Numerical Detection of Wave Breaking in the Short-Pulse Equation

Jeffrey Slepoi

Follow this and additional works at: https://digitalcommons.montclair.edu/etd

Part of the Mathematics Commons
Numerical Detection of Wave Breaking in the Short-Pulse Equation

by

Jeffrey Slepoi

A Master’s Thesis Submitted to the Faculty of
Montclair State University
In Partial Fulfillment of the Requirements
For the Degree of
Master of Pure and Applied Mathematics

May, 2016

College of Science and Mathematics
Department of Mathematics

Thesis Committee:

Dr. A. David Trubatch, Thesis Advisor

Dr. Arup Mukherjee, Committee Member

Dr. Bogdan G. Nita, Committee Member
Abstract

Numerical analysis is a powerful resource in all mathematical sciences especially in the study of partial differential equations (PDEs). It allows evaluating and demonstrating derived solutions for PDEs and whenever the solution can’t be derived analytically it provides us with ability to calculate the solution function numerically and predict its behavior over time.

This work presents a numerical method to evaluate an analytically known solution, demonstrates the needed parameters to achieve the desired accuracy, extends the methodology into the sphere of the mathematical unknown to be able to predict the results by using the same numerical methodology.

The equation in question which we are going to analyze is a short pulse equation (SPE) which is an alternative model for the nonlinear Schrödinger equation. The SPE finds applications, for example, as a model for ultrashort pulses in optical fibers and has a form: $u_{xt} = u + \frac{1}{\delta}(u^3)_{xx}$.

SPE is an integrable nonlinear partial differential equation. The soliton solutions of the NLSE have played an important role in the development of fiber-optic communications. But when the pulse becomes short, results produced by the NLSE worsen but the SPE generates good output. For this reason, it is very interesting to find the exact or numerical way to solve the SPE which represents the ultra-short light pulses.
Numerical Detection of Wave Breaking in the Short-Pulse Equation

A THESIS

Submitted in partial fulfillment of the requirements for the degree of Masters in Mathematics

by

Jeffrey Slepoi
Montclair State University
Montclair, NJ
2016
Acknowledgments

I am very grateful to Dr. A. D. Trubatch for giving me the opportunity to write this thesis under his direction. I am indebted to his help in providing a topic for my thesis, identifying the relevant literature for this research, and helping me in the areas where my knowledge was not sufficient.
I also want to thank Giorgi Nasrashvili who, being an undergraduate student, worked with me on the project for one summer and helped me with some technical and research issues.
Contents

Acknowledgments ii
List of Figures vii
List of Tables viii

1 Introduction 1

2 Background 3
  2.1 Determined Analytical solution 3
  2.2 Detection of wave breaking in Fourier space 4
  2.3 Choice of a numerical method 8
  2.4 Analytical determination of wave breaking in finite time 9

3 Determination of wave breaking in the Analytical Solution of the SPE 12
  3.1 Choice of the time step and the number of Fourier modes 12
  3.2 Analytical results. Numerical Determination of time breaking 13

4 Numerical Solution 16
  4.1 The numerical solution of the SPE and comparison to the analytical solution 16
    4.1.1 Accuracy of numerical integration 16
    4.1.2 Conserved quantities 17
  4.2 Results of δ calculations for m under the critical value 19
  4.3 Results for m larger than the critical value 22
  4.4 Oscillations in the numerical solution as another indication of breaking 24

5 Detection of wave breaking for the Gaussian initial condition 27
  5.1 Conserved quantities for Gaussian IC 27
  5.2 Oscillations in the numerical solution for higher values of a 28
  5.3 Expanding the areas of well-posedness and wave-breaking through numerical methods 30
    5.3.1 Numerical determination of borders based on b = 0.5 30
    5.3.2 Numerical verification of scaling of the results 33
  5.4 Further improvement of the results by using a higher number of Fourier modes 34
<table>
<thead>
<tr>
<th>Derivation</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Derivation of Critical value in the analytical solution $m_{cr}$</td>
<td>40</td>
</tr>
<tr>
<td>B</td>
<td>Derivation of $t_{br}$ for $m &gt; m_{cr}$</td>
<td>42</td>
</tr>
<tr>
<td>C</td>
<td>Derivation of conserved quantity $E_0$ for the Gaussian IC</td>
<td>44</td>
</tr>
<tr>
<td>D</td>
<td>Derivation of conserved quantity $E_{-1}$ for the Gaussian IC</td>
<td>46</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Breather solitons for $m = 0.1$, $m = 0.2$ - top and $m = 0.3$, $m = 0.4$ - bottom. All solitons are shown at time zero. At a later time, the soliton with $m = 0.4 (> m_{cr})$, bottom right, becomes multivalued. 4

2.2 Oscillatory nature of the Fourier coefficients. Natural log of the coefficients across mode number at time zero for $m = 0.1$, $m = 0.2$ - top, $m = 0.3$, $m = 0.4$ - bottom. Black (top) - fitting curve across peaks, red - fitting on the whole curve. $N = 2^{15}$ modes used. 5

2.3 Fitting curves for $m = 0.2$ (top) and $m = 0.3$ (bottom). Left $t = 0$, right $t = 1.0$. Blue curve - $\ln|v|$, green - fitting over the peaks, red - fitting over the whole curve. $N = 2^{15}$. 8

2.4 Fourier coefficient over time for $m = 0.39$ (above $m_{cr}$). Oscillatory bottom curve is at $t = 0$, black curve above is at $t = t_{br}/2 \approx 0.28$, and red curve (top) is at $t = t_{br} \approx 0.56$. Solid line - $\ln|n|$, -- fitted curve. Peaks which exist in the curve at $t = 0$, disappear at a later time. $N = 2^{15}$. 9

2.5 Analytical well-posedness and and wave breaking regions. Produced by using Theorems 2 and 3 in [11]. 11

3.1 Comparison of $\log(\delta)$ over time for $m < m_{cr}$ (top to bottom): green ($\circ$) $m = 0.2$, black ($\bigwedge$) $m = 0.3$, red ($\ast$) $m = 0.35$, magenta ($\circ$) $m = 0.37$, blue ($\bigwedge$) $m = 0.375$. Time period is 3 which is approximately one full cycle ($\pi_n$). $N = 2^{16}$. 13

3.2 Comparison of $\delta$ over time for $m > m_{cr}$ (right to left): green $m = 0.385$, black $m = 0.39$, red $m = 0.4$, blue $m = 0.41$. Break occurs under one time unit. Shorter $t_{br}$ corresponds to higher $m$. Top straight line - $2.8 \times 10^{-5}$, bottom - $4.3 \times 10^{-5}$. $N = 2^{16}$. 14

3.3 Erratic behavior of $\delta$ after the break point ($m > m_{cr}$), (right to left): green $m = 0.385$, black $m = 0.39$, red $m = 0.4$, blue $m = 0.41$. $N = 2^{16}$. 15

4.1 Solutions comparison for $m < m_{cr}$: $m = 0.1$ - left and $m = 0.35$ - right at the end of one full cycle ($\pi_n$). Red (thicker) line is analytic result, black - thinner - numeric production. 18

4.2 Solutions comparison for $m > m_{cr}$: $m = 383$ - left and $m = 0.4$ - right at $t = t_{br}$. Magenta (thicker) line is analytic result, black - thinner - numeric production. 18
4.3 Errors in the numerically calculated conserved quantities for $m < m_{cr}$: $m = 0.1$ - top, $m = 0.3$ - middle, $m = 0.375$ - bottom. $E_0$ on the left and $E_1$ on the right, displayed over one full breathing period. 20

4.4 Errors in the numerically calculated conserved quantities for $m > m_{cr}$: $m = 0.39$ - top, $m = 0.41$ - bottom. $E_0$ on the left and $E_1$ on the right. Breaking time (black dot) for $m = 0.39$ is 0.558 and 0.315 for $m = 0.41$. 21

4.5 Delta pattern for various $m < m_{cr}$ (top to bottom): green ($\circ$) $m = 0.2$, black ($\Delta$) $m = 0.3$, red ($*$) $m = 0.35$, magenta ($\circ$) $m = 0.37$, blue ($\bigtriangleup$) $m = 0.375$ with $N = 2^{15}$. Lower ms match perfectly, but close to $m_{cr}$ false positives are produced. 23

4.6 Delta pattern for various $m < m_{cr}$ (top to bottom): green ($\circ$) $m = 0.2$, black ($\bigtriangleup$) $m = 0.3$, red ($*$) $m = 0.35$, magenta ($\circ$) $m = 0.37$, blue ($\bigtriangleup$) $m = 0.375$ with $N = 2^{17}$. In proximity to $m_{cr}$, the numerical results improve with higher resolution ($m = 0.35$ matches perfectly, and at $m = 0.37$ there are fewer outliers than with $N = 2^{15}$). 23

4.7 Delta pattern (solid - analytic, - - numeric) for different $m > m_{cr}$ (right to left): green $m = 0.385$, black $m = 0.39$, red $m = 0.4$, blue $m = 0.41$. The breaks are evident and the timing of the break predictions is quite accurate. $N = 2^{16}$. 24

4.8 Oscillations - red stars. Top two graphs - low $m$: $m = 0.1, m = 0.3$ - no oscillations, middle - just below and just above $m_{cr}, m = 0.38, m = 0.383$ - some oscillations, bottom - above $m_{cr}, m = 0.39, m = 0.4$ - oscillations are prominent, after the break. Red stars - positions of oscillations. 26

5.1 Conserved quantities stay constant over time when $a < a_{cr}$: top $a = 0.2$, bottom $a = 1.0$. $N = 2^{15}$. 29

5.2 Conserved quantities lose their consistency in proximity to $a_{cr}$: top $a = 1.15$, middle $a = 1.17$, bottom $a = 1.18$. $N = 2^{15}$. 30

5.3 Conserved quantities break around $t_{br}$ if $a > a_{cr}$: top $a = 1.2$, middle $a = 1.5$, bottom $a = 1.8$. $N = 2^{15}$. 31

5.4 No oscillations are observed in the function when $a < a_{cr}$: left $a = 0.6$, right $a = 1.1$. $N = 2^{15}$. 32

5.5 Some oscillations can be observed near $a_{cr}$: left $a = 1.17$, right $a = 1.18$ at $N = 2^{15}$ but none were present at higher numbers of Fourier modes. 32

5.6 Oscillations are frequent and prominent when $a > a_{cr}$: left to right top to bottom $a = 1.2, a = 1.3, a = 1.5, a = 1.8$. The higher $a$ is, the more oscillations are observed. $N = 2^{15}$. 33

5.7 $b = 0.5$. For smaller values of $a$ (top to bottom: yellow $a = 0.2$, green $a = 0.4$, cyan $a = 0.6$, blue $a = 0.8$, red $a = 1.0$) all the values of $\delta$ are well above zero and therefore the wave does not break in those cases. $N = 2^{15}$. The well-posedness border is at $a \approx 0.28$. 34
5.8 $b = 0.5$. The curves terminate at the point in time where $\delta$ becomes negative. For higher values of $a$ (top to bottom: magenta $a = 1.2$, green $a = 1.4$, cyan $a = 1.6$, blue $a = 1.8$, red $a = 2.0$) breaking occurs quickly and for $a > 1.2$, $\delta$ reaches a negative value before one time unit. $N = 2^{15}$. Pelinovsky’s wave breaking border lies at $a \approx 3.47$. 35

5.9 The analytically determined well-posedness and wave-breaking regions are expanded. Stars indicate confirmation of scaling: the predicted outputs for $b = 1, 10, 20$ based on the results for $b = 0.5$ closely match the actual values. 36

5.10 Analytical well-posedness and wave breaking regions. Final results based on available computer time and resources. 38

5.11 Magnified version of Figure 5.10 to emphasize the accuracy and significance of the remaining unknown region. 39
List of Tables

2.1 Parameters in $\ln |v| = A - B \ln k - k\delta$ and the Ratio of Variances when fitted over the peaks and the whole curve. Starting at mode 200. Breaking time at $m = 0.39$ is 0.558. $N = 2^{15}$ ........................................ 7

4.1 Maximum point-wise errors of numerical integration .................................................. 17
4.2 Differences in conserved quantities between analytical and numerical methods. The comparison is done at the end of one breathing period for $m < m_{cr}$ and at breaking time $t_{br}$ for $m > m_{cr}$. $N = 2^{15}$ ................ 19
4.3 Smallest $\delta s$ for $m < m_{cr}$ and $t_{br}$ for $m > m_{cr}$ produced by analytically and numerically calculated functions. $N = 2^{15}$ ..................... 22
4.4 Oscillations for various values of $m$, which confirm occurrence of the break. The number of oscillations and their size depends on the value of $t$ and the precision of the calculations. The trend is clear: near and above $m_{cr}$, clearly noticeable oscillations are present. ............... 25

5.1 Maximum deviations in conserved quantities over time. The comparison is done at $t = 6$ for $a < a_{cr}$, and at $t_{br}$ for $a > a_{cr}$. $N = 2^{15}$ ....... 28
5.2 $\delta s$ and projected $t_{br}$ for a range of $a$. If $a \leq 1.0$, there is no break. For $a \geq 1.2$ the break is evident and $t_{br}$ can be estimated as the time of the first negative $\delta$. Data are displayed for $2^{15}$ and $2^{16}$ Fourier modes to show convergence. .................. 36
5.3 Comparison of results for equivalent pairs of $a$ and $b$ based on scaling. The base is $b = 0.5$. $N = 2^{15}$ .................. 37
5.4 $\delta s$ and inferred breaking times for different numbers of Fourier modes. $a = 1.01 - 1.07$. .................. 37
5.5 $\delta s$ and inferred breaking times for different numbers of Fourier modes. $a = 1.08 - 1.19$ .................. 38
Chapter 1

Introduction

J. Scott Russell first observed "solitons" while riding a horse along the Union Canal near Edinburgh in 1834. He was so shaken by what he had seen that he described his observations in detail, calling it the Wave of Translation. Russell experimented extensively in his laboratory-scale wave tank in order to study this observed phenomenon [1].

