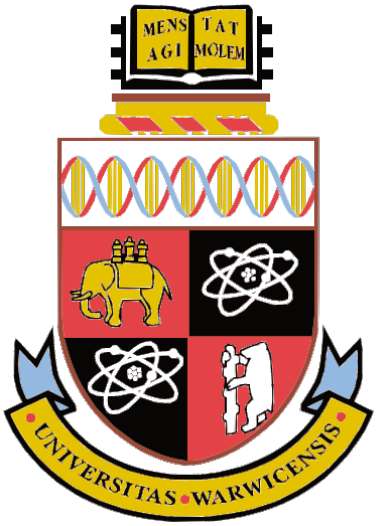


Stochastic Dynamics in Thin Fluid Films

by

Thomas Stopford

MA4K9 Dissertation

Submitted to The University of Warwick

Mathematics Institute

April, 2023

Contents

1 Introduction

The publication of the famous Reynolds equation [1], a tool used to find the pressure distribution in a thin fluid film between two bearings, is thought by many to be the origin of lubrication theory. The field has found applications across a vast range of scientific domains and length scales, relevant to the modelling of microfluidics [11], flows in the anterior chamber of the eye [3], and marine ice sheets [22]. Reynolds' approach was to reduce the Navier-Stokes equations into a much simpler set of governing equations by exploiting the dispartity between characteristic lengthscales in orthogonal directions. His simplified equations still capture the essential characteristics of the full Navier-Stokes equations at these length scales, but allow the boundary value problem to be reduced to a single PDE determining the pressure field.

Lubrication theory exploits this approach to model free surface flows; in particular the dynamics of a thin liquid film between a solid and a gas. Unlike the work of Reynolds, where the height of the fluid is prescribed to calculate the pressure, here the fluid surface is treated as a free boundary and the pressure conditions at the surface are prescribed. An evolution equation for the surface profile may be derived from the same simplified governing equations used by Reynolds. Much work has been done to study these deterministic evolution equations both theoretically and numerically, and mechanisms for instability and rupture of these films are of particular interest.

Disjoining pressure has been explored as a key mechanism for rupture in thin liquid films [24]. However, at the nanoscale, the effects of thermal fluctuations become increasingly relevant. These fluctuations can contribute to instabilities such as the development of capillary waves on the surface of the film [19] and even rupture of the film [15]. Modelling the effects of fluctuations on the film is the primary focus of this report. Their involvement motivates the use of stochastic PDEs (SPDEs). The field of fluctuating hydrodynamics was first conceived with the creation of the Landau-Lifschitz Navier-Stokes equations [18], where a noise term was added to the stress tensor of the fluid for the purpose of modelling turbulence. Grun et al. [13] used formal arguments to derive a corresponding stochastic thin film equation. Since then, work has been done to establish a more rigorous mathematical basis to the thin film equation with the introduction of spatially correlated noise [8].

Theoretical analysis has shown that the stochastic thin film equation captures much of the phenomena that are observed experimentally, e.g. the development of thermal capillary waves at the surface which can be seen after a linear stability analysis [19]. However, as with the deterministic equation, the strong non-linearity of the system makes such theoretical investigations of the full system difficult. Because of this, numerical investi-

gations become a useful tool in analysing these phenomena, and to verify the established theoretical findings.

In this work, we will first introduce the general theory of thin liquid films in Sec. (2). We derive the deterministic evolution equation for the height profile of the film, and briefly investigate its dynamics through linearisation. We then introduce the stochastic thin film equation and perform a similar linear stability analysis.

In Sec. (3), numerical schemes are constructed to solve the deterministic thin film equation. These will use the finite difference method and the Fourier pseudo-spectral method. We show that these schemes converge to known analytical solutions to the linearised system under temporal and spatial grid refinement.

These schemes will be developed to solve the stochastic thin film equation in Sec. (4) and verified by comparison to the theoretical result for the evolution of capillary waves derived in Sec. (2.2.1).

Finally, we conduct a brief investigation into ruptures in thin liquid films caused by thermal fluctuations in Sec. (5). We outline analytical results for the rupture profile and expected time of rupture [15]. Attempts are then made to use our numerical schemes to verify these theoretical results. Initially we use a Monte-Carlo approach, but this becomes computationally infeasible as the noise strength is decreased to zero, which motivates the use of importance splitting [2] and Markov chain approximation [10] to make estimates for properties of these rare events.

2 Theory for Thin Fluid Films

In this section we will introduce the relevant theory for free surface flows. To begin, we derive a thin film equation for a surface tension driven flow in the deterministic setting. We then go on to discuss the stochastic thin film equation and perform a linear stability analysis to derive the evolution of the spectrum of the stochastic thin film equation.

2.1 Deterministic Dynamics

In order to explore the dynamics of a thin film of fluid between a flat solid and vapour, we will consider only two spatial dimensions, x parallel to the solid and y perpendicular to the solid, where $y = 0$ refers to the fluid-solid interface, and $y = h(x, t)$ refers to the the fluid-vapour interface. We begin with a fluid flow problem, the goal being to obtain a single evolution equation for the height, h, of the film. The derivation that follows is heavily based on the derivation in [7, p. 1137-1141].

The fluid flow is governed by the incompressible Navier-Stokes equations. These are given by

$$\nabla \cdot \mathbf{u} = 0, \tag{1}$$

$$\rho \left[\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u}\right] = -\nabla p + \mu \nabla^2 \mathbf{u}, \tag{2}$$

where $\mathbf{u} = (u, v)$ and p are the Eulerian velocity and pressure fields of the fluid respectively, with ρ and μ denoting its density and viscosity. Note that the influence of gravity is considered negligible. This is because surface forces dominate volumetric forces at small scales.

