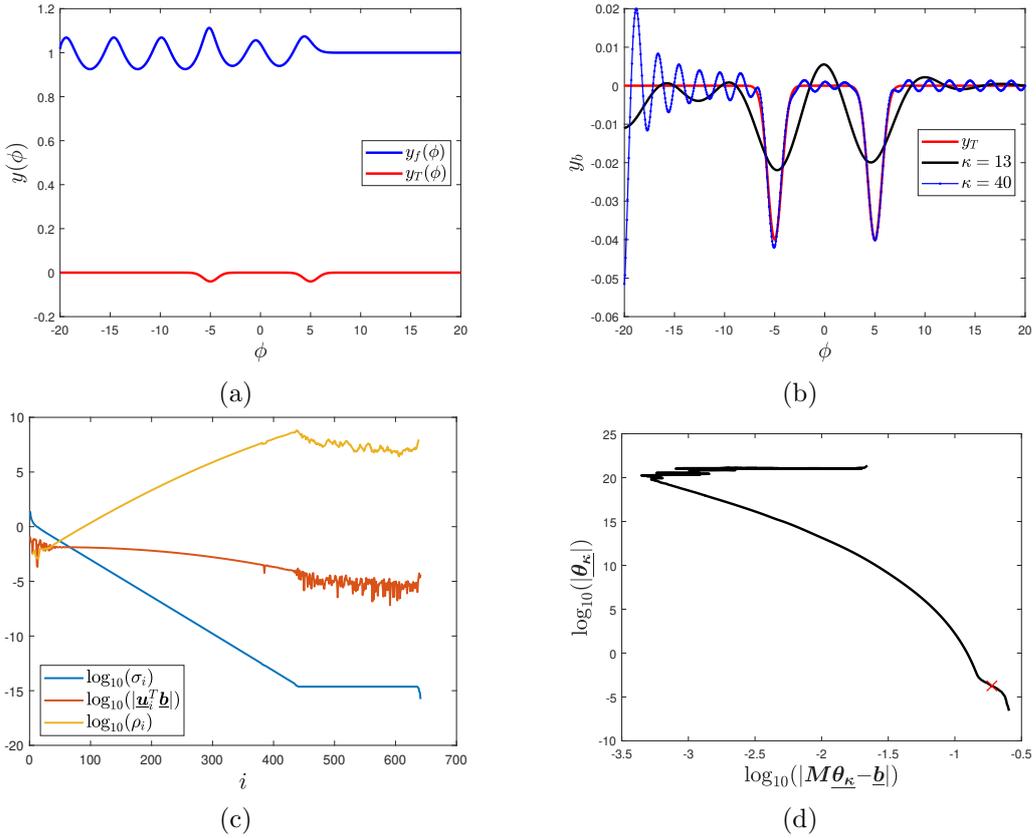


Figure 4.2.2: Inverse results for a subcritical solution with a train of waves over a Gaussian topography (4.2.1) with the parameter values $a = -0.04, b = 1, L = 20, N = 641, F = 0.8$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 13$ (black) and $\kappa = 40$ (blue) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 13$.

two topographical features is allowed to come as part of the solution, different initial guesses for the separation allows for solutions with different numbers of trapped waves (i.e., by increasing the separation we can calculate solutions with more wave peaks).

In Figure 4.2.3 we consider an example of this problem where there are two wave peaks on the free surface, for which the separation was found to be $c \approx 3.116$ (Figure 4.2.3(a)). It can be seen from Figure 4.2.3(c) that the Fourier coefficients $|\underline{u}_i^T \underline{b}|$ are on average decaying but oscillating as they do so, this is reflected in ρ_i which does not decay monotonically. The truncation rank $\kappa = 87$ has been chosen because for $i \leq 87$ the decay of ρ_i is greater on average and

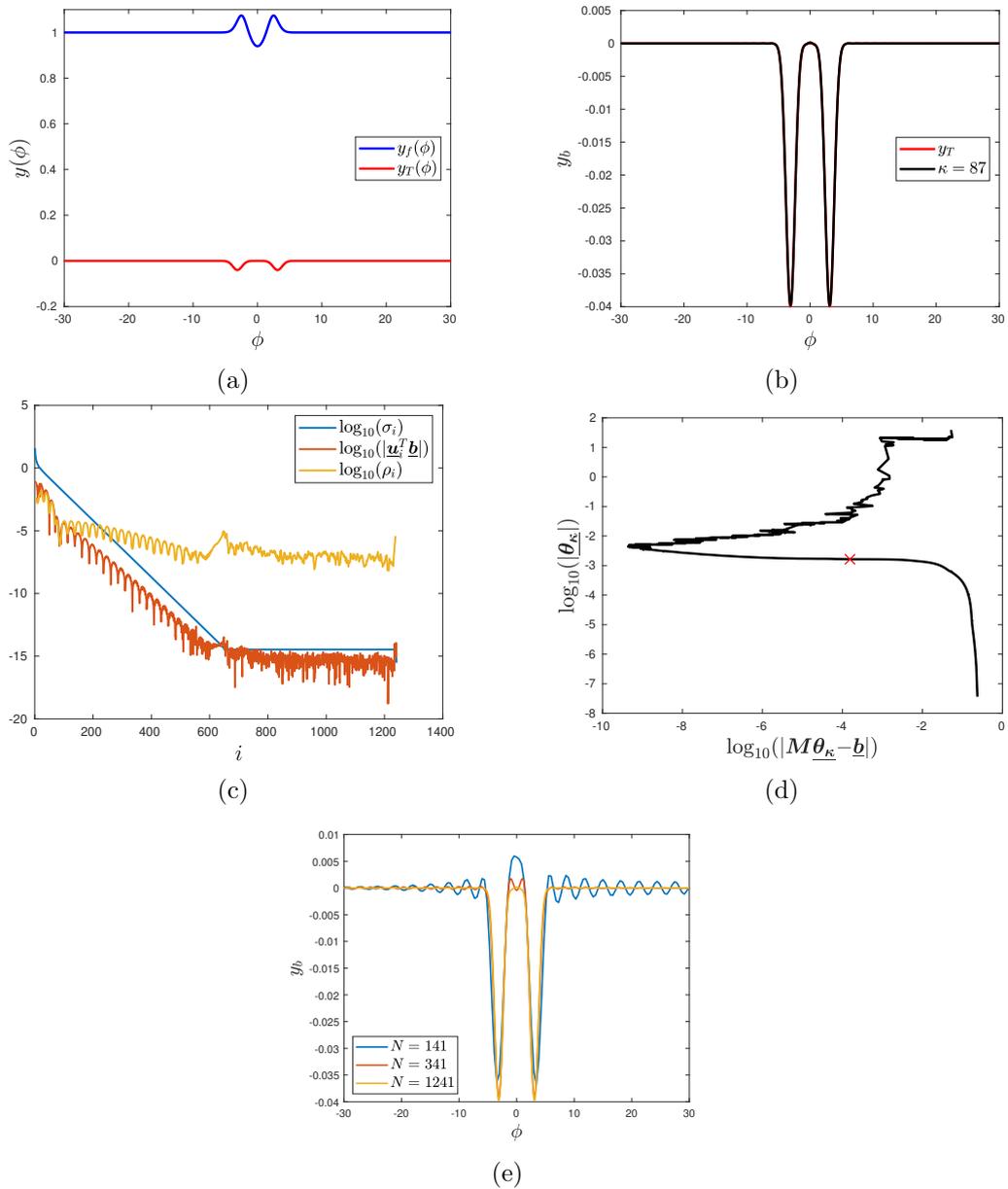


Figure 4.2.3: Inverse results for a subcritical solution with trapped waves over a Gaussian topography (4.2.1) with the parameter values $a = -0.04$, $b = 1$, $L = 30$, $N = 1241$, $F = 0.8$, $c \approx 3.116$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 87$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 87$. (e) Profiles with increasing N showing that as N is increased the inverse solution gets closer to the true solution.

the oscillations are a less dominant feature. The output for this κ is shown in Figure 4.2.3(b) and it agrees well graphically with the prescribed topography y_T . However, not visible at this scaling, small oscillations appear in the topography. Considering Figure 4.2.3(d) we see that the flat segment of the L-curve we have found to precede the corner in previous examples instead slopes upwards towards the corner, with the norm of the solution still displaying some sensitivity here to the truncation rank κ . We experimented with different values of N to study how the amplitude of these oscillations were affected; because the true topography is known in advance we are able to then select the truncation rank so as to give the “best” solution in that it minimises the norm $\|y_b - y_T\|$. In Figure 4.2.3(e) these solutions are displayed for different values of N , it can be seen that for increasing values of N the amplitude of these oscillations decreases, although it was not possible to discern a relationship between these quantities like that of the quadratic decay found previously for the unforced solitary wave previously (Figure 4.1.5(e)).

Figure 4.2.4 shows the results of repeating the procedure described above for a surface with five trapped wave peaks, for which the separation c was calculated as $c \approx 10.147$ with $N = 1241$. The discussion for this case remains almost unchanged to that of the case of two trapped wave peaks of the previous paragraph. Small undulations can be seen in the output for the topography between the two Gaussians however the size and position of the topography has been accurately captured (Figure 4.2.4(b)). It can be seen from Figure 4.2.4(e) that, as in the case of two trapped wave peaks (see Figure 4.2.3(e)), the size of these oscillations appearing in the output is related to the number of grid points used. The exact relation between these quantities was not found.

4.3 Discussion

In this chapter we have explored the ability of the TSVD method to recover the underlying topography from surface data input computed by way of the forward

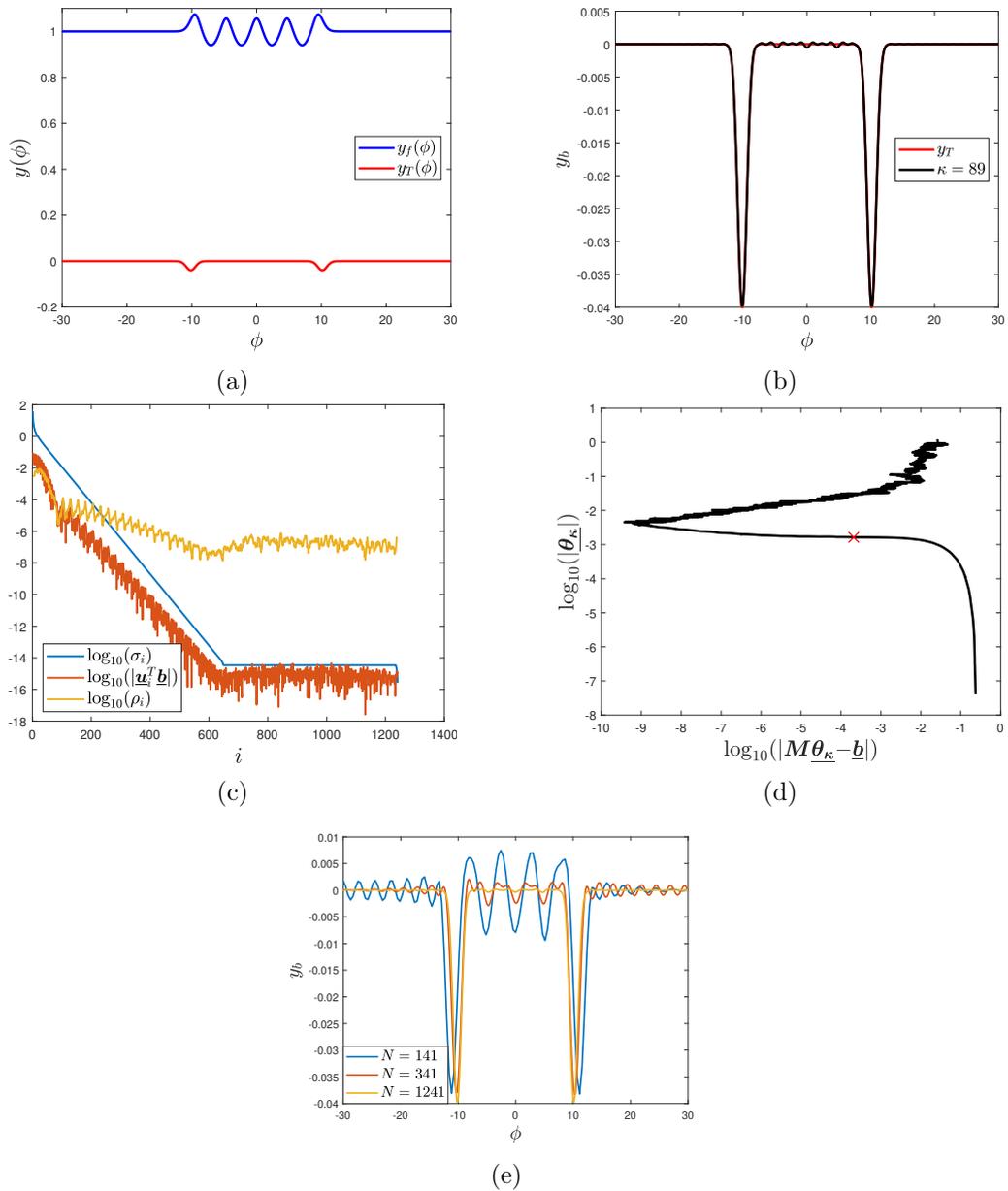


Figure 4.2.4: Inverse results for a subcritical solution with trapped waves over a Gaussian topography (4.2.1) with the parameter values $a = -0.04$, $b = 1$, $L = 30$, $N = 1241$, $F = 0.8$, $c \approx 10.147$ (a) Profiles of the prescribed topography and computed forward solution. (b) The inverse solution for the topography with $\kappa = 89$ (black) compared to the true topography (red). (c) Picard plot for use in selecting the truncation rank. (d) The L-curve for this problem with the red cross indicating the solution with $\kappa = 87$. (e) Profiles with increasing N showing that as N is increased the inverse solution gets closer to the true solution.

problem over Gaussian topography and the use of both the DPC and the L-curve in deciding a reasonable truncation rank for the system.

The method has performed well for forced supercritical flows, accurately predicting the amplitude and positioning of the topography. Selection of a truncation rank for these flows has been made by consideration of plots of ρ_i , choosing a rank as high as possible while satisfying the DPC has produced smooth solutions that lie close to the corner of the L-curve as one would like to find e.g. (Hansen, 1992a; Hansen and O’Leary, 1993). Output for the topography in supercritical flow has been more accurate as the number of points used in the model is increased, in the case of the unforced solitary wave we saw that the right-hand side of the matrix equation to which we are applying TSVD approaches zero and so accordingly does the output for the topography. This agrees with our earlier analysis concerning the unforced problem with a flat free surface; when the right-hand side is zero we found a flat topography to be the only allowable solution. Similarly to supercritical flows the results for hydraulic fall solutions have been found to be accurate reconstructions of the topography.

Application of this method to subcritical flows with surfaces featuring semi-infinite wave trains has proved more challenging. Increasing the number of mesh-points used did not lead to a reduction in the oscillations appearing in the output and these problems do not seem to satisfy the DPC. Still, outputs for the topography have seen the approximate size and positioning captured but contain errant oscillations that appear to grow as $\phi \rightarrow -\infty$. The inverse results for forward solutions that instead feature trapped waves in subcritical flow rather than semi-infinite wave trains have been markedly better behaved, with oscillations in the output again being reduced by increasing the model (and input data) resolution. This suggests that the problem lies not in subcritical flow but in the decay of the surface to the uniform stream in the far field.