A soliton is a solitary wave that retains its shape and form after colliding with another solitary wave. Since their discovery, solitons have been studied extensively. After Russell’s investigations in the 19th century, solitons were studied by Airy, Stokes, Boussinesq and Rayleigh. These investigations produced various equations but also raised further questions, for example whether inviscid equations of water waves possess solitary-wave solutions. In 1895, Korteweg and de Vries resolved this issue, deriving the following nonlinear equation for a wave propagating in a shallow channel of water:

\[ u_t + 6uu_x + u_{xxx} = 0. \]  

(1.1)

There was not much progress in the study of wave propagation in the first half of the 20th century. However, great strides were made after 1950. Many variations of the equation were introduced, describing wave propagation in different media. One field of research concentrated on describing the propagation of an ultra-short pulse in nonlinear media such as silica optical fiber. The equation

\[ u_{xt} = u + \frac{1}{6}(u^3)_{xx} \]

(1.2)

was discovered at the end of the 20th century by T. Schäfer and C.E. Wayne and is named the Schäfer-Wayne equation, or Short Pulse Equation (SPE). It was derived as an alternative to the nonlinear Schrödinger’s equation (NLSE), which is the standard model for describing propagation of pulses in nonlinear Maxwell’s equations. NLSE has the form

\[ i\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial x^2} + |u|^2u = 0 \]

[1] and its principal applications are in the propagation of light in nonlinear optical fibers. It becomes less accurate as the length of the pulse shortens. In contrast, the SPE provides a better
approximation of the solution to Maxwell's equations when the pulse is as short as a few cycles of the central frequency [10] and [7]. Numerical simulations show that the accuracy of the SPE approximation improves as the pulse shortens [6], whereas the NLSE fails to be accurate for ultra-short pulses.

For some initial conditions, analytical solutions for the SPE are known. The equation is integrable and the analytical solution (2.1) can be generated over time. In other cases, an analytical solution cannot be found and therefore a numerical one is necessary. In Chapter 2, a basis for the research is established. Afterwards, we determine and describe the pulse solution for the SPE, calculate the analytical solution over time, and, in Chapter 3, determine a breaking time based on the identified method.

In Chapter 4 we investigate a method of generating a numerical solution for the pulse IC. We demonstrate that the numeric approach works with a high level of accuracy. Then we apply the same method to an initial condition without any known analytical solution: the modified Gaussian Initial Condition (Chapter 5). To validate our findings further, we evaluate conserved quantities and review oscillations for both the pulse solution and the Gaussian (Chapters 4 and 5). The former can be recalculated analytically and numerical methods are easily evaluated. The results for the latter can be compared with the numerical results of the pulse solution.

In this work we demonstrate that numerical methods can be used to identify the break in the solution for the Short Pulse Equation. We demonstrate that a numerical solution for the pulse initial condition produces results very close to the calculated analytical ones. This accuracy justifies the use of the same methodology in the absence of an analytical solution and allows us to determine the existence of the break in the pulse numerically. An estimate of breaking time can be made at the same time. For the Gaussian IC the developed numerical methods allow us to greatly expand the areas of well-posedness and wave breaking, initially determined analytically.

The numerical solutions were implemented in Matlab. All work was done on a desktop personal computer. If a supercomputer with Matlab were available, the known areas could have been expanded even further by increasing the number of Fourier modes with a corresponding decrease in the time step. These computations would take a significant amount of time, but they are clearly possible.
Chapter 2

Background

2.1 Determined Analytical solution

The SPE is an integrable nonlinear partial differential equation. As shown by Sakovich and Sakovich [10], three possible solutions for the SPE (the loop-antiloop, two-loop, and nonsingular solutions) have very similar parametric forms derived from the sine-Gordon equation $\frac{z_{yt}}{z} = \sin(z)$.

The nonsingular solution $u(x,t)$, of the SPE is given by a pair of parametric equations:

$$u = 4mn \frac{m \sin \psi \sinh \phi + n \cos \psi \cosh \phi}{m^2 \sin^2 \psi + n^2 \cosh^2 \psi}$$

$$x = y + 2mn \frac{m \sin 2\psi - n \sinh 2\phi}{m^2 \sin^2 \psi + n^2 \cosh^2 \phi}$$

(2.1a)

(2.1b)

where $\phi = m(y + t)$, $\psi = n(y - t)$, $n = \sqrt{1 - m^2}$, $-\infty < y < \infty$ is a curve parameter and $m \in (0, 1)$.

This solution of the SPE is also called the pulse solution. This name does not mean, however, that expressions (2.1) represent a nonsingular pulse for any value of the parameter $m : 0 < m < 1$. There is a critical value $m_{cr}$ below which the pulse solution propagates at a constant speed without breaking. Then the solution is single-valued for all time. For $m > m_{cr}$, the solution formula becomes multivalued in finite time and therefore breaks.

For $m < m_{cr}$ the parameter $m$ measures the “shortness” of the pulse relative to the wavelength of oscillations under the pulse envelope [3]. As $m \to 0$, the number of oscillations under the envelope increases, the maximum amplitude decreases, and the envelope width increases (Figure 2.1).

When $m > m_{cr}$ at some finite point in time the derivative $u_x \to \infty$ in one or more points of the solution function $u$. Since $u_x = \frac{u_y}{x_y}$ then $u_x \to \infty$ as $x_y \to 0$. The
derivative can be easily computed from (2.1) and expressed as

\[
\frac{dx}{dy} = 1 + \frac{4m^2n^2(\cos 2\psi - \cosh 2\phi)}{(n^2 \cosh^2 \phi + m^2 \sin^2 \psi)} + \frac{(n \cosh \phi \sinh \phi + m \cos \psi \sin \psi)(n \sinh 2\phi - m \sin 2\psi)}{(n^2 \cosh^2 \phi + m^2 \sin^2 \psi)^2}.
\] (2.2)

This derivative is a good check if the breaking point is properly identified.

The critical value \( m_{cr} \) and the breaking time \( t_{br} \) are very important for our analysis. As shown in Appendix A, \( m_{cr} = \sin(\frac{\pi}{8}) \approx 0.3827 \), and the breaking time when \( m > m_{cr} \) is \( t_{br} = \frac{\arcsin(n\sqrt{4-2\sqrt{2}})}{2n} - \frac{\text{arccosh}(m\sqrt{4+2\sqrt{2}})}{2m} \) (Appendix B).

### 2.2 Detection of wave breaking in Fourier space

The Fourier transform is defined as

\[
\hat{u}(k) = \int_R u(x, t)e^{-ikx}dx.
\] (2.3)
The asymptotic behavior of Fourier coefficients for \( k \to \infty \) can be expressed as

\[
\hat{u} \sim \sqrt{2\pi} \mu^{\frac{\mu+1}{2}} e^{-\mu} \frac{(-i)^{\mu+1}}{k^{\mu+1}} e^{-ika-k\delta}.
\]  

(2.4)

For a single singularity \( \alpha + i\delta \) with \( \delta > 0 \), the modulus of Fourier coefficients decreases exponentially for large \( k \). If \( \delta \to 0 \) (representing a singularity on a real axis) the modulus of the Fourier coefficients has an algebraic dependence on \( k \) as shown by Klein in [8]. The initial condition of the analytical solution (2.1) has the modulus of Fourier coefficients asymptotic but oscillatory (Figure 2.2).

![Figure 2.2: Oscillatory nature of the Fourier coefficients. Natural log of the coefficients across mode number at time zero for \( m = 0.1, m = 0.2 \) - top, \( m = 0.3, m = 0.4 \) - bottom. Black (top) - fitting curve across peaks, red - fitting on the whole curve. \( N = 2^{15} \) modes used.](image)

It is important to select a good fitting technique and approach to the identification of meaningful ranges for \( k \). These have been well developed by Klein [8], taking into account the asymptotic behavior of the Fourier coefficients in order to identify the timing of the break, if it occurs.

As was established [8], there is a relationship between wave breaking (multivaluedness of the solution) and the behavior of the Fourier modes. Since the solution becomes multivalued when the derivative \( x_y \to 0 \), we compared the behavior of \( x_y \) and \( \ln |v| \) where \( |v| \) is the modulus of the Fourier coefficients. The coefficients
diminish exponentially when the solution is away from the pole but become alge-
braically dependent on the wave number in proximity of it. We can use the expres-
sion

$$\ln |v| \sim A - B \ln k - k \delta$$

(2.5)

to approximate the behavior of the Fourier coefficients by performing the least
square fitting for $\ln |v|$, where $k$ is the wave number in Fourier space. The fitting
is done for a range of wave numbers $k$. We need to identify which wave numbers to
include in the fitting. First of all, we include only positive wave numbers since the
coefficients are symmetrical.

To evaluate the accuracy of the fit, the Ratio of Variances is used:

$$RV = \frac{\sum Err_i}{\sum Var_i} , \text{ where}$$

(2.6a)

$$Err_i = \left( \ln |v| - (A - B \ln k - k \delta) \right)^2$$ and

(2.6b)

$$Var_i = \left( \ln |v| - \frac{\sum \ln |v|}{n} \right)^2$$ (2.6c)

The process to identify the best fitting for the Hopf equation is described in de-
tail by Klein [8]. The initial condition in the Hopf equation that Klein analyzes is
$\text{sech}^2(x)$. In Fourier space this function looks almost like a straight line and is rela-
tively easy to get a fitting for (after some adjustments which are described in detail
by Klein) with a very small error. Our IC with the analytic solution or the modi-
fied Gaussian IC are initially oscillatory in the Fourier space and therefore it looks
like the good fitting can only be done on the peaks of the oscillations (black top
line on each graph in Figure 2.2). We define the peaks as the points in the curve
where the function increases at least twice over a threshold ($10^{-6}$) and after that
decreases at least twice over the same threshold. The function that calculates them
returns the position of the peak (Fourier mode number) and the value of the func-
tion at that point which allows us to fit the curve over the set of identified peaks.
We need to recalculate the fitting for each time step and make sure that the crite-
rion that we choose works over time for any $m$ in the analyzed range.

Since for every value of $m$ at time zero, $\ln |v|$ has an oscillatory nature, the fitting
seems to be much better over the peaks of the oscillations. The higher $m$ is the
more peaks we have at $t = 0$ (Figure 2.2). If all modes are included, the fitting over
the peaks is not very accurate. The main difference comes from the smaller modes.
At $N = 2^{15}$ the exclusion of the first 200 modes improves the fitting significantly
(100 in the case of $N = 2^{14}$). At the same time, fitting over the whole curve gener-
ates a curve almost parallel to the one over the peaks with values of $A$ and $B$ being
quite different (2.5) but the value of $\delta$ almost the same. Since we need to track only
changes in $\delta$ over time, it does not make any difference which curve we utilize for
that purpose. In addition, after a few time steps (even at $m = 0.2$) the oscillatory
nature of $\log |v|$ becomes less prominent, the number of peaks is very small ($< 10$)
and fitting over the peaks in these cases is inaccurate and not too meaningful. The facts that we can’t use the peaks fitting for every time step and that the value of $\delta$, when peaks fitting works is almost identical in both cases, led us to use only the fitting of the whole curve even though when oscillations are prominent, peaks fitting seems to produce a good result.