We non-dimensionalise Eq. (1) and Eq. (2) with the hope of simplifying them through asymptotic analysis. First, we introduce characteristic vertical and horizontal length scales H and L so that $y = H\tilde{y}$ and $x = L\tilde{x}$ for dimensionless variables $\tilde{x}$ and $\tilde{y}$. We group terms by factors of $\varepsilon = \frac{H}{L}$ and apply the long wave approximation, $\varepsilon << 1$, to neglect terms of $O(\varepsilon^2)$ or higher. We choose characteristic horizontal velocity scale U, time scale $T = \frac{L}{U}$, and pressure scale $P = \frac{\mu UL}{H^2}$, along with dimensionless variables $\tilde{u}, \tilde{v}, \tilde{p}$ and $\tilde{t}$ so that $u = U\tilde{u}$, $v = \frac{UH}{L}$, $t = T\tilde{t}$, and $p = P\tilde{p}$.

Under these charactersitic scales Eq. (1) and Eq. (2) become

$$\partial_{\tilde{x}} \tilde{u} + \partial_{\tilde{y}} \tilde{v} = 0, \tag{3}$$

$$Re\varepsilon^2 \left[\partial_{\tilde{t}} \tilde{u} + \tilde{u}\partial_{\tilde{x}} \tilde{u} + \tilde{v}\partial_{\tilde{y}} \tilde{u}\right] = -\partial_{\tilde{x}} \tilde{p} + \varepsilon^2 \partial_{\tilde{x}}^2 \tilde{u} + \partial_{\tilde{y}}^2 \tilde{u}, \tag{4}$$

$$\frac{\rho UH}{L} \varepsilon^4 \left[\partial_{\tilde{t}} \tilde{v} + \tilde{u}\partial_{\tilde{x}} \tilde{v} + \tilde{v}\partial_{\tilde{y}} \tilde{v}\right] = -\partial_{\tilde{y}} \tilde{p} + \varepsilon^4 \partial_{\tilde{x}}^2 \tilde{v} + \varepsilon^2 \partial_{\tilde{y}}^2 \tilde{v}. \tag{5}$$

Neglecting terms grouped by factors of ε^2 or higher we obtain our simplified governing

equations

$$\partial_x u + \partial_y v = 0, \tag{6}$$

$$-\partial_x p + \mu \partial_y^2 u = 0, \tag{7}$$

$$-\partial_y p = 0. \tag{8}$$

At the interface between the film and the solid we assume no-slip and no-penetration boundary conditions,

$$u = 0 \text{ and } v = 0 \text{ at } y = 0. \tag{9}$$

On the fluid-vapour interface we have the dynamic boundary condition, obtained by balancing forces due to surface tension and pressure,

$$\mathbf{T} \cdot \hat{\mathbf{n}} = \gamma \kappa \hat{\mathbf{n}} \text{ at } y = h(x,t). \tag{10}$$

Where $\hat{\mathbf{n}}$ is the outward unit normal to the boundary given by $\hat{\mathbf{n}}^\intercal = \frac{1}{\sqrt{1+(\partial_x h)^2}}(-\partial_x h, 1)$, $\mathbf{T}$ is the stress tensor of the fluid given by $T_{ij} = \mu\left[\partial_i u_j + \partial_j u_i\right] - p\delta_{ij}$, γ is the surface tension coefficient of the fluid, and $\kappa = \nabla \cdot \hat{\mathbf{n}}$ is the curvature of the surface. Equating the normal components of the dynamic boundary condition, we see that $\hat{\mathbf{n}} \cdot \mathbf{T} \cdot \hat{\mathbf{n}} = \hat{\mathbf{n}} \cdot \gamma\kappa\hat{\mathbf{n}} = \gamma\kappa$. Ignoring terms that are $O(\varepsilon^2)$ or smaller, we see that this reduces to our boundary condition for the pressure at the film surface

$$-p = \gamma \partial_x^2 h \text{ at } y = h(x,t). \tag{11}$$

Taking the tangential components, we see that $\hat{\mathbf{t}} \cdot \mathbf{T} \cdot \hat{\mathbf{n}} = 0$, where $\hat{\mathbf{t}}$ is the unit tangent to the film surface, we obtain a boundary condition for $\partial_y u$ at the film surface

$$\partial_y u = 0 \text{ at } y = h(x,t). \tag{12}$$

Additionally, enforcing mass conservation, we obtain that

$$\partial_t h = -\partial_x \phi, \tag{13}$$

where ϕ is the flux given by $\phi(x,t) = \int_0^{h(x,t)} u(x,\tilde{y},t)\mathrm{d}\tilde{y}$.

By Eq. (8), the pressure has no y dependence. This allows us to rearrange and integrate Eq. (7) from h to y and apply boundary condition Eq. (12) to obtain $\frac{1}{\mu}(y-h)\partial_x p = \partial_y u$. This may be integrated again from 0 to y, and applying the no slip condition, Eq. (9), we get

$$\frac{1}{\mu}\left[\frac{y^2}{2} - hy\right]\partial_x p = u,$$

and applying our mass conservation condition, Eq. (13), we obtain

$$\partial_t h = \partial_x \int_0^h \frac{1}{\mu}\left[\frac{y^2}{2} - yh\right]\partial_x p \tag{14}$$

$$= \frac{1}{3\mu}\partial_x\left(h^3\partial_x p\right). \tag{15}$$

$$\tag{16}$$

Applying Eq. (11), we obtain the thin film equation

$$\partial_t h = -\frac{\gamma}{3\mu}\partial_x\left(h^3\partial_x^3 h\right). \tag{17}$$

We can choose characteristic length scale H, and time scale $\frac{3H\mu}{\gamma}$ to obtain a dimensionless form of the equation

$$\partial_{\tilde{t}}\tilde{h} = -\partial_{\tilde{x}}\left(\tilde{h}^3\partial_{\tilde{x}}^3\tilde{h}\right). \tag{18}$$

This is the deterministic evolution equation we will be solving numerically in later sections. For brevity, we will often drop the '˜' notation.