Results: Inversely Retrieving Topography from Perturbed Computed Forward Solutions

In the previous chapter input to the inverse problem was obtained by first computing solutions to the forward problem. In practice, input to the inverse problem would instead constitute measurements made of the free surface and so be subject to measurement uncertainty. Given that the inverse problem has been identified to be an ill-posed problem, taking the form of solving a Fredholm integral of the first kind, it should be expected that errors introduced to the input data by this measurement uncertainty will be magnified and without appropriate regularisation it may not be possible to obtain a meaningful solution.

In Figure 5.0.1 an example of the care that must be taken is shown; Figure 5.0.1(a) displays both a solution to a forward problem (see Figure 4.1.2 for details) and plotted on top of this the same surface onto which perturbations of $\mathcal{O}(10^{-5})$ have been added. Despite the fact that these perturbations are, to graphical accuracy, imperceptible the inverse problem is affected starkly, Figure 5.0.1(b) shows the inverse results for the two surfaces using the truncation rank $\kappa = 146$.

In this chapter we will consider we will consider the problem of recovering topography from noisy input data, to simulate this data we will take computed

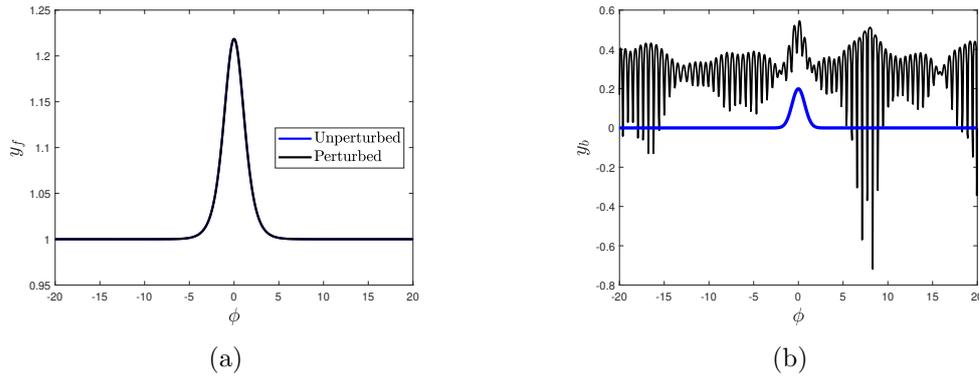


Figure 5.0.1: (a) Free surface solution for a topography of the form (5.0.1) with $a = 0.2, b = 1, F = 1.5, N = 641, L = 20$ (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-5}$ added (black). (b) TSVD solution for y_b in unperturbed problem with $\kappa = 146$ (blue), TSVD solution for y_b in perturbed problem with $\kappa = 146$ (black).

forward problem solutions and add to every interior grid point a random value independently drawn from a normal distribution with mean zero and standard deviation ϵ . The first and last grid points are left unperturbed such that the conditions of decay in the far field are satisfied, alternatively one could append to each end of the input data another data point to represent the unperturbed far field behaviour. While this is not necessary it was found that the output is drastically improved by doing so; all presented results are for input data with unperturbed endpoints. We will continue to add greater and greater amounts of noise until we can no longer retrieve the topography accurately before then discussing how taking a larger number of measurements can further push the utility of the TSVD method when given noisy data. We will at this point look at solutions for different types of free surface other than supercritical perturbation to the uniform stream type solutions past a Gaussian.

The computed forward problem solutions will be selected from those presented in the previous chapter and so the true underlying topography will be of the form

$$y_T = ae^{-(b\phi)^2} \quad (5.0.1)$$

or, for the trapped wave solutions presented previously,

$$y_T = a \left(e^{-b(\phi-c)^2} + e^{-b(\phi+c)^2} \right) \quad (5.0.2)$$

5.1 Inverse results for a single surface measurement

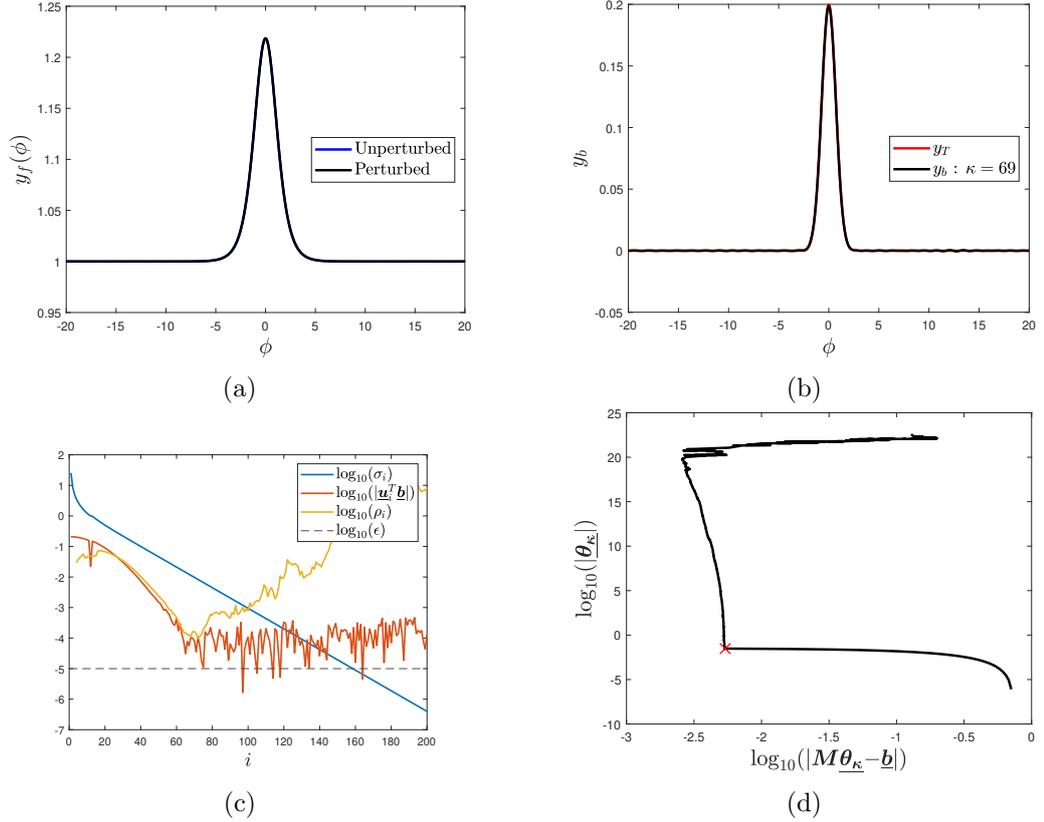


Figure 5.1.1: Inverse results for a supercritical perturbation to the uniform stream type free surface solution over a topography of the form (5.0.1) with $a = 0.2$, $b = 1$, $F = 1.5$, $N = 641$, $L = 20$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-5}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-5}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 69$.

In order to examine how the addition of noise affects the choice of truncation rank let us return to the previous example and plot the L-curve and values relevant to the DPC for this perturbed problem with $\epsilon = 10^{-5}$. From Figure 5.1.1(c) it can be seen that the addition of noise has meant that the DPC cannot be satisfied if a truncation rank is picked with $\kappa \geq 70$. Further, taking $\kappa = 69$ corresponds to a

solution on the L-curve just prior to the corner being reached (Figure 5.1.1(d)). Combined these observations suggest that $\kappa = 69$ would be a good truncation tank and indeed it can be seen from the result of doing so, in Figure 5.1.1(b), that a good approximation for the underlying topography can be recovered despite the addition of noise. There are still some small undulations present in the solution for the topography that we have not been able to eliminate but they are almost unnoticeable at this scale.

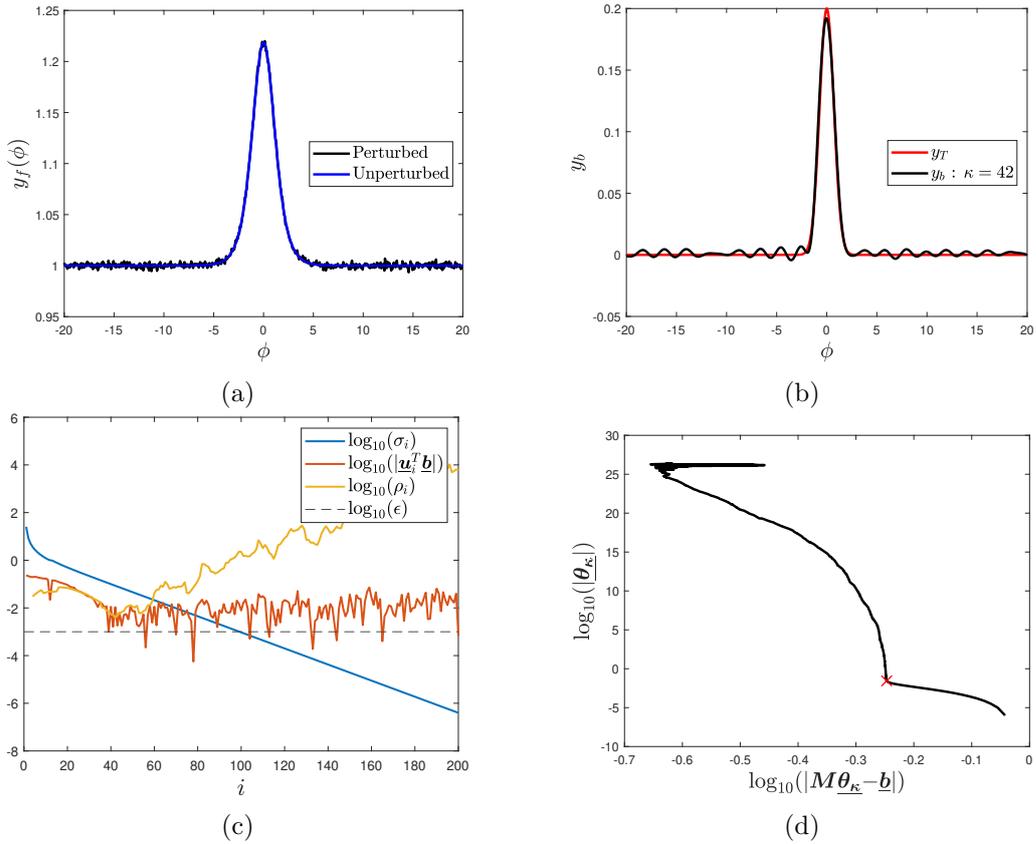


Figure 5.1.2: Inverse results for a supercritical perturbation to the uniform stream type solution over a topography of the form (5.0.1) with $a = 0.2, b = 1, F = 1.5, N = 641, L = 20$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-3}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-3}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 42$.

For the same problem let us further increase the level of noise to $\epsilon = 10^{-3}$ (Figure 5.1.2). The DPC is now only satisfied for $\kappa \leq 42$ (Figure 5.1.2(c)) with $\kappa = 42$ sitting right on the corner on the L-curve (Figure 5.1.2(d)). The horizontal segment of the L-curve seen in the previous example is now sloping upwards and

turns to become vertical at a greater value of the residual norm. Nonetheless, as can be seen in Figure 5.1.2(b) this truncation still gives a good approximation to the underlying topography using the noisy surface data.

The results of applying the same approach to a hydraulic fall solution with added noise drawn from a distribution with $\epsilon = 10^{-3}$ are shown in Figure 5.1.3. The Picard plot shown in Figure 5.1.3(c) suggests the truncation rank $\kappa = 50$ and consulting the L-curve for this problem we see that the solution for this truncation sits at the corner of the L-curve as one would hope to see. The recovered profile for the topography is contrasted to the true topography y_T in Figure 5.1.3(a) and we find that it is much like the result of the perturbation to the uniform stream solution in that the Gaussian on the topography has been accurately captured by there are small oscillations appearing on the flat regions either side that cannot be eliminated.

Increasing the standard deviation of the noise one more time to $\epsilon = 10^{-2}$ we see that in the current implementation the model is begging to struggle. Figure 5.1.4(c) shows that there is now no clear choice of truncation rank based on the DPC, although there is an interval in which ρ_i is decreasing this is for values of i at which the Fourier coefficients $|\underline{\mathbf{u}}_i^T \underline{\mathbf{b}}|$ are at the level of the noise in the system and so should not be used for assessing the DPC (see, for example, Hansen, 1990b). The L-curve has now lost any semblance of having a corner (Figure 5.1.4(d)). Despite this, in Figure 5.1.4(b) we plot the solution when the truncation rank is $\kappa = 40$ is chosen. This output still captures approximately the height and position of the Gaussian on the topography however there are large undulations in the solution and it does not correctly predict the average level of the topography away from the disturbance to be at $y_b = 0$. However, comparing the input data (Figure 5.1.4(a)) and output of this problem with that of Figure 5.0.1 which had $\epsilon = 10^{-5}$ it is clear that the quality of the output is still greatly improved by the regularisation of the problem gained by truncation.

Figure 5.1.5 shows the results of adding perturbation with standard deviation $\epsilon =$

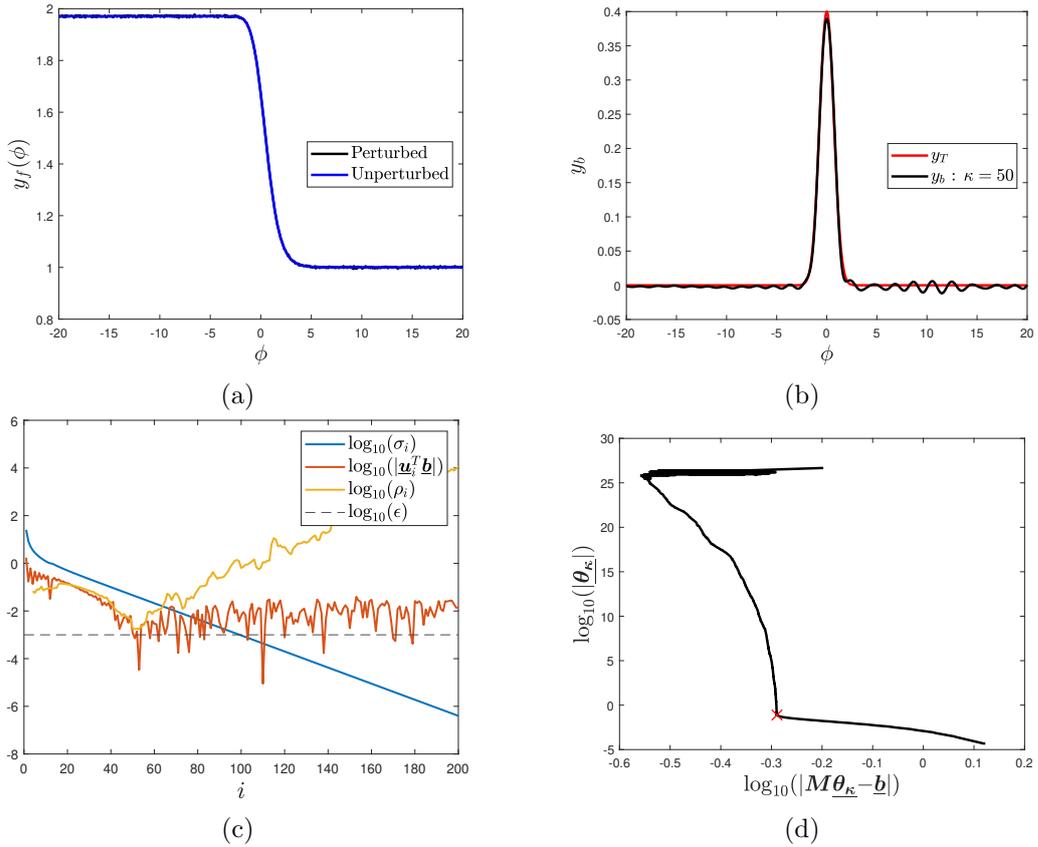


Figure 5.1.3: Inverse results for a hydraulic fall solution for a topography of the form (5.0.1) with $a = 0.4, b = 1, N = 721, L = 20, F \approx 1.617$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-3}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-3}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 50$.