It can also be observed that at high wave numbers rounding errors lead to some fuzzy behavior of the Fourier coefficients (Figure 2.2), and it is prudent to reduce the number of Fourier modes on the high side by $|v| < 10^{-12}$.

Table 2.2 reflects parameters in (2.5) for a few values of $m$ at various times in both cases of fitting. It clearly indicates that $\delta$ in both cases of fitting when number of peaks is sufficient is very close. It also confirms that peaks fitting becomes impossible soon after a few time steps.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$t$</th>
<th>#</th>
<th>$A$</th>
<th>$B$</th>
<th>$\delta$</th>
<th>RV</th>
<th>$A$</th>
<th>$B$</th>
<th>$\delta$</th>
<th>RV</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20</td>
<td>0.0</td>
<td>12</td>
<td>8.139</td>
<td>1.112</td>
<td>0.0068</td>
<td>$1.4 \times 10^{-5}$</td>
<td>1.913</td>
<td>0.179</td>
<td>0.0074</td>
<td>$2.2 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.20</td>
<td>0.25</td>
<td>5</td>
<td>7.630</td>
<td>1.010</td>
<td>0.0070</td>
<td>$3.4 \times 10^{-5}$</td>
<td>2.665</td>
<td>0.325</td>
<td>0.0069</td>
<td>$2.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.20</td>
<td>0.5</td>
<td>4</td>
<td>6.919</td>
<td>0.855</td>
<td>0.0074</td>
<td>$3.6 \times 10^{-6}$</td>
<td>2.786</td>
<td>0.342</td>
<td>0.0067</td>
<td>$1.8 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.20</td>
<td>0.75</td>
<td>3</td>
<td>6.671</td>
<td>0.799</td>
<td>0.0077</td>
<td>$2.0 \times 10^{-7}$</td>
<td>2.841</td>
<td>0.349</td>
<td>0.0067</td>
<td>$1.6 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.20</td>
<td>1.0</td>
<td>3</td>
<td>6.837</td>
<td>0.836</td>
<td>0.0075</td>
<td>$1.8 \times 10^{-7}$</td>
<td>2.819</td>
<td>0.346</td>
<td>0.0067</td>
<td>$1.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.20</td>
<td>2.0</td>
<td>5</td>
<td>7.104</td>
<td>0.896</td>
<td>0.0073</td>
<td>$1.1 \times 10^{-5}$</td>
<td>2.742</td>
<td>0.335</td>
<td>0.0068</td>
<td>$2.0 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.20</td>
<td>3.0</td>
<td>10</td>
<td>9.228</td>
<td>1.324</td>
<td>0.0065</td>
<td>$1.3 \times 10^{-4}$</td>
<td>2.640</td>
<td>0.322</td>
<td>0.0070</td>
<td>$3.1 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.30</td>
<td>0.0</td>
<td>25</td>
<td>6.812</td>
<td>0.849</td>
<td>0.0043</td>
<td>$1.2 \times 10^{-4}$</td>
<td>5.008</td>
<td>0.658</td>
<td>0.0044</td>
<td>$1.8 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.30</td>
<td>0.25</td>
<td>4</td>
<td>5.186</td>
<td>0.499</td>
<td>0.0049</td>
<td>$1.0 \times 10^{-4}$</td>
<td>6.132</td>
<td>0.831</td>
<td>0.0035</td>
<td>$2.8 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.30</td>
<td>0.5</td>
<td>3</td>
<td>6.088</td>
<td>0.746</td>
<td>0.0030</td>
<td>$1.3 \times 10^{-7}$</td>
<td>6.580</td>
<td>0.889</td>
<td>0.0029</td>
<td>$1.7 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.30</td>
<td>0.75</td>
<td>1</td>
<td>Meaningless</td>
<td></td>
<td></td>
<td></td>
<td>6.777</td>
<td>0.914</td>
<td>0.0027</td>
<td>$1.5 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.30</td>
<td>1.0</td>
<td>8</td>
<td>32.138</td>
<td>7.088</td>
<td>-0.0006</td>
<td>$2.1 \times 10^{-3}$</td>
<td>6.725</td>
<td>0.907</td>
<td>0.0028</td>
<td>$1.6 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.30</td>
<td>2.0</td>
<td>18</td>
<td>14.117</td>
<td>2.373</td>
<td>0.0022</td>
<td>$8.0 \times 10^{-2}$</td>
<td>6.349</td>
<td>0.859</td>
<td>0.0032</td>
<td>$2.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.30</td>
<td>3.0</td>
<td>18</td>
<td>18.883</td>
<td>3.317</td>
<td>0.0019</td>
<td>$6.9 \times 10^{-2}$</td>
<td>6.228</td>
<td>0.843</td>
<td>0.0033</td>
<td>$2.4 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

$m_{cr} \approx 0.3827$

<table>
<thead>
<tr>
<th>$m$</th>
<th>$t$</th>
<th>#</th>
<th>$A$</th>
<th>$B$</th>
<th>$\delta$</th>
<th>RV</th>
<th>$A$</th>
<th>$B$</th>
<th>$\delta$</th>
<th>RV</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.39</td>
<td>0.0</td>
<td>53</td>
<td>8.149</td>
<td>0.993</td>
<td>0.0016</td>
<td>$3.0 \times 10^{-4}$</td>
<td>6.988</td>
<td>0.921</td>
<td>0.0016</td>
<td>$2.8 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.39</td>
<td>0.25</td>
<td>4</td>
<td>5.371</td>
<td>0.434</td>
<td>0.0020</td>
<td>$2.5 \times 10^{-4}$</td>
<td>7.300</td>
<td>0.917</td>
<td>0.0020</td>
<td>$2.2 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.39</td>
<td>0.5</td>
<td>1</td>
<td>Meaningless</td>
<td></td>
<td></td>
<td></td>
<td>7.431</td>
<td>0.900</td>
<td>0.0001</td>
<td>$2.1 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.39</td>
<td>0.75</td>
<td>1</td>
<td>Meaningless</td>
<td></td>
<td></td>
<td></td>
<td>6.762</td>
<td>0.754</td>
<td>0.0001</td>
<td>$8.5 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Table 2.1: Parameters in $\ln |v| = A - B \ln k - k\delta$ and the Ratio of Variances when fitted over the peaks and the whole curve. Starting at mode 200. Breaking time at $m = 0.39$ is 0.558. $N = 2^{15}$.
Figure 2.3 demonstrates changes in the $\ln|v|$ curve and its fittings over time for $m = 0.2$ and $m = 0.3$.

![Figure 2.3: Fitting curves for $m = 0.2$ (top) and $m = 0.3$ (bottom). Left $t = 0$, right $t = 1.0$. Blue curve – $\ln|v|$, green -.- fitting over the peaks, red – fitting over the whole curve. $N = 2^{15}$.](image)

For $m > m_{cr}$ a few steps after $t = 0$ oscillations in the curve disappear, therefore fitting over the peaks does not produce any meaningful results (Table 2.2) Figure 2.4 demonstrates the change in the curve without peaks fitting for 3 positions in time for $m = 0.39$.

Final observation: the biggest difference in the whole curve and its fitting (when oscillations disappear after a few time steps for higher $m$) is in the first 50 included modes. The calculated curve (2.5) fits very well except for those few smaller $k$ and if in the calculation of the the ratio of the variance $RV$ we exclude them it becomes better and more meaningful since our model is asymptotic.

### 2.3 Choice of a numerical method

The Matlab package with the EXPINT library [4] of exponential integrators was used to perform the calculations. Exponential integrators are a class of numerical methods specifically designed for the numerical solution of semi-linear problems. They are essentially an alternative to implicit methods for the numerical solution of stiff or highly oscillatory differential equations [4]. These methods aim to exactly solve the linear part of the problem and then numerically solve the non-linear part, making them well suited for the problem at hand.

A number of numerical methods from the EXPINT library were evaluated for the
integration but all of them yielded more or less the same result. We chose to use ETD schemes, in particular etd4rk.

## 2.4 Analytical determination of wave breaking in finite time

The solution for the SPE does not break below the well-posedness curve and always breaks above the wave breaking curve (Figure 2.5) [11]. Theorem 2 in [11] demonstrates that wave breaking may occur only if $2\sqrt{2E_1E_2} \geq 1$. By using this criterion we can reproduce the well-posedness border cited in [11]. Theorem 3 in the same paper provides us with the condition for the wave breaking border:

\begin{align}
F_1 &:= \frac{1}{\sqrt{2}} \left( E_1^2 + (8E_0E_1 + E_1^4)^{1/2} \right)^{1/2} \\
F_0 &:= \frac{1}{\sqrt{2}} \left( E_0 + E_{-1} + \frac{1}{12} E_0 E_1^2 \right)^{1/2}
\end{align}

### References

where $E_0$, and $E_1$ are conserved quantities as in equations (4.1) and $E_{-1}$ is the energy/conserved quantity $E_{-1} := \int_R [(\partial_x^{-1} u)^2 - \frac{1}{12} u^4] \, dx$.

An analytical solution for the Gaussian Initial Condition $u_0(x) = a(1 - 2bx^2)e^{-bx^2}$ does not exist. We will consider this modified Gaussian IC because it satisfies the criterion for the theorems used in this section, i.e. $\int_R u_0(x) \, dx = 0$. As demonstrated in (D.2), $\int a(1 - 2bx^2)e^{-bx^2} \, dx = axe^{-bx^2}$ and therefore is equal to zero on $R$.

$E_0$ and $E_{-1}$ can be derived analytically for the Gaussian IC [11]:

$$E_0 = \frac{3a^2\sqrt{2\pi}}{8\sqrt{b}}, \quad E_{-1} = \frac{a^2\sqrt{\pi}(256\sqrt{2} - 51a^2b)}{2048\sqrt{b^3}}$$

(2.8)

$E_0$ is easily derivable (see Appendix C). The approach to the calculation of $E_{-1}$ is the same (Appendix D). In addition, both conserved quantities were recalculated numerically (numerical integration, calculation of anti-derivative in Fourier space) and the results matched perfectly the numbers produced by the formulas.

Based on Theorem 3, if there exists a point $x_0$ in the function's domain such that $u_0(x_0)u_0'(x_0) > 0$ and either

$$|u_0'(x_0)| > \left( \frac{F_1^2}{4F_0} \right)^{1/3}, \quad |u_0(x_0)||u_0'(x_0)|^2 > F_1 + \left( 2F_0|u_0'(x_0)|^3 - \frac{1}{2} F_1^2 \right)^{1/2}$$

(2.9a)

or

$$|u_0'(x_0)| \leq \left( \frac{F_1^2}{4F_0} \right)^{1/3}, \quad |u_0(x_0)||u_0'(x_0)|^2 > F_1,$$

(2.10a)

then the wave breaks in a finite amount of time. By applying conditions for well-posedness and wave breaking (after calculating all needed conserved quantities) we can generate the borders in Figure 2.5.

Another important point is that there is no need to run the system for all different combinations of parameters $a$ and $b$. We can hold $b$ constant, identify the border between well-posedness and wave breaking for one value of $b$, and scale it for the rest of the values of $b$. If $u(x, t) = Aw(\alpha \psi, \beta \tau)$, then our SPE $u_{xx} = u + \frac{1}{6}(u^3)_{xx}$ becomes $A\alpha\beta w_{\psi\tau} = Aw + \frac{1}{6}A^3\alpha^2(w^3)_{\psi\psi}$ or $w_{\psi\tau} = w + \frac{1}{6}(A\alpha)^2(w^3)_{\psi\psi}$ (if we fix $\alpha = 1/\beta$). If $\alpha = 1/A$ then the same SPE is returned under different parameters: $w_{\psi\tau} = w + \frac{1}{6}(w^3)_{\psi\psi}$, where $u(x, t) = Aw(x/A, At)$.