2.1.1 Gradient Flow Form

We may write the thin film equation in gradient flow form [7, p. 1141] as

$$\partial_t h = -\partial_x\left(M(h)\partial_x\left(\frac{\delta\mathcal{H}}{\delta h}(h)\right)\right). \tag{19}$$

With mobility M and free energy $\mathcal{H}$ defined by

$$M(h) = h^3, \tag{20}$$

$$\mathcal{H}(h) = \int_\Omega \frac{1}{2}\left(\partial_x h\right)^2 \mathrm{d}x. \tag{21}$$

In particular, this is a Wasserstein gradient flow [21]. This means that at all times the direction of $\partial_t h$ is the direction of maximal descent of $\mathcal{H}$ at h subject to mass conservation as in Eq. (13).

Note that from the perspective of fluid mechanics, the dimensional free energy per unit length of a film with height profile h is given by the product of the surface tension and change in area per unit length of the perturbation. This is given by

$$F(h) = \gamma\int_\Omega\left(\sqrt{1+(\partial_x h)^2} - 1\right)\mathrm{d}x = \gamma\int_\Omega \frac{1}{2}\left(\partial_x h\right)^2\mathrm{d}x + O(\varepsilon^2). \tag{22}$$

From this perspective we see that the thin film equation is simply a gradient descent on the free energy of the film.

2.1.2 Linearisation

Linearising Eq. (18) by writing $h = 1 + \delta h$ where $\delta h \ll 1$ is a perturbation we obtain

$$\partial_t(1+\delta h) = \partial_x\left((1+\delta h)^3\partial_x^3(1+\delta h)\right), \tag{23}$$

so that

$$\partial_t \delta h = \partial_x^4(\delta h) + O(|\delta h|^2). \tag{24}$$

Removing the higher order terms we obtain the linearised thin film equation

$$\partial_t \delta h = -\partial_x^4 \delta h. \tag{25}$$

We intend to find an analytical result for the evolution of the spectrum of δh. Before doing this we give a brief background on Fourier series and the Fourier transform.

2.1.3 Fourier Series

Here we give a summary of the essential facts about Fourier series given in [23, Chapter. 2]. Given a field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, we will let $L^2([0,L],\mathbb{K}) = \left\{f : [0,L] \to \mathbb{K} | \int_0^L |f|^2 \mathrm{d}x < \infty\right\}$ be the Hilbert space of square integrable functions. If the field $\mathbb{K}$ is not specified we take $\mathbb{K} = \mathbb{R}$. We equip this space with the standard L^2 inner product

$$\langle f, g\rangle = \frac{1}{L}\int_0^L f\overline{g}\mathrm{d}x, \tag{26}$$

where $\overline{g}$ denotes the complex conjugate of g. We define the trigonometric basis of $L^2([0,L],\mathbb{C})$ to be the set $\left\{\frac{1}{\sqrt{L}}E_n : n \in \mathbb{Z}\right\}$ where $E_n(x) = e^{iknx}$ with $k = \frac{2\pi}{L}$. This forms an orthonormal basis of $L^2([0,L],\mathbb{C})$, while the set $\left\{\frac{1}{\sqrt{L}}E_0\right\} \cup \left\{\frac{1}{2\sqrt{L}}(E_n + E_{-n}) : n \in \mathbb{N}\right\}$ forms an orthonormal basis of $L^2([0,L],\mathbb{R})$.

By orthonormality, given a function $u \in L^2([0,L],\mathbb{K})$, we may write u as the series

$$u = \frac{1}{L}\sum_{n=-\infty}^{\infty}\langle u, E_n\rangle E_n. \tag{27}$$

We will define $\langle u, E_n\rangle$ to be the n^{th} Fourier coefficient of u and we will denote it as $\widehat{u}_n = \frac{1}{L}\langle u, E_n\rangle = \frac{1}{L}\int_0^L u(x)e^{-inkx}dx$.

We also note the following useful identities. Given $\alpha \in \mathbb{N}$, and $f \in C^\alpha([0,L])$ we may

apply integration by parts α times to obtain

$$\widehat{\partial^\alpha f}_n = \frac{1}{L}\int_0^L \partial^\alpha f e^{-iknx}\mathrm{d}x \tag{28}$$

$$= (-1)^\alpha \frac{1}{L}\int_0^L (-ikn)^\alpha f e^{-iknx}\mathrm{d}x \tag{29}$$

$$= (ikn)^\alpha \widehat{f}_n. \tag{30}$$

2.1.4 The Fourier Transform

Given $f \in L^2(\mathbb{R}, \mathbb{K})$, we may define the Fourier transform of f, $\mathcal{F}(f)$ by

$$\mathcal{F}(f)(\xi) = \int_{-\infty}^{\infty} f(x)e^{i\xi x}\mathrm{d}x. \tag{31}$$

Given $\alpha \in \mathbb{N}$, if $f \in C^\alpha(\mathbb{R}, \mathbb{K}) \cap L^2(\mathbb{R}, \mathbb{K})$ we have the derivative formula

$$\mathcal{F}(\partial_x^\alpha f)(\xi) = (i\xi)^\alpha \mathcal{F}(f)(\xi). \tag{32}$$

Here here we have used integration by parts and the fact that such a function must decay to 0 as $|x| \to \infty$. For convenience, we will use $\widehat{f}$ to denote the Fourier transform of f. We will often refer to $\widehat{f}$ as the spectrum of f.