10^{-2} to a hydraulic fall solution instead. Inspection of Figure 5.1.5(c) leads us to try the truncation rank $\kappa = 31$, the output for which is shown in Figure 5.1.5(b), finding that the approximate form of the topography can be recovered from one measurement.

5.2 Inverse results for multiple surface measurements

The results of the previous section relied on the results of a single measurement. Ideally in an experimental set-up the surface would be measured multiple times. The usual approach would then be to average the set of surface measurements

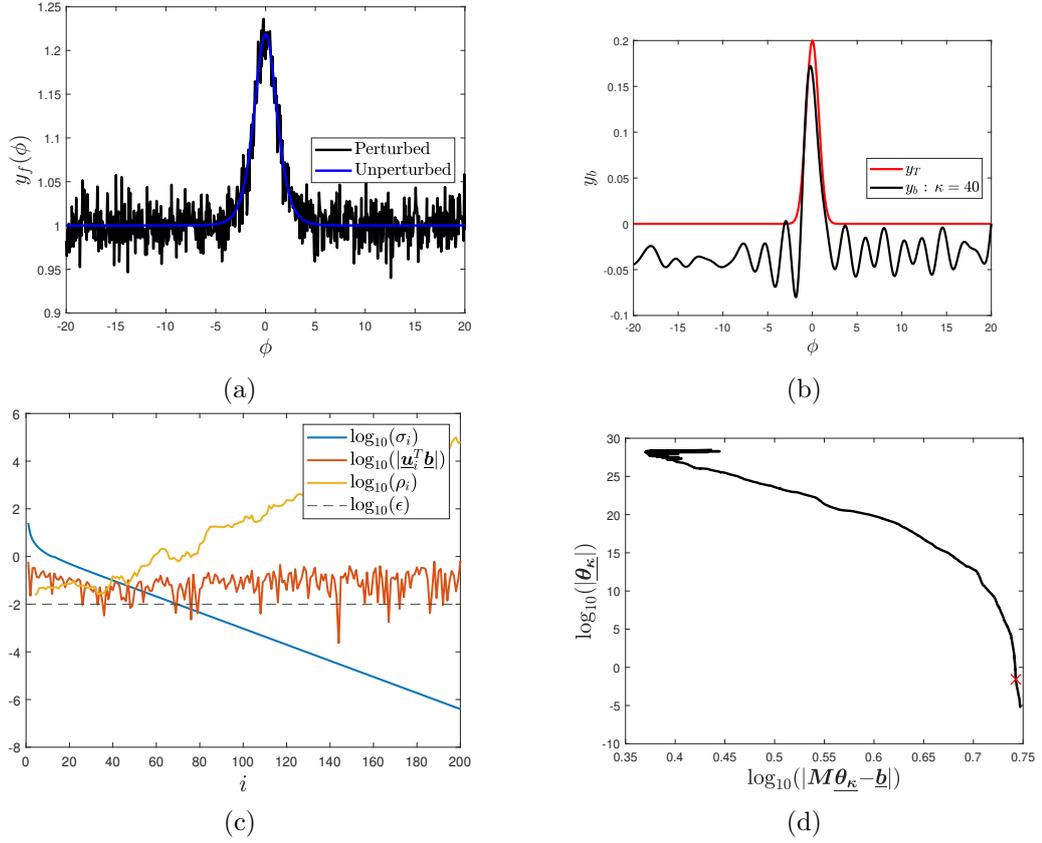


Figure 5.1.4: Results for a supercritical perturbation to the uniform stream type turbulence solution for a topography of the form (5.0.1) with $a = 0.2$, $b = 1$, $F = 1.5$, $N = 641$, $L = 20$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-2}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-2}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 40$.

in an attempt to smooth the noise and approximate the true underlying values before using this averaged surface as input. A different approach would be to instead calculate the inverse solution for each singular measurement and to then average the resultant outputs for the topography afterwards. When taking the former approach there is only one inverse problem to solve and we find that it better approximates y_T than directly inverse solving from one set of noisy data. When taking the latter approach we will have many realisations of the topography and find that while any individual realisation may be a poor estimate of the topography when repeated and averaged they better approximate y_T as more samples are taken. For these repeated measurements we choose to fix the truncation rank at $\kappa = 40$ so that we do not need to consider Picard plots for

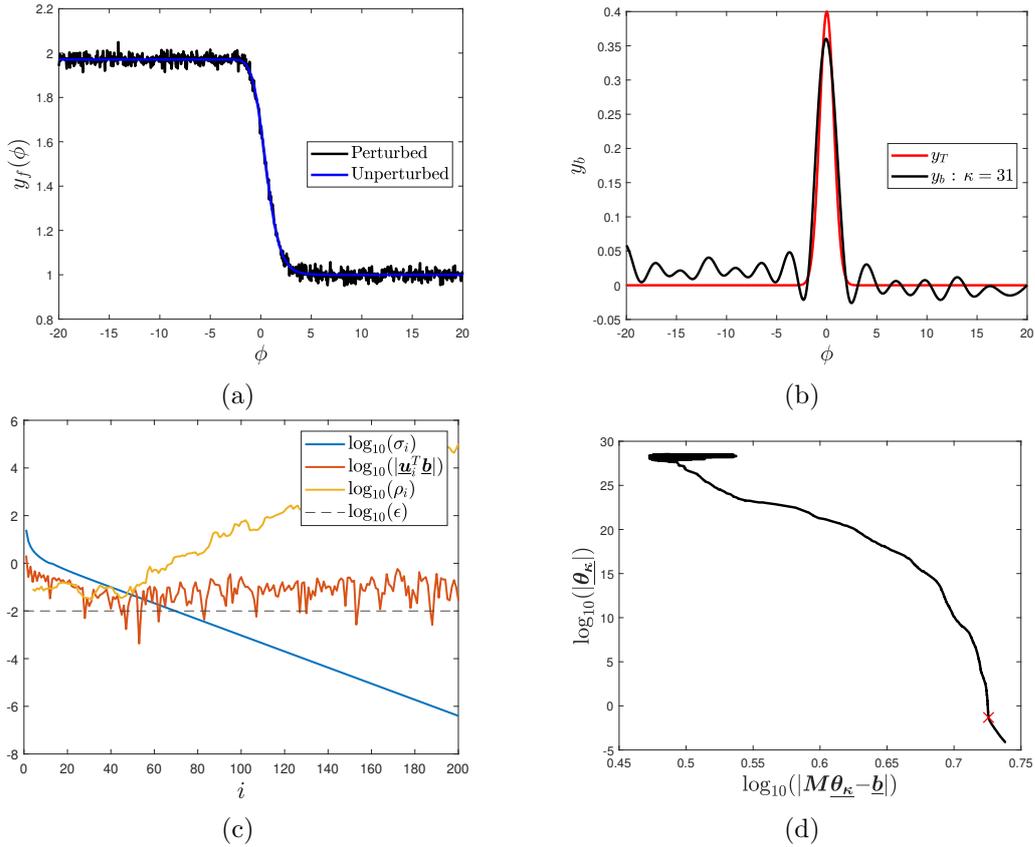


Figure 5.1.5: Results for a hydraulic fall solution for a topography of the form (5.0.1) with $a = 0.4, b = 1, N = 721, L = 20, F \approx 1.617$. (a) Free surface solution without perturbation (blue). Free surface solution with noise of standard deviation $\epsilon = 10^{-2}$ added (black). (b) The true topography (red) and the TSVD result for the perturbed problem with $\epsilon = 10^{-2}$ (c) Picard plot for perturbed problem. (d) L-curve for perturbed problem, the red cross corresponds to $\kappa = 40$.

each simulation.

Simulating 1000 measurements for the problem considered so far, a perturbation to the uniform stream solution for flow over a Gaussian, the two methods produce very similar answers for the topography (Figure 5.2.1), however it was found that computing the inverse problem for each individual measurement before averaging afterwards, i.e. taking a point-wise average of the grey curves, better approximated the true topography close to $\phi = 0$ than first averaging the noisy surface data and then solving the inverse problem once.

Similarly, Figure 5.2.2 shows the results for simulating 1000 noisy measurements of a solitary wave solution. We find again that the two averaging strategies

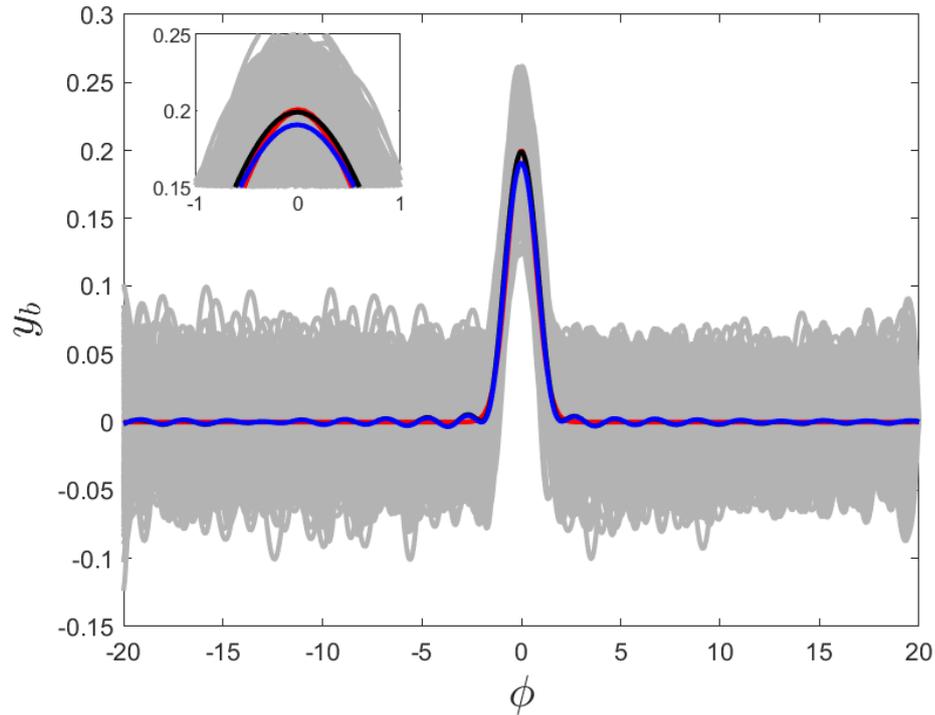


Figure 5.2.1: Inverse results for a perturbation to the uniform stream solution over a Gaussian topography (5.0.1) based on 1000 simulated measurements with the parameter values $a = 0.2$, $b = 1$, $L = 20$, $N = 641$, $F = 1.5$, $\kappa = 40$, $\epsilon = 10^{-2}$. Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue).

provide very similar results almost everywhere but deviate from one another close to $\phi = 0$. Here we find that first averaging the surface data and then solving the inverse problem using the averaged data as input better captures the true topography, however both strategies give results that closely follow the true topography.

Considered next is the case of a hydraulic fall (Figure 5.2.3). The two strategies once again are in close agreement with one another for $\phi > 1$, it is only close to the peak of the wave that they begin to deviate when following the curves from right to left. These two curves rejoin briefly on the upstream side of the obstacle but they then deviate again, with the solution found by averaging all noisy outputs for topography following the same pattern as the output for first averaging all

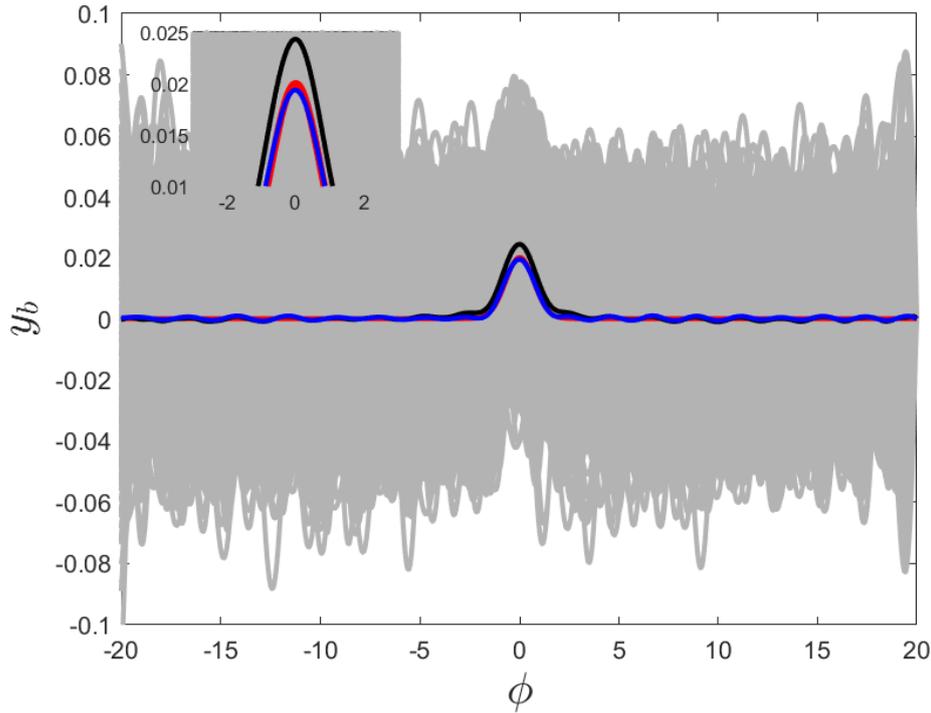


Figure 5.2.2: Inverse results for a perturbation to the solitary wave solution over a Gaussian topography (5.0.1) based on 1000 simulated measurements with the parameter values $a = 0.02, b = 1, L = 20, N = 641, F = 1.1, \kappa = 40, \epsilon = 10^{-2}$. Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue).

surfaces and then inverse solving but translated upwards for $\phi < -1$. While averaging the noisy topographies that are output for each noisy surface appears to better capture the peak height of the topographical disturbance it incorrectly predicts a change in the level of the topography upstream with the average level of the topography as predicted being non-zero. Looking at the individual noisy solutions on the right-hand side they are distributed somewhat evenly about the line $y = 0$, leading to the mean being around zero, however on the left-hand side there is a positive bias with fewer noisy solutions reaching below the zero line. While the results are not displayed here this does not occur if the simulations are repeated with $\epsilon = 10^{-3}$ or lower.