What does this all mean for the Gaussian IC? The IC $u_0 = a(1 - 2bx^2)e^{-bx^2}$ can be rewritten as $u_0 = a(1 - 2(\sqrt{bx})^2)e^{-(\sqrt{bx})^2}$ and therefore, the solution is the same if
we assume that $A = a$ and $\alpha = \sqrt{b}$ in our scaling. This tells us that as long as

$$a\sqrt{b} = 1$$

(2.11)

the solution is the same, just scaled for different $b$ where $x$ is scaled by $1/A$ and time is ‘stretched’ as $At$. See chapter 5.2 where we determine well-posedness and wave breaking in the ‘gap’.
Chapter 3

Determination of wave breaking in the Analytical Solution of the SPE

3.1 Choice of the time step and the number of Fourier modes

Numerical integration was performed in Fourier space using spectral discretization with a different number of time steps and Fourier modes over a full cycle for various values of $m$ in the analytical solution.

It appeared that even though the time step is important for proper integration, once the minimum step to ensure stability of the integration was satisfied, its reduction did not improve the output. The difference in output was below the required precision. Negative powers of two were used to determine the size of the time step.

Better resolution (doubling the number of modes – Fast Fourier transform is used by the integrators) noticeably improved the results. To compare the analytical and numerical results, $2^{15}$ modes with the appropriate time step (from $2^{-8}$ to $2^{-4}$: the higher $m$ is – the smaller the step needs to be) was used. With each time step the numerical result diverges further from the analytical one but for $m \ll m_c$ the maximum difference is still within $10^{-6}$ after a full cycle; for larger values of $m$ ($\gg m_c$) the prediction of the solution’s breaking time is also quite accurate (within 0.01 of a time unit), as will be demonstrated. In the vicinity of $m_c$ higher precision is necessary, because the results are less stable. A higher number of modes ($N = 2^{16}$ and $N = 2^{17}$) was used to ensure convergence.
3.2 Analytical results. Numerical Determination of time breaking

For all \( m < m_{cr} \) the derivative \( u_x \) never approaches \( \infty \). Because of this, we wanted to observe the behavior of \( \delta \) in (2.5) to calibrate Klein's methodology and ensure that wave breaking occurs only when \( m > m_{cr} \). Therefore, we needed to determine what would be considered a zero for \( \delta \) (i.e. precision) which would identify a break. As we used numerical methods, a value of \( \delta \) below a threshold can be viewed effectively as a zero. For each \( m \) a wave of length \( L = \frac{66}{m} \) was used (determined by experiments to avoid aliasing and have the solution resolved properly). Therefore, in theory, any value of \( \delta < 2\pi L/N \) can not be distinguished numerically from zero [8]. This threshold works perfectly when \( m \ll m_{cr} \). However, as \( m \rightarrow m_{cr} \), \( \delta \) falls below the threshold but the solution does not break. Having the analytical solution we experimented with this threshold and determined that \( \delta \sim 9 \times 10^{-5} \) should be used with \( N = 2^{15} \) Fourier modes (with a higher number of modes, the threshold should be slightly reduced). Figure (3.1) shows the \( \delta \)'s for a range of \( m < m_{cr} \) and Figure 3.2 - for \( m > m_{cr} \).

![Comparison of log(\( \delta \)) over time for \( m < m_{cr} \) (top to bottom): green (o) \( m = 0.2 \), black (\( \wedge \)) \( m = 0.3 \), red (*) \( m = 0.35 \), magenta (o) \( m = 0.37 \), blue (\( \wedge \)) \( m = 0.375 \). Time period is 3 which is approximately one full cycle (\( \frac{\pi}{n} \)). \( N = 2^{16} \).](image)

As we can see from Figure 3.1 there is no break for any \( m < m_{cr} \), and the breaks do occur for higher \( m \). Timing determination is done based on a previously derived formula (B.5). On the other hand, for \( m > m_{cr} \), the calculated breaking time coincides with the point where \( \delta \) achieves its ‘almost zero’ value for the first time (with \( N = 2^{16} \) modes this ‘small’ value of \( \delta \) is between \( 2.8 \times 10^{-5} \) and \( 4.3 \times 10^{-5} \) or \( \sim 4 \times 10^{-5} \)).

An interesting observation is that the ‘smallness’ of \( \delta \) depends on the number of
Fourier modes used. For $N = 2^{14}$ 'almost zero' $\delta$ is $\sim 10^{-4}$, for $N = 2^{15}$ it is $\sim 9 \times 10^{-5}$, and for $N = 2^{16}$ it is $\sim 3 \times 10^{-5}$. Therefore we can conclude that identification of the breaking point based on $\delta$ should be calibrated based on the number of Fourier modes used. We decided to use $N = 2^{15}$ number of modes, since it produces accurate results and allows the solution to be calculated in a reasonable amount of time with the resources available.

In conclusion, the evaluation of the curve that is fitted to the Fourier modes has embedded information that predicts the time of the break, if it exists. The precision of the calculation depends on the number of Fourier modes and needs to be evaluated carefully. If the break does not occur, the shape of the '\(\delta\)-curve' is periodic, while if it does, the shape of the curve after the break is erratic and meaningless (Figure 3.3)
Figure 3.3: Erratic behavior of $\delta$ after the break point ($m > m_{cr}$), (right to left): green $m = 0.385$, black $m = 0.39$, red $m = 0.4$, blue $m = 0.41$. $N = 2^{16}$. 
Chapter 4

Numerical Solution

4.1 The numerical solution of the SPE and comparison to the analytical solution

4.1.1 Accuracy of numerical integration

The Matlab code was written to integrate the SPE in equation (1.2) with the initial condition (2.1). A higher number of Fourier modes produces more accurate results, but because a larger number of modes necessitates a smaller time step eventually the marginal improvement in accuracy relative to the increase in running time becomes very small (as a FFT was used). In Table (4.1.1) errors are presented, with a full cycle is equal $\frac{\pi}{n}$. For $m > m_{cr}$ integration beyond the breaking point does not produce any meaningful results.

In general, the largest error gets bigger for higher $m$ but the period length differs for $m > m_{cr}$ due to the reduction in the breaking time $t_{br}$ which causes what appears to be a slight inconsistency above the critical value of $m$. With a higher number of Fourier modes the results get substantially better (the error is halved or even reduced by an order of 10, with less predictability close to $m_{cr}$). Though the errors increase for higher $m$, most of them are still small in an absolute sense and the analytical solution is virtually indistinguishable from the numerical one, as seen in Figures 4.1 and 4.2. Actually, the larger errors are present only at the points in the solution where the derivative $u_x$ is large (i.e. where the function is steep and approaches the break). If the ten points with the greatest errors (out of more than 32,000) are excluded, the largest remaining error shrinks by a factor of ten.

Figures 4.1 and 4.2 demonstrate how well the numerical solution matches the analytical. The results on these graphs were produced with $N = 2^{15}$. 
<table>
<thead>
<tr>
<th>( m )</th>
<th>( \Delta t )</th>
<th>( t_f \approx \frac{\pi}{N} )</th>
<th>at ( N = 2^{14} )</th>
<th>at ( N = 2^{15} )</th>
<th>at ( N = 2^{16} )</th>
<th>at ( N = 2^{17} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>( 2^{-4} )</td>
<td>3.157</td>
<td>( 7.4 \times 10^{-5} )</td>
<td>( 3.8 \times 10^{-5} )</td>
<td>( 1.9 \times 10^{-5} )</td>
<td>( 9.6 \times 10^{-6} )</td>
</tr>
<tr>
<td>0.20</td>
<td>( 2^{-5} )</td>
<td>3.206</td>
<td>( 1.6 \times 10^{-4} )</td>
<td>( 7.8 \times 10^{-5} )</td>
<td>( 3.9 \times 10^{-5} )</td>
<td>( 1.9 \times 10^{-5} )</td>
</tr>
<tr>
<td>0.30</td>
<td>( 2^{-6} )</td>
<td>3.293</td>
<td>( 2.6 \times 10^{-4} )</td>
<td>( 1.3 \times 10^{-4} )</td>
<td>( 6.4 \times 10^{-5} )</td>
<td>( 3.2 \times 10^{-5} )</td>
</tr>
<tr>
<td>0.35</td>
<td>( 2^{-7} )</td>
<td>3.354</td>
<td>( 3.3 \times 10^{-4} )</td>
<td>( 1.7 \times 10^{-4} )</td>
<td>( 8.3 \times 10^{-5} )</td>
<td>( 4.2 \times 10^{-5} )</td>
</tr>
<tr>
<td>0.37</td>
<td>( 2^{-7} )</td>
<td>3.382</td>
<td>( 8.7 \times 10^{-4} )</td>
<td>( 1.9 \times 10^{-4} )</td>
<td>( 9.6 \times 10^{-5} )</td>
<td>( 4.8 \times 10^{-5} )</td>
</tr>
<tr>
<td>0.375</td>
<td>( 2^{-7} )</td>
<td>3.389</td>
<td>( 2.5 \times 10^{-3} )</td>
<td>( 4.8 \times 10^{-3} )</td>
<td>( 1.0 \times 10^{-4} )</td>
<td>( 5.0 \times 10^{-5} )</td>
</tr>
<tr>
<td>0.38</td>
<td>( 2^{-7} )</td>
<td>3.396</td>
<td>( 7.4 \times 10^{-3} )</td>
<td>( 2.6 \times 10^{-2} )</td>
<td>( 7.5 \times 10^{-4} )</td>
<td>( 1.7 \times 10^{-4} )</td>
</tr>
<tr>
<td>0.382</td>
<td>( 2^{-7} )</td>
<td>3.399</td>
<td>( 1.2 \times 10^{-2} )</td>
<td>( 5.2 \times 10^{-3} )</td>
<td>( 2.2 \times 10^{-3} )</td>
<td>( 8.7 \times 10^{-4} )</td>
</tr>
<tr>
<td>( m_{cr} \approx 0.3827 )</td>
<td>( 0.8t_{br}/t_{br}^{*} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.383</td>
<td>( 2^{-7} )</td>
<td>0.630</td>
<td>( 4.3 \times 10^{-2} )</td>
<td>( 2.5 \times 10^{-2} )</td>
<td>( 7.8 \times 10^{-3} )</td>
<td>( 1.1 \times 10^{-4} )</td>
</tr>
<tr>
<td>0.383</td>
<td>( 2^{-7} )</td>
<td>0.788*</td>
<td>( 8.8 \times 10^{-2} )</td>
<td>( 6.0 \times 10^{-2} )</td>
<td>( 4.3 \times 10^{-2} )</td>
<td>( 3.9 \times 10^{-2} )</td>
</tr>
<tr>
<td>0.385</td>
<td>( 2^{-7} )</td>
<td>0.546</td>
<td>( 3.4 \times 10^{-2} )</td>
<td>( 1.2 \times 10^{-2} )</td>
<td>( 3.4 \times 10^{-3} )</td>
<td>( 3.1 \times 10^{-5} )</td>
</tr>
<tr>
<td>0.385</td>
<td>( 2^{-7} )</td>
<td>0.683*</td>
<td>( 9.9 \times 10^{-2} )</td>
<td>( 2.9 \times 10^{-2} )</td>
<td>( 8.7 \times 10^{-2} )</td>
<td>( 8.2 \times 10^{-3} )</td>
</tr>
<tr>
<td>0.390</td>
<td>( 2^{-7} )</td>
<td>0.446</td>
<td>( 2.4 \times 10^{-2} )</td>
<td>( 6.9 \times 10^{-3} )</td>
<td>( 1.2 \times 10^{-3} )</td>
<td>( 1.0 \times 10^{-5} )</td>
</tr>
<tr>
<td>0.390</td>
<td>( 2^{-7} )</td>
<td>0.558*</td>
<td>( 1.0 \times 10^{-1} )</td>
<td>( 1.6 \times 10^{-2} )</td>
<td>( 6.2 \times 10^{-2} )</td>
<td>( 3.4 \times 10^{-3} )</td>
</tr>
<tr>
<td>0.395</td>
<td>( 2^{-7} )</td>
<td>0.382</td>
<td>( 1.1 \times 10^{-2} )</td>
<td>( 4.1 \times 10^{-3} )</td>
<td>( 6.7 \times 10^{-4} )</td>
<td>( 7.5 \times 10^{-6} )</td>
</tr>
<tr>
<td>0.395</td>
<td>( 2^{-7} )</td>
<td>0.477*</td>
<td>( 6.0 \times 10^{-2} )</td>
<td>( 3.9 \times 10^{-2} )</td>
<td>( 6.2 \times 10^{-2} )</td>
<td>( 5.6 \times 10^{-3} )</td>
</tr>
<tr>
<td>0.400</td>
<td>( 2^{-8} )</td>
<td>0.330</td>
<td>( 2.1 \times 10^{-2} )</td>
<td>( 7.2 \times 10^{-3} )</td>
<td>( 8.3 \times 10^{-4} )</td>
<td>( 1.3 \times 10^{-5} )</td>
</tr>
<tr>
<td>0.400</td>
<td>( 2^{-8} )</td>
<td>0.413*</td>
<td>( 2.7 \times 10^{-2} )</td>
<td>( 9.1 \times 10^{-2} )</td>
<td>( 6.6 \times 10^{-2} )</td>
<td>( 3.0 \times 10^{-3} )</td>
</tr>
<tr>
<td>0.410</td>
<td>( 2^{-8} )</td>
<td>0.252</td>
<td>( 2.3 \times 10^{-2} )</td>
<td>( 4.9 \times 10^{-3} )</td>
<td>( 7.5 \times 10^{-4} )</td>
<td>( 3.4 \times 10^{-6} )</td>
</tr>
<tr>
<td>0.410</td>
<td>( 2^{-8} )</td>
<td>0.315*</td>
<td>( 4.4 \times 10^{-2} )</td>
<td>( 5.5 \times 10^{-2} )</td>
<td>( 6.8 \times 10^{-2} )</td>
<td>( 4.1 \times 10^{-5} )</td>
</tr>
</tbody>
</table>