It is also worth noting that if the support of f lies in $[0, L]$, then we may interpret f as an $L^2([0, 1])$ function, whose Fourier coefficients are given by

$$\widehat{f}_n = \frac{1}{L}\int_0^L f(x)e^{-iknx}\mathrm{d}x = \frac{1}{L}\int_{-\infty}^{\infty} f(x)e^{-iknx}\mathrm{d}x = \frac{1}{L}\widehat{f}(kn). \tag{33}$$

2.1.5 Application to the Linearised System

Now that we have outlined the necessary results relating to Fourier series, we are ready to derive the spectrum of δh. Assume that $h \in C^{4,1}(\mathbb{R} \times [0, \infty))$ so that the derivatives in Eq. (25) make sense, and that for each $t \geq 0$, $h(\cdot, t) \in L^2(\mathbb{R})$. Now we may apply Eq.(32) to Eq. (25) to obtain the evolution of the spectrum of δh,

$$\partial_t \widehat{\delta h}(\xi, t) = -(i\xi)^4 \widehat{\delta h}(\xi, t). \tag{34}$$

Note that here we have used the fact that ∂_t and $\mathcal{F}$ commute. This is a trivial ODE with solution

$$\widehat{\delta h}(\xi, t) = e^{-(i\xi)^4 t}\widehat{\delta h}(\xi, t). \tag{35}$$

If we enforce that h and its spatial derivatives up to order 4 are periodic on $[0, L]$ and consider h only on this domain, then we may apply Eq. (33) to obtain the evolution of it's

Fourier coefficients

$$\widehat{\delta h}_n(t) = e^{-(ikn)^4 t}\widehat{\delta h}_n(0). \tag{36}$$

We may apply this to obtain a simple analytical solution for the initial conditions $h(x,0) = 1 + \alpha\cos(kx)$. This has the form

$$h(x,t) = 1 + \alpha\cos(kx)e^{-k^4 t}. \tag{37}$$

This analytical solution will be used to test our numerical methods in Sec. (3).

2.2 Noise Model

In order to incorporate the effect of thermal fluctuations into our model, an analogous stochastic thin film equation may be derived as before, except it instead begins with the incompressible Landau-Lifschitz Navier-Stokes equations [18], given as follows

$$\nabla \cdot \mathbf{u} = 0, \tag{38}$$

$$\rho\left[\partial_t \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u}\right] = -\nabla p + \mu\nabla^2\mathbf{u} + \nabla\cdot\boldsymbol{\mathcal{S}}. \tag{39}$$

Here $\boldsymbol{\mathcal{S}}$ is the noise tensor representing the effect of thermal fluctuations. The components $\mathcal{S}_{ij}$ of $\boldsymbol{\mathcal{S}}$ are Gaussian processes with mean $E\left[\mathcal{S}_{ij}(\mathbf{x},t)\right] = 0$ and auto-correlation $E\left[\mathcal{S}_{ij}(\mathbf{x},t)\mathcal{S}_{lm}(\mathbf{x}',t')\right] = 2\mu K_B T\delta_{ij}\delta_{lm}\delta(\mathbf{x}-\mathbf{x}')\delta(t-t')$, where K_B is the Boltzmann constant and T is the temperature of the fluid. Note that we use $E[X]$ to denote the expected value of a random variable X.

Since $\boldsymbol{\mathcal{S}}$ is a spatially uncorrelated random variable, it is natural to ask what the mathematical meaning of its divergence is. Rigorously, white noise terms in stochastic PDEs are considered as measures on $\mathbb{R}^k$ $(k = 1, 2, 3)$ [5, p. 267], with stochastic differential equations being interpreted as integral equations. Since the focus of this work is in numerical solutions to stochastic PDEs, it is sufficient to only make sense of noise terms in the discrete setting, this is discussed in Sec. (4).

Treating the noise terms formally and applying a similar strategy to that used in Sec.(2.1), Grun et al. [13] found the analogous evolution equation for the film profile under the influence of noise,

$$\partial_t h = \frac{1}{\mu}\partial_x\left[\frac{h^3}{3}\partial_x p + \int_0^h (h-y)\mathcal{S}_{xy}(y)\mathrm{d}y\right],$$

where $p = -\gamma\frac{\partial^2 h}{\partial x^2}$. When applying the long wave approximation to Eq. (39), it was assumed that the components of $\boldsymbol{\mathcal{S}}$ scale in the same way as the corresponding components in the deterministic stress tensor $\mathbf{T}$. This stochastic PDE still involves integrating the noise over y, and so seeking a solution to it is impractical. In the same work, it was shown that the following much simpler stochastic PDE is identical to that given above in the sense that the time evolution of their distribution functions is equal. This stochastic PDE, which

we will refer to as the stochastic thin film equation (STFE) is given as follows,

$$\partial_t h = \frac{1}{\mu}\partial_x \left[\frac{h^3}{3}\partial_x p + \sqrt{\frac{2K_B T h^3}{3}}\mathcal{N} \right]. \tag{40}$$