Inverse solutions based on noisy measurements of trapped waves in subcritical

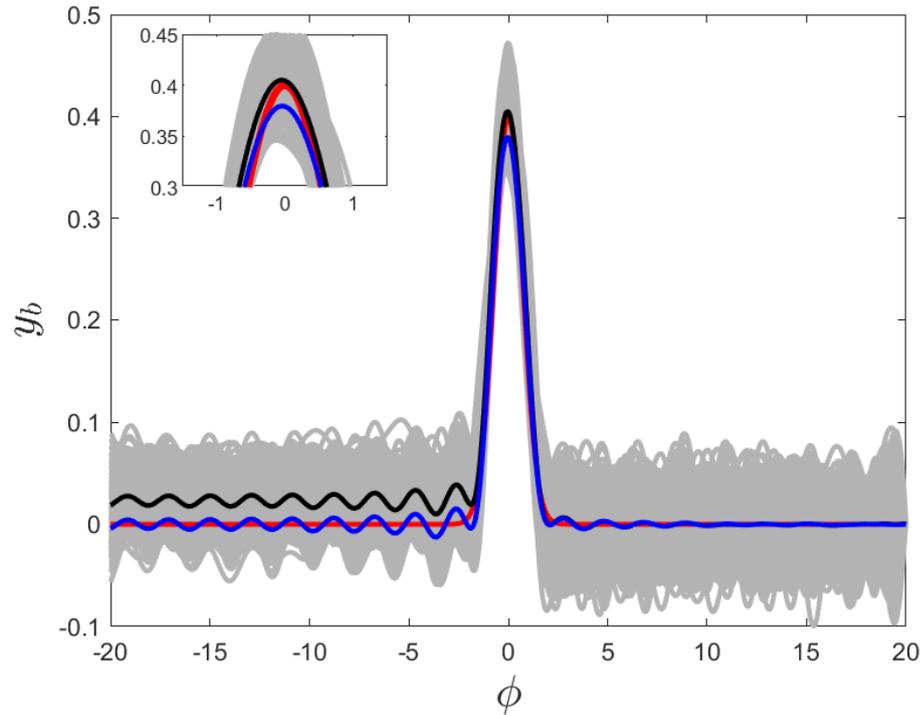


Figure 5.2.3: Inverse results for a perturbation to the hydraulic fall solution over a Gaussian topography (5.0.1) based on 1000 simulated measurements with the parameter values $a = 0.4$, $b = 1$, $L = 20$, $N = 721$, $F \approx 1.617$, $\kappa = 40$, $\epsilon = 10^{-2}$. Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue).

flow are shown in Figure 5.2.4. Once again a good estimate of the topography can be obtained employing either strategy and, while there is a separation between the two solutions for the topography, the two curves track one another's path.

Finally let us consider adding perturbation onto the uniform free stream and performing the same process of averaging. Figure 5.2.5(a) shows the inverse result to a single perturbed measurement of the surface, while the maximum disturbance to the topography is found to be small there are oscillations over the entire topography which does not accurately reflect the true solution of $y_T = 0$. Figure 5.2.5(b) shows the results for the topography when repeated measurements are made, both averaging strategies produce a result far better than that of a single simulation, the amplitude of oscillations has been decreased and the mean

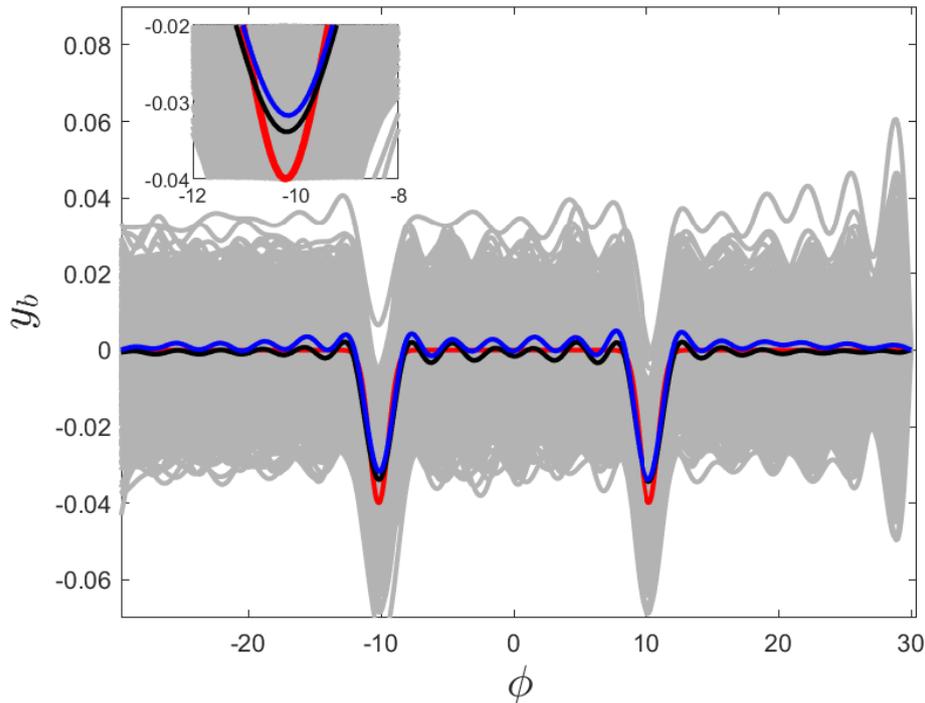


Figure 5.2.4: Inverse results for a perturbation to a trapped wave solution between two Gaussian topographical features (5.0.2) based on 1000 simulated measurements with the parameter values $a = -0.04, b = 1, L = 30, N = 641, F = 0.8, \kappa = 40, \epsilon = 10^{-2}$. Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue).

level is now at $y = 0$.

5.3 Discussion

In this chapter we have seen that even a small amount of noise added to the free surface can drastically alter the output for the topography as compared to the unperturbed problem if care is not paid to properly selecting the truncation rank so as prevent magnification of the error. When this error is small we are able to recover the topography simply by proper selection of the truncation rank. When the added perturbation is so great that an output for the topography may not resemble the true topography we have been able to improve our estimate by

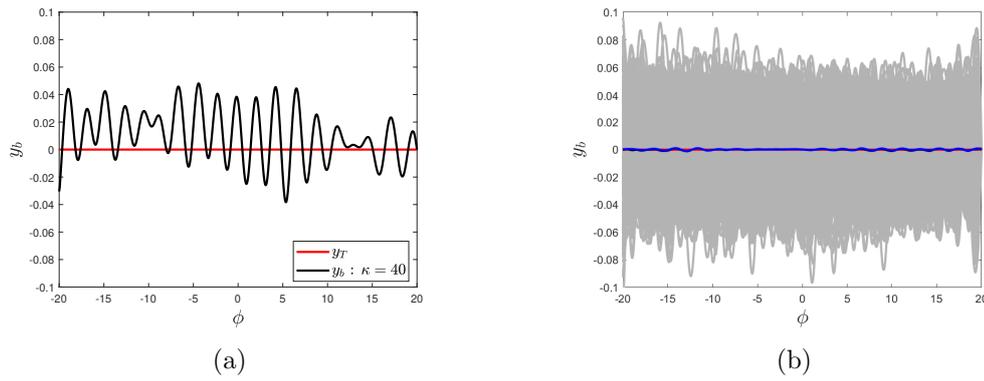


Figure 5.2.5: Inverse results for a perturbed uniform free stream with $F = 1.2$, $L = 20$, $N = 641$, $\kappa = 40$, $\epsilon = 10^{-2}$ (a) The true topography (red) and the TSVD result for a single realisation of the perturbed problem. (b) Plotted are the true topography y_T (red), the inverse solutions (grey) to 1000 individual noisy measurements, the solution when all 1000 individual realisations for the topography are averaged point-wise (black) and the inverse solution when all simulated surface measurements are first averaged to give input data to a singular inverse problem (blue).

taking further surface measurements and either repeatedly running the heavily-truncated inverse problem and averaging the resultant outputs or by averaging the surface data first.

Inverse Results for Prescribed Surfaces

In this chapter the free surface y_f will be prescribed directly rather than calculated as the solution to a forward problem. This enables for the inverse solution space to be explored in a way that wouldn't be easily facilitated by first having to compute forward solutions. We will first look at prescribing a Gaussian surface and then explore the response of the inverse solutions when the Froude number and the amplitude are changed, comparing these results to those of the fKdV. We will quantify the response of the inverse solution by tracking the norm of the topography and also the height achieved by the point $y_b(0)$.

In the second section we will prescribe a trapped wave on the free surface then analyse how the inverse solutions respond to the frequency of the wave on the surface being varied by taking Fourier transforms and considering the energy contained in each mode. Again we will use the fKdV as a guide and as a point of comparison.

6.1 Inverse results for a Gaussian free surface

For this section we take the free surface as prescribed in the form

$$y_f = 1 + \alpha e^{-(\beta\phi)^2}, \quad (6.1.1)$$

where α and $\beta > 0$ are parameters that control the amplitude and width of the free surface disturbance respectively.

6.1.1 Response of the topography to changing Froude

Figure 6.1.1 displays the fully nonlinear inverse results for the topography given a fixed free surface for different values of the Froude number. Each panel corresponds to a different value of α : the left-hand panels have $\alpha > 0$ with the free surface surface being elevated; the right-hand panels have $\alpha < 0$ with the free surface depressed. Taking first the case of $\alpha > 0$, it can be seen that the presented solutions at $F = 0.8$ for y_b all have $y_b(0) < 0$. As the Froude number is increased this value increases, eventually becoming positive. In the panels where $\alpha < 0$ the opposite is true; the solutions for $F = 0.8$ have a $y_b(0) > 0$ and this value then decreases and becomes negative for increasing F . Integration of the fKdV shows that for $F = 1$ steady solutions can only exist for a forcing $f(x)$ that satisfies $\int_{-\infty}^{\infty} f(x) dx < 0$ (see Keeler, Binder, and Blyth, 2017); every fully nonlinear solution found here satisfies this inequality. The norm $\|y_b\|$ provides a measure of the total displacement of material to construct the topography relative to the undisturbed level $y = 0$. All solutions displayed have this total displacement initially decrease before then starting to increase again as F is further increased.

The fKdV provides a model for predicting how the quantities $\|y_b\|$ and $y_b(0)$ vary with F in the fully nonlinear problem. The inverse forcing predicted by the fKdV

for a free surface given by (6.1.1), where ϕ is replaced with x , is

$$f(x) = \alpha e^{-(\beta x)^2} \left(2(F - 1) + \frac{2\beta^2}{3}(1 - 2\beta^2 x^2) - \frac{3\alpha}{2} e^{-(\beta x)^2} \right). \quad (6.1.2)$$

From this equation the central value $f(0)$ is easily obtained as

$$f(0) = \alpha \left(2(F - 1) + \frac{2\beta^2}{3} - \frac{3\alpha}{2} \right). \quad (6.1.3)$$

This is a linear equation in F and predicts that, for a fixed α and β , the height of the centre of the topography will change sign as the curve passes $f(0) = 0$. Solving for the Froude number at which this occurs provides, for $\alpha \neq 0$,

$$F_0 = 1 + \frac{9\alpha - 4\beta^2}{12}. \quad (6.1.4)$$

The norm of the forcing, $\|f\|$, for a Gaussian free surface (see Appendix A) is

$$\|f\| = \left| \alpha \pi^{\frac{1}{4}} \right| \sqrt{\frac{4(F - 1)^2}{\beta \sqrt{2}} + \frac{4\beta(F - 1) + \beta^3}{3\sqrt{2}} - \frac{6\alpha(F - 1)}{\beta \sqrt{3}} - \frac{4\alpha\beta}{3\sqrt{3}} + \frac{9\alpha^2}{8\beta}}, \quad (6.1.5)$$

excluding again the case $\alpha = 0$, for which both $\|f\|$ and $f(0)$ vanish, a minimum occurs in the norm of the forcing, for fixed α and β , when the Froude number takes the value

$$F_{min} = 1 + \frac{\alpha\sqrt{6}}{4} - \frac{\beta^2}{6}. \quad (6.1.6)$$

Figure 6.1.2(a) compares the results of (6.1.5) for the forcing in the fKdV (dashed lines) to computed values for the fully nonlinear inverse problem (solid lines) for different values of $\alpha > 0$ as the Froude number is varied. The fKdV closely predicts the value of the Froude number for which the minimum of $\|y_b\|$ occurs in the fully nonlinear problem. The actual value the norm takes is well approximated near this minimum and for values of F close to, but higher than, F_{min} . The prediction of the fKdV for the value of $\|y_b\|$ diverges much more rapidly from the true value for $F < F_{min}$ and underestimates the true value in this range. Figure 6.1.2(b) plots instead the computed values $y_b(0)$ (solid lines) and the

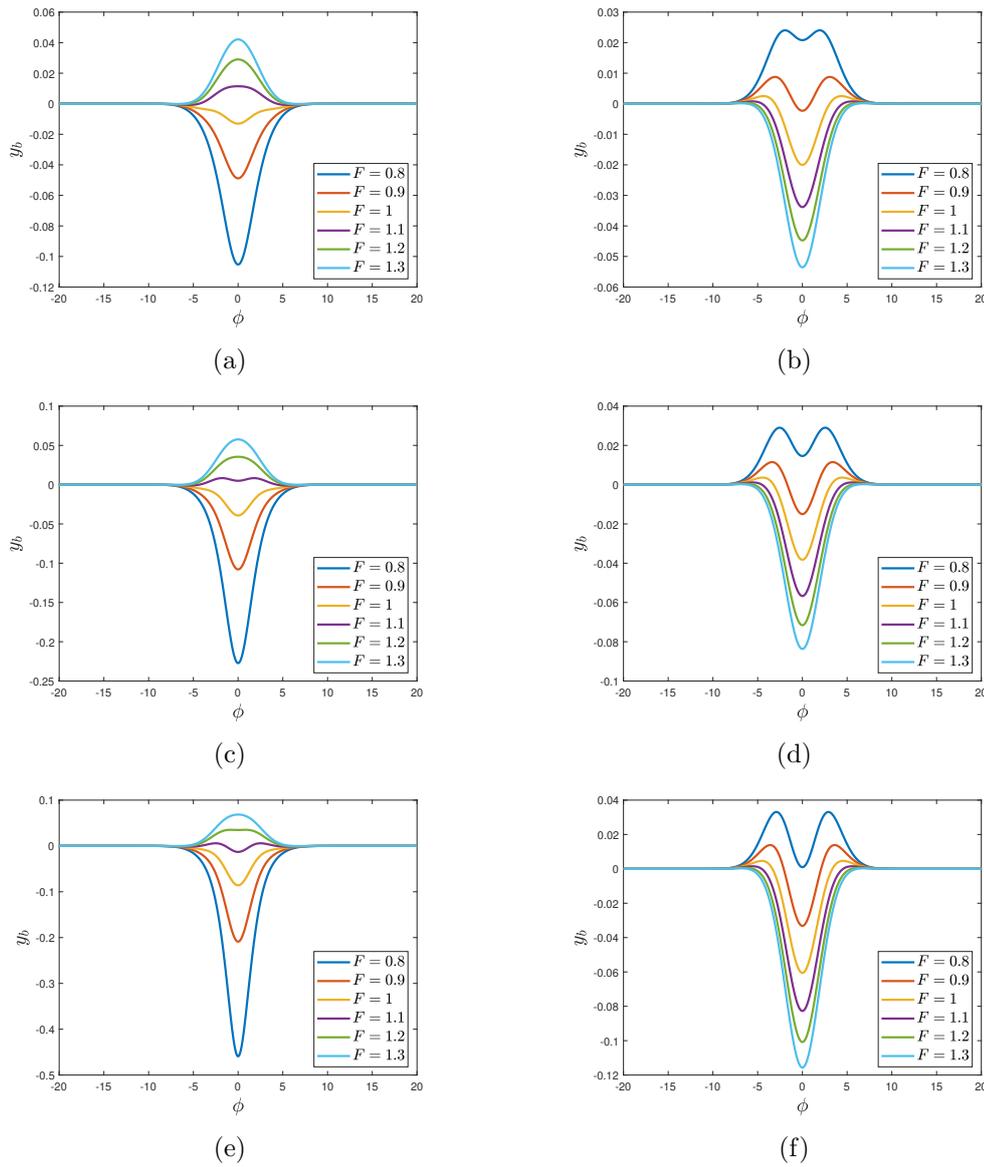


Figure 6.1.1: Inverse results for the topography as the Froude number is varied with parameters $\beta = 0.3, L = 20, N = 641$ (a) $\alpha = 0.1, \kappa = 92$. (b) $\alpha = -0.1, \kappa = 92$. (c) $\alpha = 0.15, \kappa = 104$. (d) $\alpha = -0.15, \kappa = 104$. (e) $\alpha = 0.2, \kappa = 110$. (f) $\alpha = -0.2, \kappa = 96$.

values predicted by (6.1.3) against the Froude number. The value F_0 given by (6.1.4) predicts well the Froude number at which $y(0)$ changes sign in the fully nonlinear problem with the greatest discrepancy between the predicted value and the calculated value being less than 0.004 for these amplitudes. As was the case for the norm $\|y_b\|$ the fully nonlinear results are, for these parameters and this range of F , much better predicted for $F > F_0$ than they are for $F < F_0$.