Table 4.1: Maximum point-wise errors of numerical integration

### 4.1.2 Conserved quantities

The bi-Hamiltonian formulation of the SPE provides a means to determine an infinite number of conserved functionals/quantities, and has become an active direction of research in the field of soliton theory and integrable systems. It provides constants of motion and is a good way to evaluate the numerical method. It is also a good time measure of wave breaking/multivalued evolution in finite time. We used only these two conserved quantities in our evaluation:
Figure 4.1: Solutions comparison for $m < m_{cr}$: $m = 0.1$ - left and $m = 0.35$ - right at the end of one full cycle ($\frac{\pi}{n}$). Red (thicker) line is analytic result, black - thinner - numeric production.

Figure 4.2: Solutions comparison for $m > m_{cr}$: $m = 383$ - left and $m = 0.4$ - right at $t = t_{br}$. Magenta (thicker) line is analytic result, black - thinner - numeric production.

$$E_0 = \int_R u^2 dx$$

$$E_1 = \int_R \frac{u_x^2}{1 + \sqrt{1 + u_x^2}} dx.$$  \hfill (4.1a)

Table 4.2 shows that over time numerically calculated conserved quantities do not change in value. We know that analytically calculated conserved quantities stay constant. We compare numerical results with the analytical and present the output in Table 4.2 as a maximum point-wise error in relation to the analytical calculation. The closer we get to $m_{cr}$ the conserved quantities become less stable.

Once we get close to $m_{cr}$ the results are less reliable, but for $m_{cr} - m < 0.05$ the quantities are stable. For $m > m_{cr}$ conserved quantities are stable almost until the breaking time. Some examples of the conserved quantities for $m < m_{cr}$ are displayed in Figure 4.3.
$M_{\text{max}} = \left( E_{0\text{num}} - E_{0\text{anal}} \right) / E_{0\text{anal}}$

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\Delta_{0\text{max}} = \left( E_{0\text{num}} - E_{0\text{anal}} \right) / E_{0\text{anal}}$</th>
<th>$\Delta_{1\text{max}} = \left( E_{1\text{num}} - E_{1\text{anal}} \right) / E_{1\text{anal}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>$9.9 \times 10^{-11}$</td>
<td>$9.8 \times 10^{-11}$</td>
</tr>
<tr>
<td>0.20</td>
<td>$4.3 \times 10^{-11}$</td>
<td>$4.3 \times 10^{-10}$</td>
</tr>
<tr>
<td>0.30</td>
<td>$5.4 \times 10^{-12}$</td>
<td>$5.3 \times 10^{-8}$</td>
</tr>
<tr>
<td>0.35</td>
<td>$9.7 \times 10^{-13}$</td>
<td>$2.1 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.37</td>
<td>$1.7 \times 10^{-7}$</td>
<td>$2.9 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.375</td>
<td>$3.2 \times 10^{-6}$</td>
<td>$3.9 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.38</td>
<td>$3.1 \times 10^{-5}$</td>
<td>$1.5 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.382</td>
<td>$3.1 \times 10^{-5}$</td>
<td>$5.4 \times 10^{-2}$</td>
</tr>
<tr>
<td>$m_{\text{cr}} = 0.3827$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.383</td>
<td>$2.9 \times 10^{-5}$</td>
<td>$8.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.385</td>
<td>$2.6 \times 10^{-5}$</td>
<td>$2.5 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.39</td>
<td>$7.3 \times 10^{-5}$</td>
<td>$4.3 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.40</td>
<td>$6.6 \times 10^{-5}$</td>
<td>$7.4 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.41</td>
<td>$1.4 \times 10^{-4}$</td>
<td>$4.9 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

Table 4.2: Differences in conserved quantities between analytical and numerical methods. The comparison is done at the end of one breathing period for $m < m_{\text{cr}}$ and at breaking time $t_{br}$ for $m > m_{\text{cr}}$. $N = 2^{15}$.

The graphs in Figure 4.4 demonstrate the change in the error of the numerically determined conserved quantities over time for $m > m_{\text{cr}}$. We can clearly see that up to a few steps before the break point $t_{br}$ the analytical and numerically calculated conserved quantities are identical, but right after, the numerical results are no longer constant, which indicates a break. This is further proof of the validity of the numerical method.

### 4.2 Results of $\delta$ calculations for $m$ under the critical value

For all $m < m_{\text{cr}}$ we need to observe the behavior of $\delta$ in equation (2.5) to make sure that the methodology does not produce any false positives. We want to compare breaking time predictions derived from the $\delta$s generated by the numerical and analytical methods, and observe how well the numerical method can predict the break when $m > m_{\text{cr}}$. We know that

$$t_{br} = \frac{\arcsin \left( n \sqrt{4 - 2\sqrt{2}} \right)}{2n} - \frac{\text{arccosh} \left( m \sqrt{4 + 2\sqrt{2}} \right)}{2m} \quad (4.2)$$
according to (B.5) in Appendix, so the accuracy of the numerical prediction can be evaluated.

Table 4.2 presents the results which were obtained by running various values of $m$
for various times. For \( m < m_{cr} \), a full cycle of \( \frac{\pi}{n} \) units was run. Only one time unit was necessary for \( m > m_{cr} \), as \( t = 1 \) exceeds the breaking time in all cases. The system was run with \( 2^{15} \) Fourier modes and a corresponding time step (see Table 4.1.1 for a specific step size). Since rounding errors led to some fuzzy behavior of Fourier coefficients, only \( |v| > 10^{-12} \) were used in the calculation of \( \delta \) [8].

As we see in Table 4.2 the numerical results are almost the same as the analytical ones except when \( m \) is close to \( m_{cr} \). We do get false positives when \( m \sim m_{cr} \). For \( 0.37 < m < 0.385 \) the breakage is not conclusive in the absence of the analytical results at this level of Fourier resolution. In addition, \( t_{br} \) is too high to be reliable (\( \gg 1 \)). The numerical integration produces almost identical output to the analytical as long as \( |m - m_{cr}| > 0.05 \). For any \( m \ll m_{cr} \) the error is very small (see Figure 4.5) and for any \( m \gg m_{cr} \) the graph of numerical \( \delta \) is perfect up to the breaking point (Figure 4.7).

We can clearly see that even for \( m \sim m_{cr} \) the false positive (small or even negative \( \delta \)) is followed by a recovery, which should not occur in the case of a true break.
An increase in the number of Fourier modes improves the results. With $N = 2^{16}$ the change in the pattern of numerical integration is less prominent. At $N = 2^{17}$ the pattern for $m = 0.35$ completely matches the analytical output, for $m = 0.37 - 0.375$ the results improve if calibration is taken into account ($\delta$ gets smaller with the increase in modes) (Figure 4.6). This confirms that an increase in the number of Fourier modes provides a better resolution, and in proximity to $m_{cr}$ a higher resolution is required to be able to obtain more reliable results.

### 4.3 Results for $m$ larger than the critical value

When $m > m_{cr}$ the derivative $u_x \to \infty$ at some point in time for some value of $x$. This behavior of the derivative coincides with $\delta \to 0$, assuming that any value of $\delta < \delta_0$, some small value which is effectively zero. It appears that the value of $\delta_0$ depends on the number of Fourier modes. The higher the number of modes, the lower the value of $\delta_0$, which effectively predicts the break. For example, for $2^{15}$ modes $\delta_0 = 9 \times 10^{-5}$, and for $2^{16}$ modes it decreases to $10^{-5}$. We can see in Table 4.2 that the numerical prediction of the time of the break is quite accurate for ev-

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\text{Anl min } \delta$</th>
<th>$\text{Num min } \delta$</th>
<th>$\text{Act break } t_{br}$</th>
<th>$\text{Num break } t_{br}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>$6.0 \times 10^{-3}$</td>
<td>$6.0 \times 10^{-3}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>$6.6 \times 10^{-3}$</td>
<td>$6.6 \times 10^{-3}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>$2.7 \times 10^{-3}$</td>
<td>$2.7 \times 10^{-3}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.35</td>
<td>$8.2 \times 10^{-4}$</td>
<td>$8.2 \times 10^{-4}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.37</td>
<td>$2.9 \times 10^{-4}$</td>
<td>$1.9 \times 10^{-4}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.373</td>
<td>$2.3 \times 10^{-4}$</td>
<td>$&lt; 0$</td>
<td>$\sim 1.72$</td>
<td></td>
</tr>
<tr>
<td>0.375</td>
<td>$1.8 \times 10^{-4}$</td>
<td>$&lt; 0$</td>
<td>$\sim 1.70$</td>
<td></td>
</tr>
<tr>
<td>0.38</td>
<td>$1.0 \times 10^{-4}$</td>
<td>$&lt; 0$</td>
<td>$\sim 1.65$</td>
<td></td>
</tr>
<tr>
<td>0.382</td>
<td>$0.8 \times 10^{-4}$</td>
<td>$&lt; 0$</td>
<td>$\sim 1.60$</td>
<td></td>
</tr>
<tr>
<td>$m_{cr} = \sin \frac{\pi}{8} \approx 0.3827$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.383</td>
<td></td>
<td>0.788</td>
<td>0.742</td>
<td></td>
</tr>
<tr>
<td>0.385</td>
<td></td>
<td>0.683</td>
<td>0.680</td>
<td></td>
</tr>
<tr>
<td>0.390</td>
<td></td>
<td>0.558</td>
<td>0.563</td>
<td></td>
</tr>
<tr>
<td>0.395</td>
<td></td>
<td>0.477</td>
<td>0.484</td>
<td></td>
</tr>
<tr>
<td>0.400</td>
<td></td>
<td>0.413</td>
<td>0.417</td>
<td></td>
</tr>
<tr>
<td>0.410</td>
<td></td>
<td>0.315</td>
<td>0.316</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3: Smallest $\delta$s for $m < m_{cr}$ and $t_{br}$ for $m > m_{cr}$ produced by analytically and numerically calculated functions. $N = 2^{15}$.
Figure 4.5: Delta pattern for various $m < m_{cr}$ (top to bottom): green ($\circ$)$m = 0.2$, black ($\triangle$)$m = 0.3$, red ($*$)$m = 0.35$, magenta ($\circ$)$m = 0.37$, blue ($\triangle$)$m = 0.375$ with $N = 2^{15}$. Lower $m$s match perfectly, but close to $m_{cr}$ false positives are produced.

Figure 4.6: Delta pattern for various $m < m_{cr}$ (top to bottom): green ($\circ$)$m = 0.2$, black ($\triangle$)$m = 0.3$, red ($*$)$m = 0.35$, magenta ($\circ$)$m = 0.37$, blue ($\triangle$)$m = 0.375$ with $N = 2^{17}$. In proximity to $m_{cr}$, the numerical results improve with higher resolution ($m = 0.35$ matches perfectly, and at $m = 0.37$ there are fewer outliers than with $N = 2^{15}$).