Here $\mathcal{N}$ is a Gaussian process with mean $E[\mathcal{N}(x,t)] = 0$ and auto-correlation $E[\mathcal{N}(x,t)\mathcal{N}(x',t')] = \delta(x-x')\delta(t-t')$. Non-dimensionalising with respect to equilibrium film height H, choosing time scale $\frac{3\mu H}{\gamma}$ and length scale H, we obtain the dimensionless form of the stochastic thin film equation,

$$\partial_t h = \partial_x \left[-h^3 \partial_x^3 h + \sqrt{2\psi h^3}\mathcal{N} \right], \tag{41}$$

where $\psi = \frac{2K_B T}{\gamma H}$ is a dimensionless constant. Analogously to the deterministic system, [8] shows that we may recast this into gradient flow form

$$\partial_t h = -\partial_x \left(M(h)\partial_x \left(\frac{\delta \mathcal{H}}{\delta h} \right) + \sqrt{2\psi M(h)}\mathcal{N} \right). \tag{42}$$

The details of the specific choice of the noise term are beyond the scope of this report, but are discussed in [21] and [8]. We only remark that this is a conservative stochastic gradient flow. This means that the evolution consists of a gradient flow as discussed in Sec. (2.1.1), and that the noise term is a conservative field.

2.2.1 Linearisation

It can be shown that the stochastic thin film equation captures the evolution of capillary waves on the surface of a fluid due to thermal fluctuations. The derived result has been shown already in [19] by computing the shock response of the system after linearisation. We provide an alternative derivation making use of known results in stochastic PDEs.

To begin, we linearise Eq. (41). Setting $h = 1 + \delta h$ yields the linearised form of the stochastic thin film equation

$$\partial_t \delta h = -\partial_x^4 \delta h + \sqrt{2\psi}\partial_x \mathcal{N}. \tag{43}$$

We would like to derive an analogous result to Eq. (36) for the evolution of the spectrum of δh. We must assume that the support of $h(\cdot,t)$ lies within $[0,L]$ for each t and treat $h(\cdot,t)$ as a periodic $L^2([0,L])$ function. Note that with the addition of noise, we cannot guarantee the regularity assumptions in Eq. (32). But we will apply the result formally to obtain

$$\partial_t \widehat{\delta h}_n = -(ikn)^4 \widehat{\delta h}_n + \sqrt{2\psi} ikn \widehat{\mathcal{N}}_n, \tag{44}$$

where $\widehat{\mathcal{N}}_n(t)$ is a Gaussian process with mean and correlation

$$E\left[\widehat{\mathcal{N}}_n(t)\right] = \frac{1}{L}\int_0^L E[\mathcal{N}(x,t)]e^{-iknx}\mathrm{d}x = 0, \tag{45}$$

$$E\left[\widehat{\mathcal{N}}_n(t)\overline{\widehat{\mathcal{N}}_n(t')}\right] = \frac{1}{L^2}\int_0^L\int_0^L E[\mathcal{N}(x,t)\mathcal{N}(x',t')]e^{ikn(x-x')}\mathrm{d}x\mathrm{d}x' \tag{46}$$

$$= \frac{\delta(t-t')}{L}. \tag{47}$$

We would like to analyse the long term behaviour of the spectrum of δh. To do this it helps to write this equation in stochastic ODE form,

$$\mathrm{d}\left(\widehat{\delta h}_n\right) = \left(-(ikn)^4\widehat{\delta h}_n\right)\mathrm{d}t + \left(\sqrt{\frac{2\psi}{L}}ikn\right)\mathrm{d}W_n, \tag{48}$$

where W_n is a one dimensional Wiener process. This is a form of the Ornstein-Uhlenbeck process [12] with known analytical solution

$$\widehat{\delta h}_n(t) = e^{-(ikn)^4 t}\left(\widehat{\delta h}_n(0) + \sqrt{\frac{2\psi}{L}}\int_0^t e^{-(ikn)^4 s}\mathrm{d}W_n\right). \tag{49}$$

From this we see that $\widehat{\delta h}_n$ is a Gaussian process with mean and autocorrelation given by

$$E\left[\widehat{\delta h}_n(t)\right] = e^{-(ikn)^4 t}\widehat{\delta h}_n(0), \tag{50}$$

$$E\left[\widehat{\delta h}_n(t)\overline{\widehat{\delta h}_n(t')}\right] = e^{-(ikn)^4(t-t')}\left[\widehat{\delta h}_n(0)^2 + \frac{2\psi(ikn)^2}{L}\int_0^{\min(t,t')} e^{-2(ikn)^4 s}\mathrm{d}s\right]. \tag{51}$$

In particular, for an initially uniform film ($\widehat{\delta h}_n(0) = 0$), we have

$$E\left[\left|\widehat{\delta h}_n(t)\right|^2\right] = \frac{\psi L}{4\pi^2 n^2}\left[1 - e^{-2(kn)^4 t}\right]. \tag{52}$$

Applying Eq. (33) we obtain the evolution of the spectrum of δh. Defining $\widehat{\delta h}_{\mathrm{rms}}(\xi,t) = \sqrt{E\left[\left|\widehat{\delta h}(\xi,t)\right|^2\right]}$, we should expect this to evolve as

$$\widehat{\delta h}_{\mathrm{rms}}(\xi,t) = \sqrt{\frac{\psi L}{\xi^2}\left[1 - e^{-2\xi^4 t}\right]}. \tag{53}$$

This result will be used to verify our numerical methods for solving both the full and linearised stochastic thin film equations in Sec. (4).

3 Numerical Methods for the Deterministic Thin Film Equation

In the first instance, we will formulate the problem of solving the linearised thin film equation, Eq. (25), with periodic boundary conditions. Then schemes will be derived to solve it numerically, using the finite difference method and then the Fourier spectral method. To follow we apply similar techniques to find numerical solutions to the full thin film equation, Eq. (17), this time using the finite difference method and the pseudo-spectral method. The efficacy of each method for the linear and full system will then be discussed.