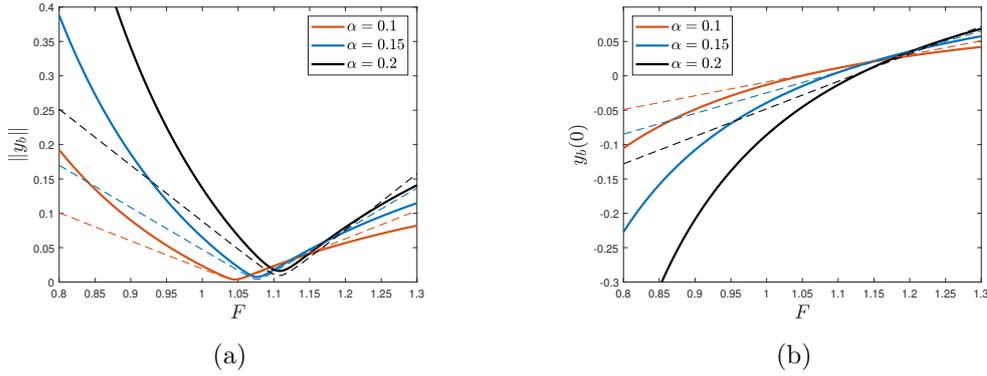


Figure 6.1.2: Inverse solutions for a fixed free surface as the Froude number is varied with $\beta = 0.3$, $L = 20$, $N = 641$, $\kappa = 92$ (a) The norm of the inverse solution against F . (b) The central value $y_b(0)$ of the inverse solution against F .

Figure 6.1.3 repeats the above comparison of the norm and central values of the fKdV forcing (dashed lines) and computed fully nonlinear (solid lines) as the Froude number is varied, this time taking $\alpha < 0$. In Figure 6.1.3(a) we can see that, similar to the case of $\alpha > 0$, the fKdV gives a good approximation for the position of the minimum in the norm $\|y_b\|$, however now the behaviour seen in Figure 6.1.2(a) is reversed with better agreement in general for $F < F_{min}$. There is an intersection between the norms of the fKdV and fully nonlinear for a value of F larger than F_{min} however beyond this value the two solutions diverge from one another with the fKdV now overestimating the value of the norm. In Figure 6.1.3(b) it is shown how the predictions of the fKdV for the height of the centre of the topography compare to those of the fully nonlinear. The fKdV does relatively well at approximating this value for $F < F_0$ and the quality of this prediction breaks down as F is increased too far. There is a maximum discrepancy in the predicted values of F_0 about four times greater than in the case of positive forcing, with the largest difference here being approximately 0.017. We note that in all panels of both Figure 6.1.2 and Figure 6.1.3 the difference between the fully nonlinear and fKdV solutions at $F = 1$ becomes smaller as the absolute value of the surface amplitude is reduced, this is an expected result as the fKdV is derived by an expansion around $F = 1$ for small amplitude waves.

Returning to Figure 6.1.1 we focus on the curves corresponding to $F = 1.1$ in

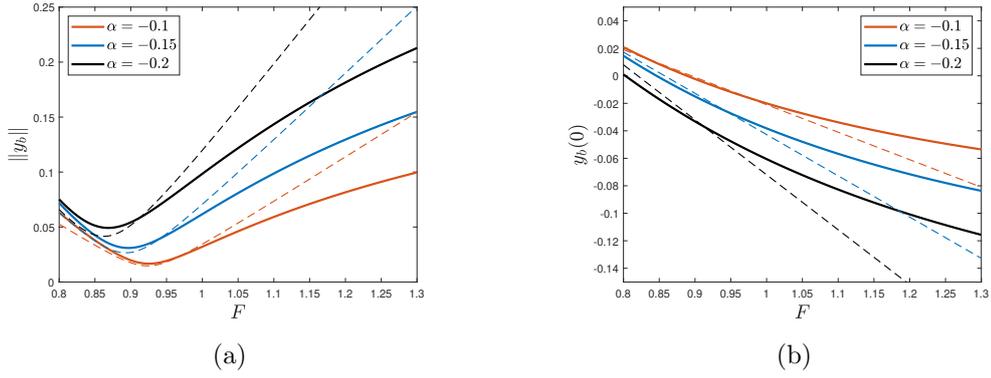


Figure 6.1.3: Inverse solutions for a fixed free surface as the Froude number is varied with $\beta = 0.3$, $L = 20$, $N = 641$, $\kappa = 92$ (a) The norm of the inverse solution against F (b) The central value $y_b(0)$ of the inverse solution against F .

the left-hand column of panels and to $F = 0.9$ in the right-hand panels. Moving downwards through the panels corresponds to an increase in the absolute value of the amplitude of the prescribed surface. The solution for $F = 1.1$ with $\alpha = 0.1$ (Figure 6.1.1(a)) is an elevation in the topography with one maximum which occurs at $\phi = 0$ with $y_b(0) > 0$. However, for the same Froude number the solution with $\alpha = 0.2$ now has a local minimum at $\phi = 0$ and one maximum on each side; at this minimum we now have $y_b(0) < 0$. The solution for $F = 0.9$ with $\alpha = 0.1$ shown in Figure 6.1.1(b) has a local minimum at $\phi = 0$ with $y_b(0) < 0$. Looking down the column of panels it can be seen that increasing the amplitude is causing $y_b(0)$ to become more negative. We know that as $\alpha \rightarrow 0$ we will have $y_b(0) \rightarrow 0$ because $\alpha = 0 \implies y_b(\phi) = 0$, but not if it will become positive before going to zero.

We look again at the expression for $f(x = 0)$ given by (6.1.3), we now fix F and β , considering $f_0 = f(x = 0)$ to be a function of the amplitude we note that it is a quadratic in α with a negative leading coefficient and with roots

$$\alpha_1 = 0 \quad \text{and} \quad \alpha_2 = \frac{4(F-1)}{3} + \frac{4\beta^2}{9}.$$

There are three distinct cases to consider here which depend on the relative size of the terms in α_2 . For each of these cases an example of the $(\alpha, f(0))$ curve (dotted line) is shown and compared to the fully nonlinear $(\alpha, y_b(0))$ curve (solid

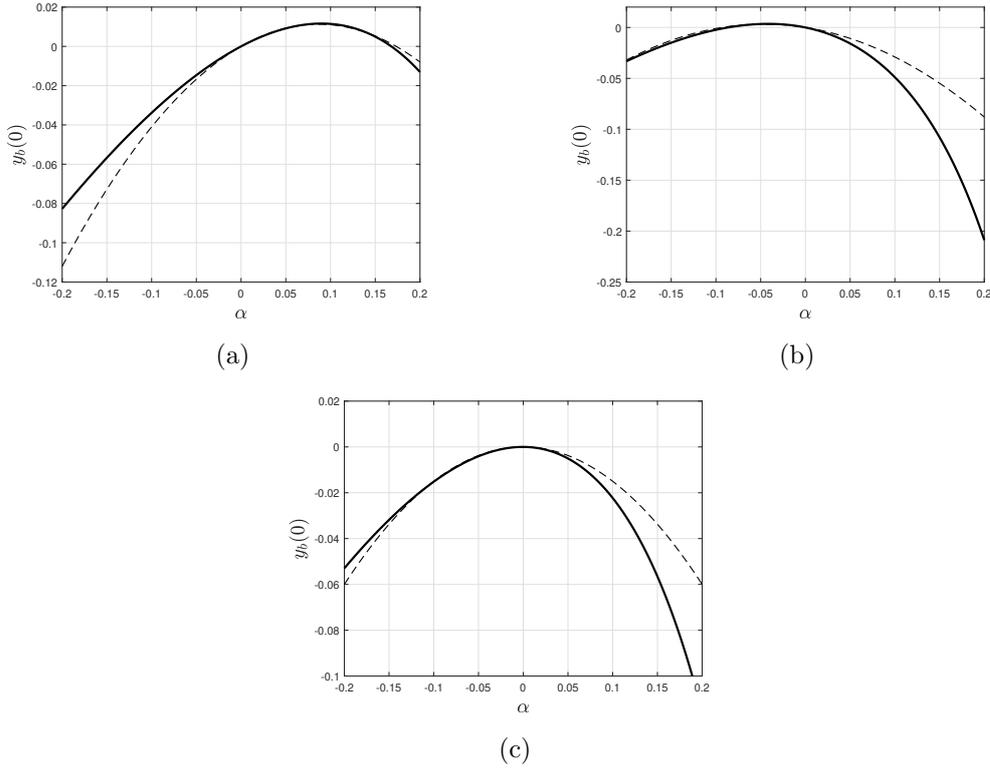


Figure 6.1.4: Plots of the central response of the inverse solution for a fixed Froude number as the amplitude of the prescribed surface is varied. $\beta = 0.3, L = 20, N = 641, \kappa = 58$ (a) $F = 1.1 > 1 - \frac{\beta^2}{3}$ (b) $F = 0.9 < 1 - \frac{\beta^2}{3}$ (c) $F = 0.97 = 1 - \frac{\beta^2}{3}$.

line) in Figure 6.1.4. The three cases are as follows:

- Case 1: $F > 1 - \frac{\beta^2}{3} \implies \alpha_2 > 0$. We find in this case two distinct values of α satisfying $f(0) = 0$. If $0 < \alpha < \alpha_2$ then we must have $f(0) > 0$ and if $\alpha < 0$ or $\alpha > \alpha_2$ then $f(0) < 0$ (Figure 6.1.4(a)).
- Case 2: $F < 1 - \frac{\beta^2}{3} \implies \alpha_2 < 0$. We find again in this case two distinct values of α satisfying $f(0) = 0$. In this case we find that the topography has $f(0) > 0$ if $\alpha_2 < \alpha < 0$ and $f(0) < 0$ if $\alpha < \alpha_2$ or $\alpha > 0$ (Figure 6.1.4(b)).
- Case 3: $F = 1 - \frac{\beta^2}{3} \implies \alpha_2 = 0$. Now there is only one value, $\alpha = 0$, satisfying $f(0) = 0$. In this case for any non-zero α the fKdV predicts that $f(0) < 0$ (Figure 6.1.4(c)).

Figure 6.1.4(a) shows good agreement between results predicted by the fKdV and the fully nonlinear calculations; finding that $y_b(0) > 0$ for $0 < \alpha < 0.168$ where

the fKdV suggested $0 < \alpha < 0.173$. Similarly, Figure 6.1.4(a) finds $y_b > 0$ for $-0.086 < \alpha < 0$ where the fKdV predicts $-0.093 < \alpha < 0$. For the values of α tested there is very good agreement for $\alpha < 0$. Figure 6.1.4(c) shows that the fKdV has correctly predicted that there is not a value α for which $y_b(0) > 0$.

Figure 6.1.5 plots the corresponding norms to their respective panels in Figure 6.1.4. The behaviour of each of these cases is quite different. Figure 6.1.5(a) shows that for $\alpha < 0$ the fKdV overestimates the value of $\|y_b\|$ until the two curves intersect at the global minimum positioned at $(\alpha, \|y_b\|) = (0, 0)$. For $\alpha > 0$ the fKdV then underestimates $\|y_b\|$, both curves increase with α until $\alpha \approx 1$ before falling again with increasing α to a local minimum that occurs for $\alpha \approx 1.79$. All cases see a global minimum at the origin, however this is the only case for which a minimum occurs for a different value of α . This happens close to the value $\alpha = 1.68$ for which $y_b(0)$ was found to change sign from positive to negative. Figure 6.1.4(b) has the fKdV globally (except for at the origin) underestimate the true value of $\|y_b\|$, although for $\alpha < 0$ the two curves closely match. There are no minima present apart from the origin and any reduction in the value of $|\alpha|$ causes a reduction in $\|y_b\|$. Figure 6.1.4(c), the case for which $y_b(0) \leq 0$, appears almost as a hybrid of the behaviour of the previous two panels, similar to Figure 6.1.4(b) for $\alpha < 0$ and Figure 6.1.4(b) for $\alpha > 0$.

6.2 Inverse results for artificial trapped waves

6.2.1 KdV

The forced KdV is given by

$$\eta_{xxx} + 9\eta\eta_x - 6\mu\eta_x = -3f_x, \quad (6.2.1)$$

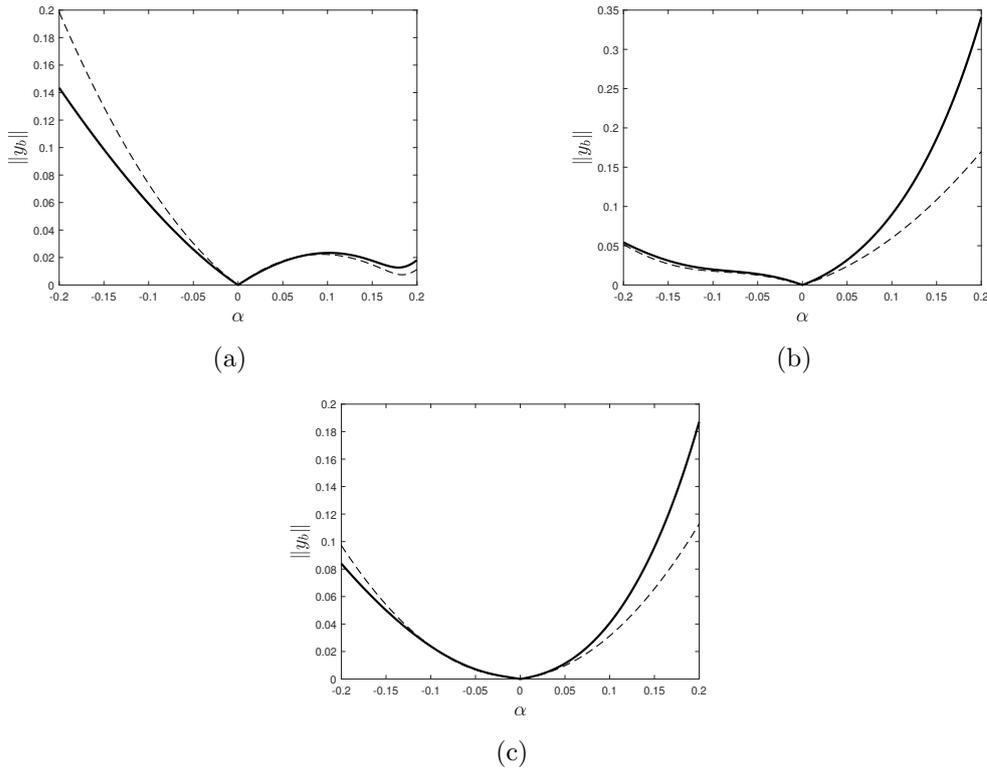


Figure 6.1.5: Plots of the norm of the inverse solution for a fixed Froude number as the amplitude of the prescribed surface is varied. $\beta = 0.3$, $L = 20$, $N = 641$, $\kappa = 58$ (a) $F = 1.1$ (b) $F = 0.9$ (c) $F = 0.97$.