Every value of $m > m_{cr}$ especially if $m$ is not too close to the critical value. Figure 4.7 depicts the $\delta$s above the critical point.

From the output in Figure 4.7 we can see that for $2^{16}$ modes our criterion for the
break is when $\delta_0 = 10^{-5}$. The numerical prediction is almost the same as the analytic.

### 4.4 Oscillations in the numerical solution as another indication of breaking

Another indicator of a possible breakage was observed, making us more confident in our methodology: as the pulse approaches and passes the breaking time, multiple minor oscillations of the solution are present (the Gibbs Effect). Oscillations are small peaks and troughs that appear in the function and it loses its ‘smoothness’. If oscillations appear, they come in large numbers and make the pulse look broken. To identify these oscillations we found the regions of the pulse which had two or more consecutive peak-trough patterns.

As expected, no oscillations occur if $m \ll m_{cr}$ but some do appear as $m$ approaches the critical value and are prominent when $m > m_{cr}$. In Figure 4.8 we show the solution at $t_{br}$ for $m > m_{cr}$ and at $t = \frac{\pi}{n}$ for $m < m_{cr}$ with the red stars indicating small oscillations in the solution.

The summary Table 4.4 presents observed oscillations for different values of $m$.

If any doubt of the breakage exists based on analysis of the indicator of exponential $\delta$ decay, we can always review the oscillations to arrive at even more certainty in our conclusions. Again, we can see that for $m \sim m_{cr}$ ($0.382 \leq m \leq 0.383$) the
Table 4.4: Oscillations for various values of \( m \), which confirm occurrence of the break. The number of oscillations and their size depends on the value of \( t \) and the precision of the calculations. The trend is clear: near and above \( m_{cr} \), clearly noticeable oscillations are present.

results are inconclusive, but a step below \( (m < 0.382) \) or a step above \( (m > 0.383) \) the picture is clear. The visual representation of oscillations for various values of \( m \) is in Figure 4.8.
Figure 4.8: Oscillations - red stars. Top two graphs – low \(m\) : \(m = 0.1, m = 0.3\) - no oscillations, middle - just below and just above \(m_{cr}\), \(m = 0.38, m = 0.383\) - some oscillations, bottom - above \(m_{cr}\), \(m = 0.39, m = 0.4\) - oscillations are prominent, after the break. Red stars - positions of oscillations.
Chapter 5

Detection of wave breaking for the Gaussian initial condition

As we established, our numeric integration is reliable and we can use the value of the exponential coefficient $\delta$ in the curve fitting formula (2.5) to accurately predict wave breaking. If we perform similar analysis for an IC with no analytical solution with respect to time, we should be able to determine with certainty whether the solution breaks at a finite point in time. Consider the modified Gaussian IC:

$$u_0(x) = a(1 - 2bx^2)e^{-bx^2}$$  \hspace{1cm} (5.1)

The areas of well-posedness and wave breaking for this IC have been identified analytically and were presented previously in Chapter 2.4. The gap between these two areas is quite large and has not yet been diminished by analytical methods. We will demonstrate how the gap can be significantly reduced using our numerical method.

5.1 Conserved quantities for Gaussian IC

We previously evaluated and calculated the conserved quantities $E_0$ and $E_1$ for the pulse solution (4.1). We wish to repeat the operation for the Gaussian IC to confirm the validity of the calculations. Let us assume there exists a critical value $a_{cr}$, analogous to $m_{cr}$ for the pulse solution. By our numerical calculation in Chapter 5.4, $1.12 < a_{cr} < 1.17$ at $b = 0.5$.

Table 5.1 shows that over time, the conserved quantities do not change in value for small $a$. As $a_{cr}$ is approached, the conserved quantities become less stable. In the case of the pulse solution we calculated the $\Delta_0^{\text{max}}$ as the percent difference between the numerical solution and the analytical one. In the case of the Gaussian IC, an analytical solution does not exist and therefore the evaluation of stability is done against the initial value of each conserved quantity (e.g. $E_0^{(t=0)}$).
Table 5.1: Maximum deviations in conserved quantities over time. The comparison is done at
$t = 6$ for $a < a_{cr}$, and at $t_{br}$ for $a > a_{cr}$, $N = 2^{15}$.

<table>
<thead>
<tr>
<th>$a$</th>
<th>$\Delta_0^{\text{max}} = (E_0^{t=\text{end}} - E_0^{t=0})/E_0^{t=0}$</th>
<th>$\Delta_1^{\text{max}} = (E_1^{t=\text{end}} - E_1^{t=0})/E_1^{t=0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20</td>
<td>$2.1 \times 10^{-7}$</td>
<td>$1.2 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.40</td>
<td>$2.1 \times 10^{-7}$</td>
<td>$3.5 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.60</td>
<td>$2.0 \times 10^{-7}$</td>
<td>$4.8 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.80</td>
<td>$1.9 \times 10^{-7}$</td>
<td>$7.7 \times 10^{-7}$</td>
</tr>
<tr>
<td>1.00</td>
<td>$1.7 \times 10^{-7}$</td>
<td>$1.6 \times 10^{-6}$</td>
</tr>
<tr>
<td>1.10</td>
<td>$1.8 \times 10^{-7}$</td>
<td>$2.0 \times 10^{-6}$</td>
</tr>
<tr>
<td>1.12</td>
<td>$1.8 \times 10^{-7}$</td>
<td>$2.1 \times 10^{-6}$</td>
</tr>
<tr>
<td>$a_{cr}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.17</td>
<td>$7.0 \times 10^{-8}$</td>
<td>$5.1 \times 10^{-6}$</td>
</tr>
<tr>
<td>1.18</td>
<td>$7.0 \times 10^{-8}$</td>
<td>$3.5 \times 10^{-5}$</td>
</tr>
<tr>
<td>1.20</td>
<td>$9.2 \times 10^{-9}$</td>
<td>$1.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.30</td>
<td>$3.1 \times 10^{-9}$</td>
<td>$1.6 \times 10^{-1}$</td>
</tr>
<tr>
<td>1.50</td>
<td>$2.1 \times 10^{-9}$</td>
<td>$3.7 \times 10^{-1}$</td>
</tr>
<tr>
<td>1.80</td>
<td>$8.7 \times 10^{-8}$</td>
<td>$5.4 \times 10^{-1}$</td>
</tr>
</tbody>
</table>

The maximum error for $E_0$ does not change much with $a$, but the error in $E_1$ is
quite similar to what we experienced with the pulse IC: the error increases as $a$
does. However, in the critical zone the change in maximum error is less significant
than what we saw before for the pulse solution. Some difference from the pulse
case is expected due to the difference in the evaluation method.

Figure 5.1 displays $E_0$ and $E_1$ over six time units for $a = 0.2$ and $a = 1.0$.

Once we are within 0.02 of $a_{cr}$, conserved quantity $E_1$ eventually loses consistency
(Figure 5.2).

At $a \gg a_{cr}$, $E_1$ drastically changes just before the projected $t_{br}$, which serves to
confirm the break (Figure 5.3).

### 5.2 Oscillations in the numerical solution for higher values of $a$

We saw that for the analytical solution of the SPE the solution function experi­
enced significant oscillations at higher values of $m$. This phenomenon served as a
confirmation of the break. If we run the Gaussian solution for higher values of $a$
Figure 5.1: Conserved quantities stay constant over time when $a \ll a_{cr}$: top $a = 0.2$, bottom $a = 1.0$. $N = 2^{15}$.

(which we believe are in the breaking zone), we would expect to see similar oscillations as a confirmation of the break. If the break does not occur, there should be no oscillations present at any time (for $a \ll a_{cr}$), but if it does, oscillations should be evident near $t_{br}$.

The results for $a \ll a_{cr}$ are in Figure 5.4.

In the vicinity of $a_{cr}$ the results are inconclusive. At lower numbers of Fourier modes we observed some oscillations, but there were none for higher $N$ (Figure 5.5).

The results for $a \gg a_{cr}$ are self-explanatory (Figure 5.6).

Figures 5.4 and 5.6 confirm our conclusions.
5.3 Expanding the areas of well-posedness and wave-breaking through numerical methods

5.3.1 Numerical determination of borders based on $b = 0.5$

We would like to use the previously developed numerical methodology to better identify the areas of well-posedness and wave-breaking.
Figure 5.3: Conserved quantities break around $t_{br}$ if $a > a_{cr}$: top – $a = 1.2$, middle – $a = 1.5$, bottom – $a = 1.8$. $N = 2^{15}$.

Since $b$ and $x$ are scalable we can fix the value of $b$ at 0.5 and run the system for the set of values of $a$ from 0.2 (definitely well-posed) to 2.0 (definitely breaking). The gap can be shrunk by scaling the values of $a$ identified for $b = 0.5$. For smaller values of $a$ (0.2 – 1.0) we get the following $\delta s$ on a logarithmic scale over time (Figure 5.7).

If we run the system for higher values of $a$ (1.3 – 2.0) the break is evident and we can easily detect it when $\delta$ becomes zero. If $a$ is in the wave breaking zone, the break occurs almost instantaneously; the higher the value of $a$, the sooner break
occurs. Over time, $\delta$ diminishes and never recovers (Figure 5.8). Numerous oscillations are also present at the breaking time (the Gibbs Effect). Results are almost identical for $2^{14} - 2^{16}$ Fourier modes and therefore we can conclude that they converge. Breaking times for $a \geq 1.2$ are quite clear from Figure 5.8. For example, for $a = 1.2$, $t_{br} \approx 1.25$ (first occurrence of very small $\delta$) or $t_{br} \approx 1.95$ (first occurrence of negative $\delta$), but regardless of the exact value of $t_{br}$ the break is evident.

Table 5.2 shows the results of numerical integration for low and high values of $a$ for $N = 2^{15}$ and $N = 2^{16}$ to demonstrate the convergence of the integration.

So far, we have expanded the areas of well-posedness and wave-breaking. The unknown range is now much smaller: for $b = 0.5$ it has been reduced from $0.28 - 3.47$ to $1.0 - 1.2$ (Figure 5.9).
Figure 5.6: Oscillations are frequent and prominent when \( a \gg a_{cr} \): left to right top to bottom \( a = 1.2, a = 1.3, a = 1.5, a = 1.8 \). The higher \( a \) is, the more oscillations are observed. \( N = 2^{15} \).

### 5.3.2 Numerical verification of scaling of the results

Based on our calculations so far for \( b_{orig} = b = 0.5 \), the border of the region of well-posedness was expanded to \( a = 1 \), and the border of the wave-breaking region was expanded to \( a = 1.2 \).

To numerically verify that scaling works, we ran the system for three values of \( a \) corresponding to \( b = 1, 10 \) and \( 20 \), which should have produced scaled but effectively the same results as for \( b = 0.5 \). We expected a minimum \( \delta \) (when a break did not occur) and a breaking time (when it did) to be calculable based on scaling:

\[
\delta_{new} = \delta_{orig} \cdot a\sqrt{b_{new}}/\sqrt{b_{orig}} \quad \text{and} \quad t_{br} = t_{br}^{orig} \cdot a\sqrt{b_{new}}/\sqrt{b_{orig}}.
\]

Both values behaved as expected and are summarized in Table 5.3. Therefore, \( \text{min}(\delta) \) and \( t_{br} \) can be calculated for any value of \( b \) based on the initially calculated and calibrated \( \text{min}(\delta) \) and \( t_{br} \) (adjusted for \( \delta_0 \)) for \( b = 0.5 \).

The corresponding values of \( a \) are, of course, approximate and no perfect match is expected due to a multitude of rounding errors, but Table 5.3 confirms the accuracy of scaling. The stars in Figure 5.9 indicate confirmed results.
Figure 5.7: $b = 0.5$. For smaller values of $a$ (top to bottom: yellow $a = 0.2$, green $a = 0.4$, cyan $a = 0.6$, blue $a = 0.8$, red $a = 1.0$) all the values of $\delta$ are well above zero and therefore the wave does not break in those cases. $N = 2^{15}$. The well-posedness border is at $a \approx 0.28$.

5.4 Further improvement of the results by using a higher number of Fourier modes

We would now like to narrow the gap even further. Depending on the number of Fourier modes used, we get different outputs in this range ($1.0 - 1.2$). We need to make sure that the produced results converge, i.e. that the output does not change with an increase in the number Fourier modes.