Both the full thin film equation and its linearised form are fourth order PDEs, so implicit time integration is used improve the stability of the methods for larger timestep sizes. This is simple for the linearised thin film equation since the problem reduces to solving a system of linear equations in each time step. In the full thin film equation the system of equations is nonlinear and so must be solved with the aid of root finding algorithms, in our case the Newton-Kantorovich method [20]. We find that anti-aliasing is crucial when solving nonlinear PDEs using fully implicit pseudo-spectral methods, since without it the Newton-Kantorovich method fails to converge.

3.1 Linear Thin Film Equation

The linearised thin film equation will be solved on a finite interval $\Omega = [0, L]$ with periodic boundary conditions. To allow our numerical schemes to be represented as compactly as possible, and for consistency in notation between the linearised and full system, we define the linear flux operator $\phi(h) = -\partial_x^3 h$ and the linear spatial derivative operator $\mathcal{L}(h) = \partial_x^4 h = \partial_x \phi(h)$. So that we may write Eq. (25) in the form

$$\partial_t h = \mathcal{L}(h). \tag{54}$$

The problem will be solved on temporal domain $I = [0, \infty)$. Now that we have defined the necessary terms, the initial value problem we wish to solve is to find $h \in C^\infty(\Omega \times I)$ such that

$$\begin{cases} \partial_t h(x,t) = \mathcal{L}(h)(x,t) \text{ for all } (x,t) \in \Omega \times I, \\ h(x,0) = f(x) \text{ for all } x \in \Omega, \\ \partial_x^\alpha h(0,t) = \partial_x^\alpha h(L,t) \text{ for all } \alpha \in \{0,\dots,4\}, t \in I. \end{cases} \tag{55}$$

We will assume that the initial condition f is $C^\infty(\Omega)$ and is periodic on Ω in all of its derivatives. With this exact formulation we may derive our numerical schemes.

3.1.1 Finite Difference

For our finite difference method we discretise Ω to the uniform spatial grid $\mathbf{\Omega} = \{x_i = i\Delta x : i \in \Theta_N\}$ with N gridpoints, where $\Theta_N = \{0, \dots, N-1\}$ is the set of grid indexes, and $\Delta x = \frac{L}{N}$ is the grid spacing. We will be using a fixed timestep denoted by Δt. Given $j \in \mathbb{N}$, we denote the time $j\Delta t$ with t_j. For a function $u : \Omega \times I \to \mathbb{R}$, we will denote its discrete counterpart with the sequence $\left(\mathbf{u}^{(j)}\right)_{j=1}^{\infty}$, where for each j, the vector $\mathbf{u}^{(j)} = \left(u_0^{(j)}, \dots, u_{N-1}^{(j)}\right) \in \mathbb{R}^N$ denotes the discrete approximation to $h\left(\cdot, t_j\right)$.

We denote the discrete version of our spatial differential operator $\mathcal{L}$ and flux operator ϕ with the discrete operators $\mathbf{L}$ and $\boldsymbol{\phi} : \mathbb{R}^N \to \mathbb{R}^N$. To find a suitable approximations of these we will approximate $\partial_x^3 \approx \mathbf{D}_x^{3,c}$, where $\mathbf{D}_x^{3,c}$ is the third centred difference operator given by

$$\left(\mathbf{D}_x^{3,c}\mathbf{v}\right)_i = \frac{-v_{i-2} + 2v_{i-1} - 2v_{i+1} + v_{i+2}}{2\Delta x^3}, \tag{56}$$

so that our discrete flux operator takes the form $\boldsymbol{\phi}(\mathbf{h}) = \mathbf{D}_x^{3,c}\mathbf{h}$. We approximate $\partial_x \approx \mathbf{D}_x^c$, where $\mathbf{D}_x^c : \mathbb{R}^N \to \mathbb{R}^N$ is the first central difference operator defined by

$$(\mathbf{D}_x^c\mathbf{v})_i = \frac{v_{i+1} - v_{i-1}}{2\Delta x}, \tag{57}$$

so that our discrete spatial differential operator becomes

$$\mathbf{L}(\mathbf{h}) = \mathbf{D}_x^c\boldsymbol{\phi}(\mathbf{h}). \tag{58}$$

Note that the subscripts are all computed modulo N. For the time derivative we use the Euler approximation $\partial_t \approx \frac{\mathbf{h}^{(j+1)} - \mathbf{h}^{(j)}}{\Delta t}$. Now our goal is to solve the discrete problem of finding a sequence $(\mathbf{h}^{(j)})$ of vectors $\mathbf{h}^{(j)} \in \mathbb{R}^N$ so that

$$\begin{cases} \frac{\mathbf{h}^{(j+1)} - \mathbf{h}^{(j)}}{\Delta t} = \mathbf{L}(\mathbf{h}^{(k)}), \\ \mathbf{h}^{(0)} = \mathbf{f}, \end{cases} \tag{59}$$

where $\mathbf{f}$ is vector representing f sampled on the discrete grid $\mathbf{\Omega}$, and $k \in \{j, j+1\}$ depending on if we choose to use implicit or explicit time integration. For its stability properties, we will be using implicit time integration, so we take $k = j + 1$. This a is system of linear equations which we may write in the form $\mathbf{A}\mathbf{h}^{j+1} = \mathbf{h}^j$, where

$$\mathbf{A} = \mathbf{I} + \Delta t\mathbf{L}. \tag{60}$$

Note that $\mathbf{L}$ refers to the matrix form of our spatial differential operator. This may be solved using simple row reduction algorithms. Expanding Eq. (58), we find that

$$(\mathbf{L}(\mathbf{h}))_i = -\frac{h_{i-3} - 2h_{i-2} - h_{i-1} + 4h_i - h_{i+1} - 2h_{i+2} + h_{i+3}}{4\Delta x^4}, \tag{61}$$

so our matrix has the stencil $\frac{1}{\Delta x^4}[1, -2, -1, 4, -1, -2, 1]$.