We shall consider inverse solutions to the fKdV where the surface, η , takes a sinusoidal form given by

$$\eta = A \cos(\nu x) . \quad (6.2.2)$$

Integrating the fKdV and then rearranging we arrive at

$$f = -\frac{1}{3}\eta_{xx} - \frac{3}{2}\eta^2 + 2\mu\eta + C , \quad (6.2.3)$$

where previously we have set the constant $C = 0$ by applying the condition of the decay of η as the flow approaches the uniform stream in the far field; however, this condition does not now apply. Substituting for η into the above expression and expanding the squared term by way of the double angle formula then leads

to the expression for the forcing being written as

$$f(x) = A \left(2\mu + \frac{\nu^2}{3} \right) \cos(\nu x) - \frac{3A^2}{4} \cos(2\nu x) + \tilde{C} \quad (6.2.4)$$

where the new constant \tilde{C} is defined as $\tilde{C} = C - \frac{3A^2}{4}$. It is at this point we set $\tilde{C} = 0$, this can be performed without loss of generality as solutions will differ only by the addition of a uniform height level to the surface (Binder, Blyth, and Balasuriya, 2015). As such we have

$$f(x) = A \left(2\mu + \frac{\nu^2}{3} \right) \cos(\nu x) - \frac{3A^2}{4} \cos(2\nu x). \quad (6.2.5)$$

There are two modes to the topography and so if we Fourier transform the topography we expect to see peaks at the (angular) frequencies ν and 2ν with the energies we shall call c_1 and c_2 respectively, which are given by:

$$c_1 = \left| A \left(2\mu + \frac{\nu^2}{3} \right) \right|; \quad c_2 = \left| \frac{3A^2}{4} \right|. \quad (6.2.6)$$

We shall refer to the mode with energy c_1 the first mode and that with energy c_2 the second mode.

Which of these modes will be most energetic and at what frequencies can be separated into the following cases (a schematic diagram is shown in Figure 6.2.1):

1. Supercritical flow, i.e. $\mu > 0$

- If $\mu > \frac{3|A|}{8}$ the most energetic mode is c_1 for all values of ν .
- If $\mu \leq \frac{3|A|}{8}$ then c_1 and c_2 intersect at $\nu = \sqrt{\frac{9|A|}{4} - 6\mu}$. The most energetic mode is c_1 for $\nu > \sqrt{\frac{9|A|}{4} - 6\mu}$ and c_2 for $\nu < \sqrt{\frac{9|A|}{4} - 6\mu}$.

2. Critical flow, i.e. $\mu = 0$

- An intersection always exists at $\nu = \sqrt{\frac{9|A|}{4}}$, the most energetic mode is c_1 for $\nu > \sqrt{\frac{9|A|}{4}}$ and c_2 for $\nu < \sqrt{\frac{9|A|}{4}}$.

3. Subcritical flow, i.e. $\mu < 0$

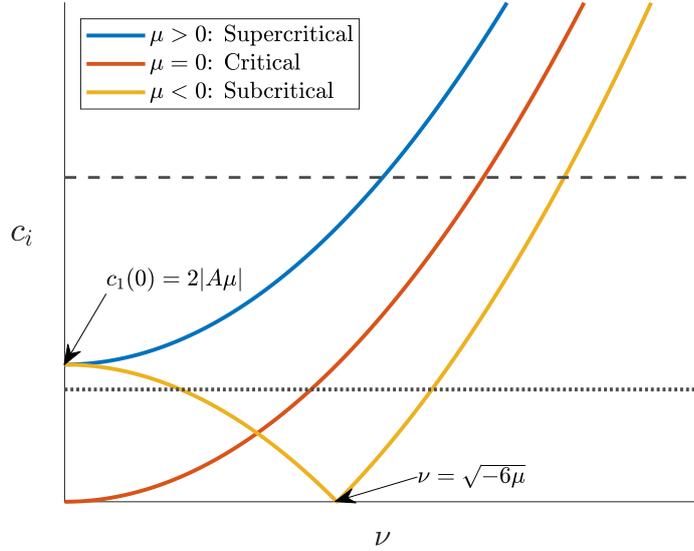


Figure 6.2.1: Schematic of the c_i curves against frequency ν for different regimes of flow. The solid lines represent the possible forms of the c_1 curves, depending on the Froude number, and the broken lines represent c_2 for different arbitrary values of A , plotted only to show how the value of A can alter the number of possible intersections between the c_1 and c_2 curves and the values of ν for which one mode will be dominant over the other.

- If $|\mu| < \frac{3|A|}{8}$ then c_1 and c_2 interact once at $\nu = \sqrt{\frac{9|A|}{4} - 6\mu}$ with c_1 being the most energetic mode for $\nu > \sqrt{\frac{9|A|}{4} - 6\mu}$ and c_2 the most energetic mode for $\nu < \sqrt{\frac{9|A|}{4} - 6\mu}$
- If $|\mu| \geq \frac{3|A|}{8}$ an additional intersection occurs at $\nu = \sqrt{-\left(\frac{9|A|}{4} + 6\mu\right)}$ and the most energetic mode is c_1 for $\nu < \sqrt{-\left(\frac{9|A|}{4} + 6\mu\right)}$ and $\nu > \sqrt{\frac{9|A|}{4} - 6\mu}$ whereas c_2 is the most energetic mode for $\sqrt{-\left(\frac{9|A|}{4} + 6\mu\right)} < \nu < \sqrt{\frac{9|A|}{4} - 6\mu}$.

In order to compare these results to those of the fully nonlinear problem a free surface was prescribed in the form

$$y_f = 1 + \frac{A}{2} \left(\tanh(b(c - \phi)) + \tanh(b(c + \phi)) \right) \cos(\nu\phi) \quad (6.2.7)$$

where b is constant controlling how quickly these waves are forced to decay at the edge of their interval of support and c controls how large that interval is.

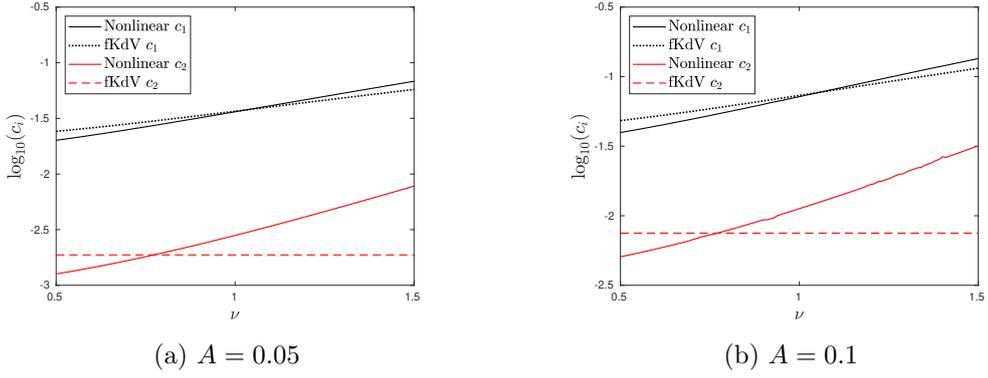


Figure 6.2.2: The energies carried in the Fourier modes of the inverse solution for the fully nonlinear (solid) and fKdV (dashed) models with $F = 1.2$, $N = 1624$. No intersection occurs in supercritical flow when $\mu > \frac{3|A|}{8}$. (a) $A = 0.05$ (b) $A = 0.1$.

Due to the assumption of uniform flow as $\phi \rightarrow \infty$ in the fully nonlinear model it was not possible to take an infinite train like it was for the fKdV. To still be able to compare the results c was taken large enough that multiple wave cycles would be seen in the interval of support and then the inverse solution truncated to $[l, l]$ with $l < c$ in order to remove edge effects from the tanh functions used. The resultant truncated wave was then Fourier transformed in order to find the energy carried in each mode and to compare to the predictions of the fKdV.

The fKdV predicted two possible scenarios for supercritical flow, the first of which was that there would be no intersections between c_1 and c_2 with the first mode being dominant everywhere. This behaviour is reflected in the fully nonlinear calculations with Figure 6.2.2 displaying this for two values of the amplitude that were chosen to satisfy $\mu > \frac{3|A|}{8}$.

The other possibility for supercritical flow was that if $\mu \leq \frac{3|A|}{8}$ an intersection occurs, this scenario is plotted in Figure 6.2.3. We can see that an intersection does indeed occur, it was predicted by the fKdV, for the parameters chosen, to happen at $\nu = \sqrt{\frac{3}{20}} \approx 0.3873$ with energy $c_1 = c_2 = 0.03$ and found in the fully nonlinear calculations to occur with $\nu \approx 0.76$ and $c_1 = c_2 \approx 0.0466$.

For critical flow the fKdV predicts that the c_1 and c_2 modes will share the same energy for $\nu = \sqrt{\frac{9|A|}{4}}$ with c_1 most energetic for greater frequencies and c_2 most

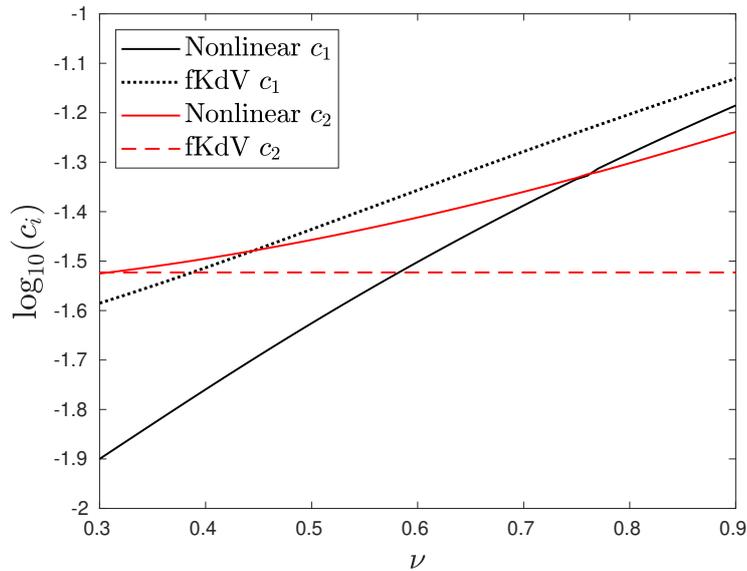


Figure 6.2.3: The energies carried in the Fourier modes of the inverse solution for the fully nonlinear (solid) and fKdV (dashed) models with $A = 0.2$, $N = 1624$, $F = 1.05$. An intersection occurs in supercritical flow when $\mu \leq \frac{3|A|}{8}$.

energetic for lower frequencies. In Figure 6.2.4 the predicted fKdV energies are compared to those found by the fully-nonlinear method for $A = 0.1$. The intersection point was predicted by the fKdV to occur at $\nu = \sqrt{0.225} \approx 0.4743$ with $c_1 = c_2 = 0.0075$, whereas the intersection calculated for the fully nonlinear problem occurred for $\nu \approx 0.615$ with $c_1 = c_2 \approx 0.011$.

Remaining is the case of subcritical flow with $\mu < 0$. We have been unable to find a set of parameters satisfying $|\mu| < \frac{3|A|}{8}$, such that the prediction of the fKdV is that there is only one intersection point, that yield an inverse solution that when then used as input into the forward problem would converge and return the originally prescribed surface. This leaves only the case $\mu \geq \frac{3|A|}{8}$ to be able to test against the predictions of the fKdV (Figure 6.2.5). The fKdV predicts the left-hand intersection to happen at $\nu \approx 0.7450$ with energy $c_1 = c_2 \approx 0.0003$ where it is found in the nonlinear calculations to happen for $\nu \approx 0.7925$ with an energy of $c_1 = c_2 \approx 0.00077$. The fKdV predicts the right-hand intersection to occur for $\nu \approx 0.8032$ with the same energy $c_1 = c_2 \approx 0.0003$, the fully nonlinear calculations find this to occur for $\nu \approx 0.935$ with energy $c_1 \approx 0.00092$. Most

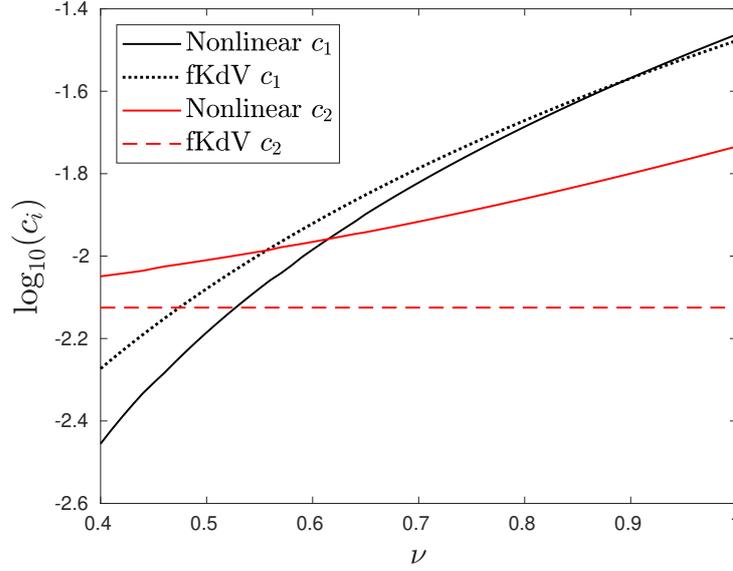


Figure 6.2.4: Comparison of the energy in the Fourier modes between nonlinear (solid) and fKdV (dashed) for critical flow with $A = 0.1, N = 1624$. One intersection occurs between the two energies.

interesting about this case is that because $\mu < 0$ the fKdV predicts that the energy of the first mode can go exactly to zero for $\mu = -\frac{\nu^2}{6}$. It can be seen from Figure 6.2.5) that the energy in the first mode does indeed sharply decrease for some value of ν .