Table 5.4 reflects produced minimum $\delta$s and breaking times, if applicable, for the range in question. We split the range into two parts based on the convergence factor. The first part covers all values of $a$ for which $\delta$ does not reach zero at $N = 2^{15}$. It produces the range $1.01 < a < 1.07$.

As we can see for all values of $a$ in this range, $2^{15}$ and $2^{16}$ Fourier modes produce almost identical results (the difference in minimum $\delta$ is less than $10^{-5}$). The smallest $\delta$ is around $10^{-3}$ which is definitely above the threshold. Also, there are no oscillations in the solution within this range.

The second part contains the rest of the values of $a$ in the range in question, $1.08 \leq a \leq 1.19$. All values of $a$ in this range produce wave breaking at $2^{15}$ modes. We need to see if the break still occurs at the same time if $2^{16}$ and $2^{17}$ Fourier modes are used. Then, we can assert that results converge and draw our conclusions (Ta-
Figure 5.8: $b = 0.5$. The curves terminate at the point in time where $\delta$ becomes negative. For higher values of $a$ (top to bottom: magenta $a = 1.2$, green $a = 1.4$, cyan $a = 1.6$, blue $a = 1.8$, red $a = 2.0$) breaking occurs quickly and for $a > 1.2$, $\delta$ reaches a negative value before one time unit. $N = 2^{15}$. Pelinovsky’s wave breaking border lies at $a \approx 3.47$.

As we can see, the results converge for $a \leq 1.12$, with no break (as the smallest $\delta$ is still around $10^{-3}$). The results also converge for $a \geq 1.17$, as there is a break at almost the same time regardless of the value of $N$. For $1.12 < a < 1.17$ the results are inconclusive: At $2^{16}$ modes we see a break, but at $2^{17}$ modes we can’t be sure if there is a break or not since the $\delta$ is small, but remains above 0. Without running more analysis with an even higher number of modes, as well as further evaluation of acceptable minimum $\delta$s, no conclusion can be made with certainty.

Nevertheless, we have been able to shrink the area of the unknown even further (Figures 5.10 and 5.11).
Table 5.2: $\delta$s and projected $t_{br}$ for a range of $a$s. If $a \leq 1.0$, there is no break. For $a \geq 1.2$ the break is evident and $t_{br}$ can be estimated as the time of the first negative $\delta$. Data are displayed for $2^{15}$ and $2^{16}$ Fourier modes to show convergence.

<table>
<thead>
<tr>
<th>$a$</th>
<th>$2^{15}$ modes. min($\delta$)/$t_{br}$</th>
<th>$2^{16}$ modes. min($\delta$)/$t_{br}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>$3.88 \times 10^{-2}$</td>
<td>$3.87 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.4</td>
<td>$2.20 \times 10^{-2}$</td>
<td>$2.20 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.6</td>
<td>$1.25 \times 10^{-2}$</td>
<td>$1.25 \times 10^{-2}$</td>
</tr>
<tr>
<td>0.8</td>
<td>$6.22 \times 10^{-3}$</td>
<td>$6.21 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.0</td>
<td>$2.15 \times 10^{-3}$</td>
<td>$2.15 \times 10^{-3}$</td>
</tr>
<tr>
<td>$a_{cr}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>$\delta &lt; 0, t_{br} = 1.97$</td>
<td>$\delta &lt; 0, t_{br} = 1.91$</td>
</tr>
<tr>
<td>1.4</td>
<td>$\delta &lt; 0, t_{br} = 0.76$</td>
<td>$\delta &lt; 0, t_{br} = 0.77$</td>
</tr>
<tr>
<td>1.6</td>
<td>$\delta &lt; 0, t_{br} = 0.55$</td>
<td>$\delta &lt; 0, t_{br} = 0.55$</td>
</tr>
<tr>
<td>1.8</td>
<td>$\delta &lt; 0, t_{br} = 0.42$</td>
<td>$\delta &lt; 0, t_{br} = 0.42$</td>
</tr>
<tr>
<td>2.0</td>
<td>$\delta &lt; 0, t_{br} = 0.34$</td>
<td>$\delta &lt; 0, t_{br} = 0.34$</td>
</tr>
</tbody>
</table>

Figure 5.9: The analytically determined well-posedness and wave-breaking regions are expanded. Stars indicate confirmation of scaling: the predicted outputs for $b = 1, 10, 20$ based on the results for $b = 0.5$ closely match the actual values.
### Table 5.3: Comparison of results for equivalent pairs of $a$ and $b$ based on scaling. The base is $b = 0.5$. $N = 2^{15}$.

<table>
<thead>
<tr>
<th>$b$</th>
<th>$a$</th>
<th>Predicted min($\delta$)/$t_{br}$</th>
<th>Actual min($\delta$)/$t_{br}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1.0</td>
<td>well-posedness border</td>
<td>2.2 \times 10^{-3}</td>
</tr>
<tr>
<td>1.0</td>
<td>0.707107</td>
<td>3.0 \times 10^{-3}</td>
<td>3.0 \times 10^{-3}</td>
</tr>
<tr>
<td>10.0</td>
<td>0.223607</td>
<td>9.7 \times 10^{-3}</td>
<td>9.7 \times 10^{-3}</td>
</tr>
<tr>
<td>20.0</td>
<td>0.158114</td>
<td>13.9 \times 10^{-3}</td>
<td>13.9 \times 10^{-3}</td>
</tr>
<tr>
<td>0.5</td>
<td>1.2</td>
<td>wave breaking border</td>
<td>$\delta &lt; 0, t_{br} = 1.97$</td>
</tr>
<tr>
<td>1.0</td>
<td>0.848528</td>
<td>$\delta &lt; 0, t_{br} = 2.7$</td>
<td>$\delta &lt; 0, t_{br} = 2.8$</td>
</tr>
<tr>
<td>10.0</td>
<td>0.268328</td>
<td>$\delta &lt; 0, t_{br} = 8.7$</td>
<td>$\delta &lt; 0, t_{br} = 8.2$</td>
</tr>
<tr>
<td>20.0</td>
<td>0.189737</td>
<td>$\delta &lt; 0, t_{br} = 12.5$</td>
<td>$\delta &lt; 0, t_{br} = 11.6$</td>
</tr>
</tbody>
</table>

### Table 5.4: $\delta$s and inferred breaking times for different numbers of Fourier modes. $a = 1.01 - 1.07$.

<table>
<thead>
<tr>
<th>$a$</th>
<th>$2^{15}$ modes. min($\delta$)/$t_{br}$</th>
<th>$2^{16}$ modes. min($\delta$)/$t_{br}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.01</td>
<td>$\delta = 2.00 \times 10^{-3}$</td>
<td>$\delta = 2.00 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.02</td>
<td>$\delta = 1.85 \times 10^{-3}$</td>
<td>$\delta = 1.85 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.03</td>
<td>$\delta = 1.71 \times 10^{-3}$</td>
<td>$\delta = 1.70 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.04</td>
<td>$\delta = 1.57 \times 10^{-3}$</td>
<td>$\delta = 1.56 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.05</td>
<td>$\delta = 1.44 \times 10^{-3}$</td>
<td>$\delta = 1.43 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.06</td>
<td>$\delta = 1.31 \times 10^{-3}$</td>
<td>$\delta = 1.30 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.07</td>
<td>$\delta = 0.83 \times 10^{-3}$</td>
<td>$\delta = 1.18 \times 10^{-3}$</td>
</tr>
</tbody>
</table>
Table 5.5: $\delta$s and inferred breaking times for different numbers of Fourier modes. $a = 1.08 - 1.19$

<table>
<thead>
<tr>
<th>$a$</th>
<th>$2^{15}$ modes. min($\delta$)/$t_{br}$</th>
<th>$2^{16}$ modes. min($\delta$)/$t_{br}$</th>
<th>$2^{17}$ modes. min($\delta$)/$t_{br}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.08</td>
<td>$\delta &lt; 0/t_{br} = 3.13$</td>
<td>$\delta = 1.06 \times 10^{-3}$</td>
<td>$\delta = 1.06 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.09</td>
<td>$\delta &lt; 0/t_{br} = 2.47$</td>
<td>$\delta = 0.94 \times 10^{-3}$</td>
<td>$\delta = 0.94 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.10</td>
<td>$\delta &lt; 0/t_{br} = 2.42$</td>
<td>$\delta = 0.84 \times 10^{-3}$</td>
<td>$\delta = 0.83 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.11</td>
<td>$\delta &lt; 0/t_{br} = 2.38$</td>
<td>$\delta = 0.73 \times 10^{-3}$</td>
<td>$\delta = 0.73 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.12</td>
<td>$\delta &lt; 0/t_{br} = 2.33$</td>
<td>$\delta = 0.64 \times 10^{-3}$</td>
<td>$\delta = 0.63 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.13</td>
<td>$\delta &lt; 0/t_{br} = 2.28$</td>
<td>$\delta &lt; 0/t_{br} = 2.42$</td>
<td>$\delta = 0.54 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.14</td>
<td>$\delta &lt; 0/t_{br} = 2.22$</td>
<td>$\delta &lt; 0/t_{br} = 2.19$</td>
<td>$\delta = 0.45 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.15</td>
<td>$\delta &lt; 0/t_{br} = 1.09$</td>
<td>$\delta &lt; 0/t_{br} = 2.14$</td>
<td>$\delta = 0.37 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.16</td>
<td>$\delta &lt; 0/t_{br} = 2.13$</td>
<td>$\delta &lt; 0/t_{br} = 2.09$</td>
<td>$\delta = 0.30 \times 10^{-3}$</td>
</tr>
<tr>
<td>1.17</td>
<td>$\delta &lt; 0/t_{br} = 2.09$</td>
<td>$\delta &lt; 0/t_{br} = 2.03$</td>
<td>$\delta &lt; 0/t_{br} = 2.00$</td>
</tr>
<tr>
<td>1.18</td>
<td>$\delta &lt; 0/t_{br} = 1.05$</td>
<td>$\delta &lt; 0/t_{br} = 1.98$</td>
<td>$\delta &lt; 0/t_{br} = 1.95$</td>
</tr>
<tr>
<td>1.19</td>
<td>$\delta &lt; 0/t_{br} = 2.02$</td>
<td>$\delta &lt; 0/t_{br} = 1.94$</td>
<td>$\delta &lt; 0/t_{br} = 1.89$</td>
</tr>
</tbody>
</table>

Figure 5.10: Analytical well-posedness and wave breaking regions. Final results based on available computer time and resources.
Figure 5.11: Magnified version of Figure 5.10 to emphasize the accuracy and significance of the remaining unknown region.
Appendix A

Derivation of Critical value in the analytical solution $m_{cr}$

Since the solution for the SPE is derived from the solution for the sine-Gordon Equation (SGE) the derivative of the function $u_x$ can be expressed as [10]

$$u_x = \tan \left( -4 \arctan \frac{m \sin \psi}{n \cosh \phi} \right) \quad (A.1)$$

By using the double angle tangent formula $\tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$ we can get the formula for $\tan 4x$ as

$$\tan 4x = \frac{2 \tan 2x}{1 - \tan^2 2x} = \frac{4 \tan x}{1 - \tan^2 x \left( 1 - \frac{4 \tan^2 x}{(1 - \tan^2 x)^2} \right)} \quad (A.2)$$

$$= \frac{4 \tan x (1 - \tan^2 x)}{1 - 6 \tan^2 x + \tan^4 x}$$

$$= \frac{4 \tan x (1 - \tan x) (1 + \tan x)}{(\tan^2 x - 3 - 2\sqrt{2})(\tan^2 x - 3 + 2\sqrt{2})}$$

Based on this result, we can conclude that $u_x \to \infty$ when $\tan x \to \pm \sqrt{3 \pm 2\sqrt{2}}$. Therefore (by applying this result to (A.1)), $u_x \to \infty$ when

$$\frac{\sin \psi}{\cosh \phi} \to \pm \frac{n}{m} \sqrt{3 \pm 2\sqrt{2}}$$

or to be precise, we need to see when

$$\frac{\sin \psi}{\cosh \phi} = \pm \frac{n}{m} \sqrt{3 \pm 2\sqrt{2}} \quad (A.3)$$
The absolute value of the left side of (A.3) ≤ 1 (sin θ ≤ 1 and cosh θ ≥ 1). On the right side, m < n (it’s known that for m > 0.5 the solution is multi-valued, the proof of which is not difficult but rather long and beside the point); the ± sign on the right side comes into play only if ψ < 0, which we should keep in mind for future derivations. Therefore we can simplify (A.3) to

\[ \frac{\sin \psi}{\cosh \phi} = \pm \frac{n}{m} \sqrt{3 - 2\sqrt{2}}. \]  

(A.4)

The lowest value of m that achieves the equality in (A.4) is when \( \frac{\sin \phi}{\cosh \psi} = 1 \). We also know that \( m = \sqrt{1 - n^2} \); this means that now we can obtain the critical value of m as

\[ \frac{m_{cr}}{\sqrt{1 - m_{cr}^2}} = \sqrt{3 - 2\sqrt{2}} \implies m_{cr}^2 = 3 - 2\sqrt{2} - 3m_{cr}^2 + 2\sqrt{2}m_{cr}^2 \implies m_{cr} = \sqrt{\frac{3 - 2\sqrt{2}}{4 - 2\sqrt{2}}} = \frac{\sqrt{2 - \sqrt{2}}}{2} \implies m_{cr} = \sqrt{\frac{1 - \frac{\sqrt{2}}{2}}{2}} = \sin \left( \frac{\pi}{8} \right) \approx 0.3827. \]

(A.5)
Appendix B

Derivation of $t_{br}$ for $m > m_{cr}$

In the previous section we identified the critical value of $m_{cr}$, below which the solution is single valued at all times and above which the solution necessarily becomes multivalued. We need to identify what that breaking time is, since that is the point of shock.