3.1.2 Notes on Computational Efficiency

Solving a general linear system of N equations in N unknowns has time complexity $O(N^2)$ using the standard row reduction algorithm. If the matrix A associated with the system is banded, that is $A_{ij} = 0$ whenever $|i - j| > \ell$ for some $\ell \in \mathbb{N}_{>0}$, then solving the system has complexity $O(\ell N)$ instead.

Our matrix is almost banded, except due to our periodic boundary conditions there are non-zero elements close to every corner. This may be fixed by applying a coordinate transform to our gridpoints and operating on the transformed gridpoints instead. This coordinate transform is given by

$$q(i) = \begin{cases} 2i \text{ if } i < \frac{N}{2}, \\ 2(N-i) - 1 \text{ if } i \geq \frac{N}{2}. \end{cases}$$

We then instead solve the system $\tilde{A}\tilde{\mathbf{h}}^{j+1} = \tilde{\mathbf{h}}^j$, where $\tilde{h}_{q(i)} = h_i$. The matrix $\tilde{A}$ is banded with 6 bands above and 6 bands below the diagonal. The new transformed system may be solved more efficiently using banded row reduction algorithms.

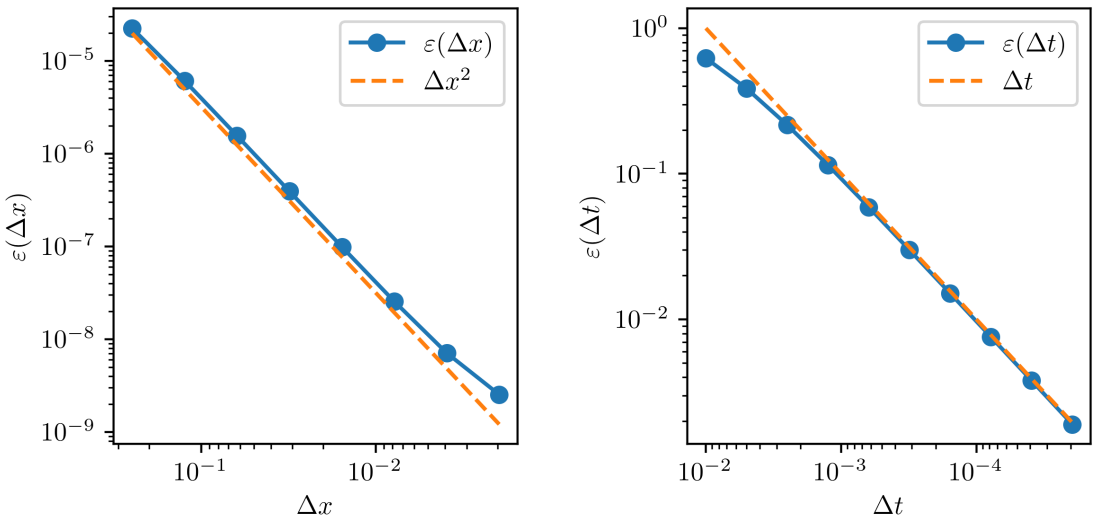
3.1.3 Stability and Convergence

For the linear system we may show stability and convergence analytically. A Von Neumann stability analysis [14, Chapter. 9] shows that the explicit method is stable on the condition that $\frac{\Delta t}{\Delta x^4} < \frac{1}{6}$, while the implicit method is unconditionally stable. This illustrates the need for implicit time integration for this PDE. The size of timestep needed for stability on very fine grids makes explicit methods computationally infeasible. In both methods, a simple Taylor expansion shows that we should expect first order consistency in time and second order consistency in space. This can be seen experimentally in figure (1). For all of our convergence plots we will be using the discrete L^∞ norm. That is, given discrete spatial mesh Ω_N, exact solution $\bar{h}$ and approximate solution $(\mathbf{h}^j)_{j=1}^\infty$, we define the approximation error ε by

$$\varepsilon := \sup_j \|\bar{h}(\cdot, t_j) - \mathbf{h}^j\|_{\infty,N} = \sup_{i,j} |\bar{h}(x_i, t_j) - h_i^j|. \tag{62}$$

3.2 Fourier Spectral Methods

The Fourier spectral method is a popular technique used to solve PDEs. It is favourable to the finite difference method in many cases due to its high order of convergence, and often can be easier to implement.



(a) (b)

Figure 1: Convergence of method (59) to analytical solution (37). (1a) Convergence in space with $\Delta t = 1 \times 10^{-11}$. (1a) Convergence in time with $\Delta x \approx 1 \times 10^{-3}$.

Before we derive the numerical method used for our specific problem, we will discuss the conceptual basis of the Fourier spectral method. Recall the details about Fourier series and the trigonometric basis of $L^2([0, L])$ discussed, Sec. (2.1.3). We now address the discrete counterpart to this, discrete Fourier transforms.