To understand where this energy goes to zero the fully nonlinear problem consider the dimensional surface $\hat{\eta} = A \cos(k\hat{x})$ where A has units of length and k of inverse length. We nondimensionalise this surface with $\hat{x} = Hx$ and $\hat{\eta} = H\eta$ where H is the typical depth of the dimensional flow. This leads to the nondimensional form $\eta = \epsilon \cos(\nu x)$ where $k = \frac{\nu}{H}$ and $\epsilon = \frac{A}{H}$. The unforced steady fKdV was given by

$$\mu\eta_x - \frac{3}{2}\eta\eta_x - \frac{1}{6}\eta_{xxx} = 0. \quad (6.2.8)$$

Integrate the above expression once to obtain

$$\mu\eta - \frac{3}{4}\eta^2 - \frac{1}{6}\eta_{xx} = 0 \quad (6.2.9)$$

before substituting in for the small amplitude surface $\eta = \epsilon \cos(\nu x)$ with $\epsilon \ll 1$

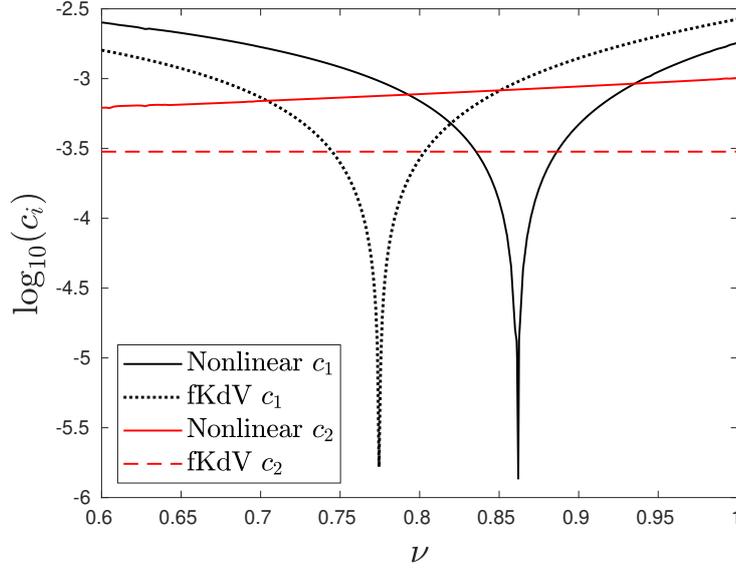


Figure 6.2.5: Comparison of the energy in the Fourier modes between nonlinear (solid) and fKdV (dashed) for subcritical flow with $A = 0.02$, $N = 1624$, $F = 0.9$. Two intersections occur in subcritical flow and the energy in the c_1 mode appears to go to zero for a particular value of ν .

causing the η^2 term to vanish leaving

$$\cos(\nu x) \left(\frac{\nu^2}{6} + \mu \right) = 0 \quad (6.2.10)$$

or

$$\nu = \sqrt{-6\mu}. \quad (6.2.11)$$

The dispersion relation for linear gravity waves in the absence of surface tension (see, for example, Billingham and A. C. King, 2001 or Vanden-Broeck, 2010) is

$$c^2 = \frac{g}{k} \tanh(kH) = \frac{gH}{\nu} \tanh(\nu). \quad (6.2.12)$$

With the Froude number given by $F = \frac{U}{\sqrt{gH}}$ we can square the F and use the typical fluid speed $U = c$ to obtain

$$F^2 = \frac{U^2}{gH} = \frac{c^2}{gH} = \frac{gH}{\nu gH} \tanh \nu. \quad (6.2.13)$$

Now we find that the solution to the transcendental equation

$$\nu F^2 = \tanh(\nu) \tag{6.2.14}$$

predicts the root for nonlinear curve in Figure 6.2.5 as $\nu \approx 0.8589$.

Conclusion

In Chapter 2 we explored how to perform truncated singular value decomposition and the use of the discrete Picard condition and L-curve in selecting an appropriate truncation rank in order to try get good approximate solutions to ill-posed problems based on the discretisation of Fredholm integral equations of the first kind.

In Chapter 3 we showed how prescribing $y_f(\phi)$ lets the inverse problem for the topography be expressed as a linear problem in $\theta_b(\phi)$ albeit an ill-posed problem due to its nature as a Fredholm integral equation of the first kind. We introduced a discretisation to the problem and showed how we can apply our knowledge of truncated singular value decomposition and the Moore-Penrose psuedoinverse to obtain least squares solutions to a regularised problem that, so long as the discrete Picard condition is satisfied, should be a good approximation to the true underlying problem allowing for the grid scale issues that were present in applying Newton's method to the problem, experienced by Binder, Blyth, and McCue, 2013; Tam, Yu, Kelso, and Binder, 2015; and Vasan and Deconinck, 2013.

Chapter 4 saw us test the truncated singular value decomposition method by attempting to retrieve a known topography featuring one or two Gaussians from computed forward solutions. We found that when the free surface decayed in the far field we were able to achieve good results for the topography by our method for multiple types of free surface solutions, namely solutions that are

perturbations to the uniform stream and those that are perturbations to the solitary wave solutions, the unforced solitary wave, hydraulic falls and trapped waves in subcritical flow.

In Chapter 5 the ability of the truncated singular value decomposition to give good regularised solutions when the input data is perturbed with white noise was trialled for differing amounts of noise. It was found that while the problem is sensitive to noise the regularisation of solutions achieved by the method meant solutions were still relatively accurate for reasonable amount of noise. When very noisy input data was given we found that repeating the problem for a different sets of perturbed data, akin to taking more set of experimental measurements, and averaging the output still allowed for the primary characteristics of the Gaussian topography to be recovered, albeit with some small undulations persisting about the zero line. In the case of the hydraulic fall with a large amount of added noise we found that the Gaussian was captured well in the inverse solution but that the level of the undisturbed topography was not predicted correctly as $\phi \rightarrow -\infty$.

In Chapter 6 we saw how being able to prescribe a surface directly and then quickly obtain inverse solutions enables exploration of the inverse problem's solution space that would have been difficult before as we can now directly alter properties of the surface. To this purpose we considered how changing individually the amplitude of a Gaussian free surface or the Froude number affects the output for the topography. We then studied trapped wave solutions via Fourier transforms of the inverse solutions finding that the fKdV can help to predict which mode will carry the most energy, while it did not correctly predict the frequencies at which one mode becomes dominant it does correctly predict that the dominant mode will change as the frequency is changed. We also found that the energy in the first mode in fully nonlinear subcritical flow has a frequency at which the energy contained goes to zero, making the second mode the only mode present, and that the frequency at which this occurs is predicted by the dispersion equation for linear waves.

Future work would include further exploration of the inverse solution space now that the truncated singular value decomposition method has been shown provide fast and accurate results, although it is important that the outputs are confirmed to be physical and converge again in the forward sense as the inverse method is quite indifferent to if the input data it is fed corresponds to a physical surface. It would be interesting to explore the combined problem where a non-zero pressure is imposed on the flow, certainly to explore how the developed method could be used to construct topographies such that the effects of the non-zero pressure and the topography cancel one another to produce a flat free surface despite the presence of forcing.

A

Norm of inverse fKdV forcing for a Gaussian free surface

The inversely found forcing from the fKdV was given by the equation

$$f(x) = 2\mu\eta - \frac{1}{3}\eta_{xx} - \frac{3}{2}\eta^2. \quad (\text{A.0.1})$$

For a Gaussian free surface given by

$$\eta = \alpha e^{-(b\phi)^2} \quad (\text{A.0.2})$$

the forcing is then given explicitly by

$$f(x) = \alpha e^{-(bx)^2} \left(2(F - 1) + \frac{2b^2}{3}(1 - 2b^2x^2) - \frac{3a}{2}e^{-(bx)^2} \right). \quad (\text{A.0.3})$$

Seeking $\|f\|$, the norm of the forcing, we use the definition

$$\|f\|^2 = \int_{-\infty}^{\infty} f^2 dx. \quad (\text{A.0.4})$$

and first find

$$f^2 = \left[\alpha e^{-(bx)^2} \left(2\mu + \frac{2b^2}{3}(1 - 2b^2x^2) - \frac{3a}{2}e^{-(bx)^2} \right) \right]^2,$$

expanding before collecting powers of x to give

$$f^2 = \left[\left(4\mu^2 + \frac{4b^4}{9} + \frac{8\mu b^2}{3} \right) - \left(\frac{16b^6}{9} + \frac{16\mu b^4}{3} \right) x^2 + \frac{16b^8}{9} x^4 \right. \\ \left. - (6\mu a + 2ab^2) e^{-(bx)^2} + 4ab^4 x^2 e^{-(bx)^2} + \frac{9a^2}{4} e^{-2(bx)^2} \right] a^2 e^{-2(bx)^2}.$$

Substituting f^2 into (A.0.4) we first note that $a = 0 \implies \|f\| = 0$ and so exclude this case as we use the linearity of the integral to divide by a^2 , expand the integrals and factor out constants to yield

$$\frac{\|f\|^2}{a^2} = \left(4\mu^2 + \frac{4b^4}{9} + \frac{8\mu b^2}{3} \right) \int_{-\infty}^{\infty} e^{-2(bx)^2} dx - (6\mu a + 2ab^2) \int_{-\infty}^{\infty} e^{-3(bx)^2} dx \\ + \frac{9a^2}{4} \int_{-\infty}^{\infty} e^{-4(bx)^2} dx - \left(\frac{16b^6}{9} + \frac{16\mu b^4}{3} \right) \int_{-\infty}^{\infty} x^2 e^{-2(bx)^2} dx \\ + 4ab^4 \int_{-\infty}^{\infty} x^2 e^{-3(bx)^2} dx + \frac{16b^8}{9} \int_{-\infty}^{\infty} x^4 e^{-2(bx)^2} dx. \quad (\text{A.0.5})$$

Now, in order to evaluate the integrals in (A.0.5) we will need to consider integrals of the form

$$I_m = \int_{-\infty}^{\infty} x^{2m} e^{-n(bx)^2} dx \quad (\text{A.0.6})$$

the solution for which shall be proved by induction. We propose that

$$I_m = \left(\frac{1}{2nb^2} \right)^m \frac{(2m)!}{2^m m!} \sqrt{\frac{\pi}{nb^2}}. \quad (\text{A.0.7})$$

for $m \in \mathbb{N}$ where $b \in \mathbb{R}$ and $n > 0$. For the base case we make use of the result that

$$I_0 = \int_{-\infty}^{\infty} e^{-n(bx)^2} dx = \sqrt{\frac{\pi}{nb^2}}$$

and check that

$$I_0 = \left(\frac{1}{2nb^2} \right)^0 \frac{(0)!}{2^0 0!} \sqrt{\frac{\pi}{nb^2}} = \sqrt{\frac{\pi}{nb^2}}. \quad (\text{A.0.8})$$

Assume now that (A.0.7) is true for $m = k$ and consider integrating (A.0.6) by parts with $u = x^{2k-1}$ and $dv = x e^{-n(bx)^2}$, and so $du = (2k-1)x^{2k-2}$ and

$v = -\frac{1}{2nb^2}e^{-n(bx)^2}$, giving

$$\begin{aligned} I_k &= \left(x^{2k-1} \frac{-1}{2nb^2} e^{-n(bx)^2} \right) \Big|_{-\infty}^{\infty} + \frac{(2k-1)}{2nb^2} \int_{-\infty}^{\infty} x^{2k-2} e^{-n(bx)^2} dx \\ &= 0 + \frac{(2k-1)}{2nb^2} \int_{-\infty}^{\infty} x^{2(k-1)} e^{-n(bx)^2} dx = \frac{(2k-1)}{2nb^2} I_{k-1}, \end{aligned}$$

or alternatively

$$I_{k+1} = \frac{(2k+1)}{2nb^2} I_k. \quad (\text{A.0.9})$$

Then by the assumption on I_k we have

$$I_{k+1} = \frac{(2k+1)}{2nb^2} \left(\frac{1}{2nb^2} \right)^k \frac{(2k)!}{2^k k!} \sqrt{\frac{\pi}{nb^2}}$$

and by noting that

$$(2k+1) \frac{(2k)!}{2^k k!} = (2k+1) \frac{(2k)!}{2^k k!} \frac{2k+2}{2(k+1)} = \frac{(2k+2)!}{2^{k+1} (k+1)!}$$

we find

$$I_{k+1} = \left(\frac{1}{2nb^2} \right)^{k+1} \frac{(2(k+1))!}{2^{k+1} (k+1)!} \sqrt{\frac{\pi}{nb^2}}$$

and so given that the assumption holds for $m = k$ it also holds $m = k + 1$. As (A.0.7) is for $m = 0$ it must then be true for $m \in \mathbb{N}$.

Returning to (A.0.5) we can now evaluate

$$\begin{aligned} \frac{\|f\|^2}{a^2} &= \left(4\mu^2 + \frac{4b^4}{9} + \frac{8\mu b^2}{3} \right) \sqrt{\frac{\pi}{2b^2}} - (6\mu a + 2ab^2) \sqrt{\frac{\pi}{3b^2}} \\ &\quad + \frac{9a^2}{4} \sqrt{\frac{\pi}{4b^2}} - \left(\frac{16b^6}{9} + \frac{16\mu b^4}{3} \right) \sqrt{\frac{\pi}{2}} \frac{1}{4b^3} \\ &\quad + 4ab^4 \sqrt{\frac{\pi}{3}} \frac{1}{6b^3} + \frac{16b^8}{9} \sqrt{\frac{\pi}{2}} \frac{3}{16b^5}. \end{aligned}$$

Collecting terms and rearranging we obtain an expression for the norm of the

forcing as

$$\|f\| = |a\pi^{\frac{1}{4}}| \sqrt{\frac{4\mu^2}{b\sqrt{2}} + \frac{4b\mu + b^3}{3\sqrt{2}} - \frac{6a\mu}{b\sqrt{3}} - \frac{4ab}{3\sqrt{3}} + \frac{9a^2}{8b}}. \quad (\text{A.0.10})$$

It is not obvious that the quantity contained within the square root above should always be positive as the parameters a , b and μ are varied. A contradiction here may indicate a surface that is non-physical. Let

$$h(a, b, \mu) = \frac{4\mu^2}{b\sqrt{2}} + \frac{4b\mu + b^3}{3\sqrt{2}} - \frac{6a\mu}{b\sqrt{3}} - \frac{4ab}{3\sqrt{3}} + \frac{9a^2}{8b} \quad (\text{A.0.11})$$

We would like to show that $h \geq 0$ given $b > 0$, so we shall suppose instead that somewhere $h < 0$. Collecting terms in h to highlight how it is quadratic in a as

$$h = \left(\frac{9}{8b}\right) a^2 - \left(\frac{6\mu}{b\sqrt{3}} + \frac{4b}{3\sqrt{3}}\right) a + \left(\frac{4b\mu + b^3}{3\sqrt{2}} + \frac{4\mu^2}{b\sqrt{2}}\right) \quad (\text{A.0.12})$$

we see that as the coefficient of a^2 is positive then there must be real roots to the equation obtained by setting (A.0.12) equal to zero as we have assumed that somewhere $h < 0$. We call h_1 the discriminant of applying the quadratic formula to $h = 0$ finding

$$h_1 = \left(\frac{6\mu}{b\sqrt{3}} + \frac{4b}{3\sqrt{3}}\right)^2 - 4\left(\frac{9}{8b}\right) \left(\frac{4b\mu + b^3}{3\sqrt{2}} + \frac{4\mu^2}{b\sqrt{2}}\right).$$

For real roots to exist we must have that $h_1 \geq 0$. We expand h_1 and collect powers of b to write

$$h_1 = \frac{1}{36b^2} \left((64 - 81\sqrt{2})b^4 + (576 - 324\sqrt{2})\mu b^2 + (1296 - 972\sqrt{2})\mu^2 \right) \quad (\text{A.0.13})$$

which is a positive number multiplying a quadratic in b^2 . Writing $B = b^2$ and introducing h_2 for the bracketed terms then for $h_1 \geq 0$ we must have

$$h_2 = (64 - 81\sqrt{2})B^2 + (576 - 324\sqrt{2})\mu B + (1296 - 972\sqrt{2})\mu^2 \geq 0. \quad (\text{A.0.14})$$

We see that for $h_2 \geq 0$ to be true somewhere then the equation $h_2 = 0$ must have real roots as the coefficient of B^2 is negative. However, writing h_3 as the discriminant of applying the quadratic formula to $h_2 = 0$ we find

$$h_3 = ((576 - 324\sqrt{2})\mu)^2 - 4(64 - 81\sqrt{2})(1296 - 972\sqrt{2})\mu^2$$

which rearranges to

$$h_3 = (295488 - 419904\sqrt{2})\mu^2 \approx -2021\mu^2 < 0.$$

As h_3 is negative there are no real roots to the equation $h_2 = 0$. This means that $h_2 < 0$ and therefore $h_1 < 0$ which in turn implies that there are no real roots to the equation $h = 0$ and therefore we cannot have values of $h < 0$. This contradicts our initial assumption and as such we may conclude that $h \geq 0$.