We know that $\phi = m(y + t) = f(y)$, therefore $y = f^{-1}(\phi)$. Since $\psi = n(y - t)$ we get

$$\psi = n \left( \frac{\phi}{m} - 2t \right) = \frac{n}{m} \phi - 2nt$$  \hspace{1cm} (B.1)

Based on (A.4) and (B.1) we can obtain

$$\sin \left( \frac{n}{m} \phi - 2nt \right) = \pm \cosh(\phi) \cdot \frac{n}{m} \sqrt{3 - \sqrt{2}} \implies$$

$$\frac{n}{m} \phi - 2nt = \pm \arcsin \left( \cosh(\phi) \cdot \frac{n}{m} \sqrt{3 - \sqrt{2}} \right) \implies$$  \hspace{1cm} (B.2)

$$t = \frac{\phi}{2m} \mp \frac{\arcsin \left( \cosh(\phi) \cdot \frac{n}{m} \sqrt{3 - \sqrt{2}} \right)}{2n}$$

To simplify further derivations let’s assign $r = \frac{n}{m} \sqrt{3 - \sqrt{2}}$. The break occurs when $t_\phi = 0$.

Based on (B.2) we get

$$t_\phi = \frac{1}{2m} - \frac{r \sinh \phi}{2n \sqrt{1 - \frac{r^2}{2} - \frac{r^2}{2} \cosh(2\phi)}}$$  \hspace{1cm} (B.3)

If we make the derivative equal to zero we obtain:
\[
\frac{1}{2m} = \frac{n \sinh \phi \sqrt{3 - 2\sqrt{2}}}{2nm \sqrt{1 - \frac{r^2}{2} - \frac{r^2}{2} \cosh(2\phi)}} \implies \\
(3 - 2\sqrt{2}) \sinh^2 \phi = 1 - \frac{r^2}{2} - \frac{r^2}{2} \cosh(2\phi) \implies \\
\text{(since: } \cosh(2\phi) = 2 \cosh^2 - 1 \text{ and } \sinh^2 \phi = \cosh^2 \phi - 1) \\
(3 - 2\sqrt{2})(\cosh^2 \phi - 1) = 1 - r^2 \cosh^2 \phi \implies \\
\cosh^2 \phi (3 - 2\sqrt{2} + r^2) = 4 - 2\sqrt{2} \implies \\
\cosh^2 \phi = \frac{4 - 2\sqrt{2}}{(3 - 2\sqrt{2})(1 + \frac{m^2}{n^2})} = m^2 \cdot \frac{4 - 2\sqrt{2}}{3 - 2\sqrt{2}} \implies \text{ and finally,} \\
\phi = \pm \arccosh \left( m \sqrt{\frac{4 - 2\sqrt{2}}{3 - 2\sqrt{2}}} \right) = \pm \arccosh \left( m \sqrt{4 + 2\sqrt{2}} \right)
\]

It appears that if \( \phi \) is positive it makes \( t \) negative, which means that we need to select a minus sign in (B.4) as the final result for \( \phi \). This finalizes our formula for the breaking time in (B.2) as

\[
t_{br} = \frac{\phi}{2m} + \frac{\arcsin \left( \cosh(\phi) \cdot \frac{n}{m} \sqrt{3 - \sqrt{2}} \right)}{2n},
\]
where \( \phi = - \arccosh \left( m \sqrt{4 + 2\sqrt{2}} \right) \),

and if we substitute \( \phi \) in \( \cosh \) we get:

\[
t_{br} = \frac{\arcsin \left( n \sqrt{4 - 2\sqrt{2}} \right)}{2n} - \frac{\arccosh \left( m \sqrt{4 + 2\sqrt{2}} \right)}{2m}
\]
Appendix C

Derivation of conserved quantity $E_0$ for the Gaussian IC

We need to calculate

$$E_0 = \int_R u^2 dx = \int_R a^2 (1 - 2bx^2)^2 e^{-2bx^2} dx$$

(C.1)

Based on integration by parts and the fact that the Gaussian integral is

$$\int_R e^{-x^2} dx = \sqrt{\pi}, \text{ or in our case } \int_R e^{-2bx^2} dx = \sqrt{\frac{\pi}{2b}} = \frac{\sqrt{2\pi}}{2\sqrt{b}}$$

(C.2)

we get

$$\int_R u^2 dx = a^2 \int_R (1 - 4bx^2 + 4b^2 x^4) e^{-2bx^2} dx =$$

$$a^2 \left( \int_R e^{-2bx^2} dx - \int_R 4bx^2 e^{-2bx^2} dx + \int_R 4b^2 x^4 e^{-2bx^2} dx \right).$$

(C.3)

Let's integrate the second integral in parentheses by parts:

$$\int_R (-4bx^2) e^{-2bx^2} dx = \int_R x e^{-2bx^2} d(-2bx^2) dx =$$

$$xe^{-2bx^2} \big|_R - \int_R e^{-2bx^2} dx = - \int_R e^{-2bx^2} dx$$

(C.4)

Therefore, based on (C.3), the first two integrals in parentheses cancel each other out and we have
\[
\int_R u^2 \, dx = a^2 \int_R 4b^2 x^4 e^{-2bx^2} \, dx = -ba^2 \int_R (-4bx)x^3 e^{-2bx^2} \, dx
\]
\[
- ba^2 x^3 e^{-2bx^2} \bigg|_R + 3a^2 \int_R bx^2 e^{-2bx^2} \, dx = 3a^2 \int_R bx^2 e^{-2bx^2} \, dx
\]

and based on (C.4) and then (C.2) we obtain

\[
3a^2 \int_R bx^2 e^{-2bx^2} \, dx = \frac{3a^2}{4} \int_R e^{-2bx^2} \, dx = \frac{3a^2 \sqrt{2\pi}}{4 \cdot 2\sqrt{b}} = \frac{3a^2 \sqrt{2\pi}}{8\sqrt{b}}.
\]
Appendix D

Derivation of conserved quantity $E_{-1}$ for the Gaussian IC

We need to show that

$$E_{-1} := \int_R \left[ (\partial_x^{-1}u)^2 - \frac{1}{12}u^4 \right] \, dx = \int_R (\partial_x^{-1}u)^2 \, dx - \frac{1}{12} \int_R u^4 \, dx \quad (D.1)$$

Let's consider the anti-derivative for the Gaussian:

$$\int a(1 - 2bx^2)e^{-bx^2} \, dx = a \int e^{-bx^2} \, dx - a \int 2bx^2e^{-bx^2} \, dx =$$

$$a \int e^{-bx^2} \, dx + a \int xe^{-bx^2}d(-bx^2) =$$

$$a \int e^{-bx^2} \, dx + axe^{-bx^2} - a \int e^{-bx^2} \, dx = axe^{-bx^2} \quad (D.2)$$

The integral of the square of anti-derivative on $R = (-\infty, +\infty)$ is:

$$\int_R a^2x^2e^{-2bx^2} \, dx = \frac{a^2}{4b} \left( -xe^{-2bx^2} \right)_{R} + \int_R e^{-2bx^2} \, dx \quad (D.3)$$

The first term is equal to zero on $R$, the second one is a Gaussian integral and therefore for the Gaussian IC we get

$$\int_R (\partial_x^{-1}u)^2 \, dx = \frac{a^2}{4b} \cdot \frac{\sqrt{2\pi}}{2\sqrt{b}} = \frac{a^2\sqrt{\pi} \cdot 256\sqrt{2}}{2048\sqrt{b^3} \quad (D.4)}$$

Now let's consider the second term of the integral:

$$-\frac{a^4}{12} \int_R (1 - 2bx^2)^4 \, dx = -\frac{a^4}{12} \int_R (1 - 8bx^2 + 24b^2x^4 - 32b^3x^6 + 16b^4x^8) \, dx \quad (D.5)$$
Let’s integrate each term separately. We will integrate each by parts, keeping in mind that $\int_{-\infty}^{\infty} x^n e^{-4bx^2} \, dx = 0$, and that formula for the Gaussian integral (C.2):

$$\int_{-\infty}^{\infty} e^{-4bx^2} \, dx = \frac{\sqrt{\pi}}{2\sqrt{b}} \quad \text{Gaussian integral} \quad (D.6)$$

Second term in (D.5):

$$\int_{-\infty}^{\infty} -8bx^2 e^{-4bx^2} \, dx = \int_{-\infty}^{\infty} xe^{-4bx^2} \, d(-4bx^2) = -\int_{-\infty}^{\infty} e^{-4bx^2} \, dx = -\frac{\sqrt{\pi}}{2\sqrt{b}} \quad (D.7)$$

Third term in (D.5) (same methodology of integration by parts and eliminating the zero term with $\left. x^n e^{-4bx^2} \right|_R$):

$$\int_{-\infty}^{\infty} 24b^2 x^4 e^{-4bx^2} \, dx = 3 \cdot 3 \int_{-\infty}^{\infty} bx^2 e^{-4bx^2} \, dx = \frac{9\sqrt{\pi}}{16\sqrt{b}} \quad (D.8)$$

Fourth term in (D.5):

$$\int_{-\infty}^{\infty} -32b^3 x^6 e^{-4bx^2} \, dx = -5 \int_{-\infty}^{\infty} 4b^2 x^4 e^{-4bx^2} \, dx = -\frac{15}{2} \int_{-\infty}^{\infty} bx^2 e^{-4bx^2} \, dx = -\frac{15\sqrt{\pi}}{32\sqrt{b}} \quad (D.9)$$

The last, fifth term in (D.5):

$$\int_{-\infty}^{\infty} 16b^4 x^8 e^{-4bx^2} \, dx = 7 \int_{-\infty}^{\infty} 2b^3 x^6 e^{-4bx^2} \, dx = \frac{35}{4} \int_{-\infty}^{\infty} b^2 x^4 e^{-4bx^2} \, dx = \frac{105}{32} \int_{-\infty}^{\infty} bx^2 e^{-4bx^2} \, dx = \frac{105\sqrt{\pi}}{512\sqrt{b}} \quad (D.10)$$

Now we can sum up the five terms of the second part of the integral:

$$-\frac{a^4}{12} \cdot \frac{\sqrt{\pi}}{2\sqrt{b}} \left( 1 \right) - 1 + \frac{9}{16} + \frac{15}{32} + \frac{105}{512} = -\frac{a^2}{12} \cdot \frac{\sqrt{\pi}}{2\sqrt{b} \cdot 512} = -\frac{a^2 \sqrt{\pi} \cdot 51b}{2048\sqrt{b^3}} \quad (D.11)$$

Finally, if we combine the results of integration of both terms of $E_{-1}$ we obtain:

$$\frac{a^2 \sqrt{\pi} \cdot 256\sqrt{2}}{2048\sqrt{b^3}} - \frac{a^4 \sqrt{\pi} \cdot 51b}{2048\sqrt{b^3}} = \frac{a^2 \sqrt{\pi} (256\sqrt{2} - 51a^2b)}{2048\sqrt{b^3}} \quad (D.12)$$
Bibliography