The fKdV predicts that a minimum will occur in the norm for a value $F = F_{min}$.

To see this differentiate (A.0.11) with respect to μ to obtain

$$\frac{dh}{d\mu} = \frac{8\mu}{b\sqrt{2}} + \frac{4b}{3\sqrt{2}} - \frac{6a}{b\sqrt{3}}. \quad (\text{A.0.15})$$

Setting the above equal to zero and substituting for $\mu = F - 1$ a stationary point is found for the value

$$F_{min} = 1 + \frac{a\sqrt{6}}{4} - \frac{b^2}{6}. \quad (\text{A.0.16})$$

This is easily shown to be a minimum as

$$\frac{d^2h}{d^2\mu} = \frac{8}{b\sqrt{2}} > 0. \quad (\text{A.0.17})$$

B

Repeated integration by parts of a Gaussian

Consider integrating by parts the expression

$$G(\xi) = \int e^{-\xi^2} d\xi = \int \frac{-2\xi}{-2\xi} e^{-\xi^2} d\xi \quad (\text{B.0.1})$$

N times giving an expression of the form

$$G(\xi) = \left(\sum_{m=0}^N E_m(\xi) \right) + I_{N+1}(\xi) \quad (\text{B.0.2})$$

where E_m , which represent the boundary terms, and I_m , the resultant integral term, are functions to be found. Observe that for $N = 0$ from (B.0.1) we have $E_0 = 0$ and

$$I_1 = \int e^{-\xi^2} d\xi.$$

By definition the different I_m (for $m \geq 1$) are related by the recursive formula

$$I_m = E_m + I_{m+1}. \quad (\text{B.0.3})$$

We propose that

$$I_r = \left(\prod_{m=1}^r 2m - 3 \right) \frac{(-1)^r}{2^{r-1}} \int \xi^{-(2r-2)} e^{-\xi^2} d\xi \quad (\text{B.0.4})$$

and prove it by induction. We shall make use of the result (obtained by integrating by parts once) that

$$\begin{aligned} \int \xi^{-(2k-2)} e^{-\xi^2} d\xi &= \int \xi^{-(2k-2)} \frac{-2\xi}{-2\xi} e^{-\xi^2} d\xi \\ &= -\frac{\xi^{-(2k-1)}}{2} e^{-\xi^2} - \int \frac{(2k-1)\xi^{-2k}}{2} e^{-\xi^2} d\xi. \end{aligned} \quad (\text{B.0.5})$$

For $r = 1$, i.e. after integrating by parts zero times, this formula yields:

$$I_1 = \left(\prod_{m=1}^1 2m - 3 \right) \frac{(-1)^1}{2^{1-1}} \int x^{-(2-2)} e^{-x^2} dx = \int e^{-x^2} dx \quad (\text{B.0.6})$$

and so we see the base case is verified. Let us assume now that (B.0.4) holds for $r = k$ such that

$$I_k = \left(\prod_{m=1}^k 2m - 3 \right) \frac{(-1)^k}{2^{k-1}} \int x^{-(2k-2)} e^{-x^2} dx. \quad (\text{B.0.7})$$

Now, applying the result of (B.0.5)

$$I_k = \left(\prod_{m=1}^k 2m - 3 \right) \frac{(-1)^k}{2^{k-1}} \left[-\frac{x^{-(2k-1)}}{2} e^{-x^2} - \int \frac{(2k-1)x^{-2k}}{2} e^{-x^2} dx \right]. \quad (\text{B.0.8})$$

Expanding, using $(2k-1) = (2(k+1)-3)$ to rewrite the product in the second term and then simplifying we find

$$I_k = \left(\prod_{m=1}^k 2m - 3 \right) \frac{(-1)^{k+1}}{2^k} x^{-(2k-1)} e^{-x^2} + \left(\prod_{m=1}^{k+1} 2m - 3 \right) \frac{(-1)^{k+1}}{2^k} \int x^{-2k} e^{-x^2} dx \quad (\text{B.0.9})$$

By comparison with $I_k = E_k + I_{k+1}$ we then find

$$E_k = \left(\prod_{m=1}^k 2m - 3 \right) \frac{(-1)^{k+1}}{2^k} x^{-(2k-1)} e^{-x^2} \quad (\text{B.0.10})$$

$$I_{k+1} = \left(\prod_{m=1}^{k+1} 2m - 3 \right) \frac{(-1)^{k+1}}{2^k} \int x^{-2k} e^{-x^2} dx. \quad (\text{B.0.11})$$

Now, (B.0.11) shows that if (B.0.4) is true for $r = k$ then it is also true for

$r = k + 1$. As we have shown it to be true for $r = 1$ we can conclude by induction that it is true for $r \in \mathbb{N}$. Importantly, in the process of proving this result, we have derived (B.0.10) which gives an exact expression for each term E_k .

For $k \geq 1$ we find that

$$\frac{E_{k+1}}{E_k} = \frac{(1 - 2k)}{2\xi^2} \quad (\text{B.0.12})$$

and so as $\xi \rightarrow \infty$ we have $E_{k+1} \ll E_k$ and therefore

$$G(\xi) = \int e^{-\xi^2} d\xi \sim E_1(\xi) = \frac{-1}{2\xi} e^{-\xi^2}. \quad (\text{B.0.13})$$

The behaviour of $\int G(\xi) d\xi$ for large ξ can be found by similar argument to be

$$\int G(\xi) d\xi \sim \int \frac{-1}{2\xi} e^{-\xi^2} d\xi \sim \frac{1}{4\xi^2} e^{-\xi^2}. \quad (\text{B.0.14})$$

Bibliography

- Akylas, T. R. (1984). “On the excitation of long nonlinear water waves by a moving pressure distribution”. In: *Journal of Fluid Mechanics* 141, pp. 455–466.
- Ben-Israel, A. and Greville, T. N. E. (2003). *Generalized inverses: theory and applications*. Vol. 15. Springer Science & Business Media.
- Billingham, J. and King, A. C. (2001). *Linear Water Waves*. Cambridge Texts in Applied Mathematics. Cambridge University Press.
- Binder, B. J., Blyth, M. G., and Balasuriya, S. (2015). “Steady free-surface flow over spatially periodic topography”. In: *Journal of Fluid Mechanics* 781, R3.
- Binder, B. J., Dias, F., and Vanden-Broeck, J.-M. (2008). “Influence of rapid changes in a channel bottom on free-surface flows”. In: *IMA Journal of Applied Mathematics* 73.1, pp. 254–273.
- Binder, B. J. and Vanden-Broeck, J.-M. (2007). “The effect of disturbances on the flows under a sluice gate and past an inclined plate”. In: *Journal of Fluid Mechanics* 576, pp. 475–490.
- Binder, B. J., Vanden-Broeck, J.-M., and Dias, F. (2005). “Forced solitary waves and fronts past submerged obstacles”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 15.3.
- Binder, B. J. (2019). “Steady Two-Dimensional Free-Surface Flow Past Disturbances in an Open Channel: Solutions of the Korteweg-De Vries Equation and Analysis of the Weakly Nonlinear Phase Space”. In: *Fluids* 4.1.

- Binder, B. J., Blyth, M. G., and McCue, S. W. (2013). “Free-surface flow past arbitrary topography and an inverse approach for wave-free solutions”. In: *IMA Journal of Applied Mathematics* 78.4, pp. 685–696.
- Binder, B. J. and Vanden-Broeck, J.-M. (2011). “Hybrid free-surface flows in a two-dimensional channel”. In: *Physical Review E* 84.1.
- Capelli, A. (1892). “Sopra la compatibilita o incompatibilita di piu equazioni di primo grado fra piu incognite”. In: *Revista di Matematica* 2, pp. 54–58.
- Cole, S. (1985). “Transient waves produced by flow past a bump”. In: *Wave Motion* 7.6, pp. 579–587.
- Croeze, A., Pittman, L., and Reynolds, W. (2012). “Solving Nonlinear Least-Squares Problems with the Gauss-Newton and Levenberg-Marquardt Methods”. In: URL: <https://www.math.lsu.edu/system/files/MunozGroup1%20-%20Paper.pdf>.
- Dias, F. and Vanden-Broeck, J.-M. (2004). “Trapped waves between submerged obstacles”. In: *Journal of Fluid Mechanics* 509, pp. 93–102.
- Dias, F. and Vanden-Broeck, J.-M. (1989). “Open channel flows with submerged obstructions”. In: *Journal of Fluid Mechanics* 206, pp. 155–170.
- (2002). “Generalised critical free-surface flows”. In: *Journal of Engineering Mathematics* 42.3-4, pp. 291–301.
- Eckart, C. and Young, G. (1936). “The approximation of one matrix by another of lower rank”. In: *Psychometrika* 1.3, pp. 211–218.
- Forbes, L. K. and Schwartz, L. W. (1982). “Free-surface flow over a semicircular obstruction”. In: *Journal of Fluid Mechanics* 114, pp. 299–314.
- Griffel, D. H. (1989). *Linear algebra and its applications*. Ellis Horwood.
- (2002). *Applied functional analysis*. Courier Corporation.
- Grimshaw, R. H. J. and Smyth, N. (1986). “Resonant flow of a stratified fluid over topography”. In: *Journal of Fluid Mechanics* 169, pp. 429–464.
- Groetsch, C. W. (2007). “Integral equations of the first kind, inverse problems and regularization: a crash course”. In: *Journal of Physics: Conference Series*. Vol. 73. IOP Publishing.

- Hansen, P. C. (1987). “The truncated SVD as a method for regularization”. In: *BIT Numerical Mathematics* 27, pp. 534–553.
- (1990a). “The discrete Picard condition for discrete ill-posed problems”. In: *BIT Numerical Mathematics* 30.4, pp. 658–672.
- (1990b). “Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank”. In: *SIAM Journal on Scientific and Statistical Computing* 11.3, pp. 503–518.
- (1992a). “Analysis of discrete ill-posed problems by means of the L-curve”. In: *SIAM review* 34.4, pp. 561–580.
- (1992b). “Numerical tools for analysis and solution of Fredholm integral equations of the first kind”. In: *Inverse Problems* 8.6, p. 849.
- Hansen, P. C. and O’Leary, D. P. (1993). “The use of the L-curve in the regularization of discrete ill-posed problems”. In: *SIAM Journal on Scientific Computing* 14.6, pp. 1487–1503.
- Holmes, R. J., Hocking, G. C., Forbes, L. K., and Baillard, N. Y. (2013). “Waveless subcritical flow past symmetric bottom topography”. In: *European Journal of Applied Mathematics* 24.2, pp. 213–230.
- Hunter, J. and Vanden-Broeck, J.-M. (1983). “Accurate computations for steep solitary waves”. In: *Journal of fluid Mechanics* 136, pp. 63–71.
- James, M. (1978). “The generalised inverse”. In: *The Mathematical Gazette* 62.420, pp. 109–114.
- Kabanikhin, S. I. (2008). “Definitions and examples of inverse and ill-posed problems”. In: *Journal of Inverse and Ill-posed Problems* 16, pp. 317–357.
- Keeler, J. S., Binder, B. J., and Blyth, M. G. (2017). “On the critical free-surface flow over localised topography”. In: *Journal of Fluid Mechanics* 832, pp. 73–96.
- Keeler, J. S., Binder, B. J., and Blyth, M. G. (2018). “Steady two-dimensional free-surface flow over semi-infinite and finite-length corrugations in an open channel”. In: *Physical Review Fluids* 3.11.
- Keeler, J. S., Blyth, M. G., and King, J. R. (2021). “Termination points and homoclinic glueing for a class of inhomogeneous nonlinear ordinary differential equations”. In: *Nonlinearity* 34.1, p. 532.

- Keller, J. B. (1976). “Inverse Problems”. In: *The American Mathematical Monthly* 83.2, pp. 107–118.
- Lawson, C. L. and Hanson, R. J. (1974). *Solving least squares problems*. Prentice Hall.
- Lustri, C. J., McCue, S. W., and Binder, B. J. (2012). “Free surface flow past topography: a beyond-all-orders approach”. In: *European Journal of Applied Mathematics* 23.4, pp. 441–467.
- Miller, K. (1970). “Least squares methods for ill-posed problems with a prescribed bound”. In: *SIAM Journal on Mathematical Analysis* 1.1, pp. 52–74.
- Penrose, R. (1955). “A generalized inverse for matrices”. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 51. 3. Cambridge University Press, pp. 406–413.
- Phillips, D. L. (1962). “A technique for the numerical solution of certain integral equations of the first kind”. In: *Journal of the ACM (JACM)* 9.1, pp. 84–97.
- Planitz, M. (1979). “Inconsistent systems of linear equations”. In: *The Mathematical Gazette* 63.425, pp. 181–185.
- Robbins, C., Blyth, M. G., Maclean, J., and Binder, B. J. (2023). “A method to calculate inverse solutions for steady open channel free surface flow”. In: *Submitted*.
- Schot, S. H. (1992). “Eighty years of Sommerfeld’s radiation condition”. In: *Historia Mathematica* 19.4, pp. 385–401.
- Shen, S. S. P. (1995). “On the accuracy of the stationary forced Korteweg-de Vries equation as a model equation for flows over a bump”. In: *Quarterly of Applied Mathematics* 53.4, pp. 701–719.
- Stokes, G. G. (1880). “Supplement to a paper on the theory of oscillatory waves”. In: *Mathematical and Physical Papers vol. 1*, pp. 314–326.
- Tam, A., Yu, Z., Kelso, R. M., and Binder, B. J. (2015). “Predicting channel bed topography in hydraulic falls”. In: *Physics of Fluids* 27.11.
- Vanden-Broeck, J.-M. (1987). “Free-surface flow over an obstruction in a channel”. In: *Physics of Fluids* 30.8, pp. 2315–2317.

-
- (1997). “Numerical calculations of the free-surface flow under a sluice gate”. In: *Journal of Fluid Mechanics* 330, pp. 339–347.
 - (2010). *Gravity-capillary free-surface flows*. Cambridge University Press.
 - Varah, J. M. (1973). “On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems”. In: *SIAM Journal on Numerical Analysis* 10.2, pp. 257–267.
 - (1979). “A practical examination of some numerical methods for linear discrete ill-posed problems”. In: *Siam Review* 21.1, pp. 100–111.
 - (1983). “Pitfalls in the numerical solution of linear ill-posed problems”. In: *SIAM Journal on Scientific and Statistical Computing* 4.2, pp. 164–176.
 - Vasan, V. and Deconinck, B. (2013). “The inverse water wave problem of bathymetry detection”. In: *Journal of Fluid Mechanics* 714, pp. 562–590.
 - Wade, S. L., Binder, B. J., Mattner, T. W., and Denier, J. P. (2017). “Steep waves in free-surface flow past narrow topography”. In: *Physics of Fluids* 29.6.
 - Whitham, G. B. (2011). *Linear and nonlinear waves*. John Wiley & Sons.
 - Wu, T. Y.-T. (1987). “Generation of upstream advancing solitons by moving disturbances”. In: *Journal of Fluid Mechanics* 184, pp. 75–99.