3.2.1 Discrete Fourier Transforms

As with Fourier transforms, the following information is a summary of that seen in [23, Chapter. 2]. Let u be a $C^\infty(\Omega)$ function that is periodic in all of its derivatives, and let $\mathbf{u}$ be its discrete counterpart on the uniform discrete spatial grid $\mathbf{\Omega}$. We define the discrete Fourier transform operator $\mathcal{F}_N : \mathbb{R}^N \to \mathbb{C}^N$ by

$$\mathcal{F}_N(\mathbf{u})_n = \widehat{u}_n, \tag{63}$$

where $\widehat{u}_n$ are the discrete Fourier coefficients of $\mathbf{u}$ given by

$$\widehat{u}_n = \frac{1}{L} \sum_{j=0}^{N-1} u_j e^{-iknx_j} \Delta x. \tag{64}$$

We use the same notation for the discrete Fourier coefficients of $\mathbf{u}$ and the Fourier coefficients of u because they are equal for $n \in \{-N/2, \ldots, N/2 - 1\}$. Given a discretised function $\mathbf{u}$ on a discrete grid with N grid points, we will denote its discrete Fourier trans-

form onto N Fourier coefficients by

$$\widehat{\mathbf{u}} = \mathcal{F}_N(\mathbf{u}) = \left(\widehat{u}_{-N/2}, \ldots, \widehat{u}_{n/2-1}\right). \tag{65}$$

And for convenience we denote the set of reciprocal index points as $\widehat{\Theta}_N = \{-N/2, \ldots, \mathbf{N}/2 - 1\}$. The inverse discrete Fourier transform operator $\mathcal{F}_N^{-1} : \mathbb{C}^N \to \mathbb{R}^N$ is defined by

$$\mathcal{F}_N^{-1}\left(\widehat{\mathbf{u}}\right)_j = \sum_{n=-N/2}^{N/2-1} \widehat{u}_n e^{iknx_j}. \tag{66}$$

Note that $\mathcal{F}_N^{-1}\left(\mathcal{F}_N(\mathbf{u})\right) = \mathbf{u}$ for any $\mathbf{u} \in \mathbb{R}^N$.

The application of discrete Fourier transforms to the Fourier spectral method is explained most effectively by example. It will be applied to the linearised system in the following section.

3.2.2 Application to the Linearised System

The application of the Fourier spectral method to the linearised system is very simple. First recall the evolution formula we obtained for the evolution of the Fourier coefficients, Eq. (34). We simply truncate the Fourier series to the terms whose index lies in $\widehat{\Theta}_N$ and write this compactly as

$$\partial_t \widehat{\mathbf{h}} = (i\boldsymbol{k})^4 \widehat{\mathbf{h}}, \tag{67}$$

where $\boldsymbol{k}$ is the frequency vector given by $k_n = kn$ for $n \in \widehat{\Theta}_N$. We interpret multiplication of vectors as component-wise multiplication. Applying the initial condition $h(x, 0) = f(x)$ for each $x \in [0, L]$ gives us our scheme in discrete space

$$\widehat{h}_n^{(j)} = e^{-(ikn)^4 j\Delta t} \widehat{f}_n \text{ for } n \in \widehat{\Theta}_N. \tag{68}$$

Again we write this in the compact form

$$\widehat{\mathbf{h}}^{(j)} = e^{(i\boldsymbol{k})^4 j\Delta t} \widehat{\mathbf{f}}, \tag{69}$$

with $e^{i\boldsymbol{k}^4 j\Delta t}$ again being a component-wise operation, i.e. $\left(e^{i\boldsymbol{k}^4 j\Delta t}\right)_n = e^{i(nk)^4 j\Delta t}$. At any time t_j, we can obtain our approximation $\mathbf{h}^{(j)}$ with $\mathbf{h}^{(j)} = \mathcal{F}_N^{-1}\left(\widehat{\mathbf{h}}^{(j)}\right)$. This solution has no error in the case where f has only finitely many Fourier coefficients. The only cause of spacial error is in the truncation of the Fourier series of $\widehat{f}$. We see this in Figure (2) where the spatial convergence is super-exponential.

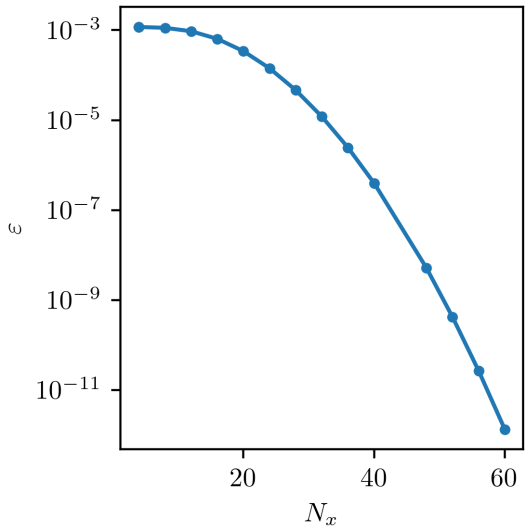


Figure 2: Convergence of method with number of Fourier nodes for the method given by Eq. (69) applied to the initial value problem with $u_0(x) = 1 + e^{-100x^2}$ on domain $\Omega = [-1, 1]$.

.

3.3 Full Thin Film Equation

Setting $\Omega = [0, L]$ and $I = [0, \infty)$ as earlier, the full thin film equation problem is formulated as follows. We seek h such that

$$\begin{cases} \partial_t h(x,t) = -\partial_x \left(h(x,t)^3 \partial_x^3 h(x,t) \right) \text{ for all } (x,t) \in \Omega \times I, \\ h(x,0) = f(x) \text{ for all } x \in \Omega, \\ \partial_x^\alpha h(0,t) = \partial_x^\alpha h(1,t) \text{ for all } \alpha \in \{0, \dots, 4\}, t \in I. \end{cases} \tag{70}$$

Again, we will write our differential equation as $\partial_t h = \mathcal{L}(h)$, where

$$\mathcal{L}(h) = \partial_x \phi(h), \tag{71}$$

with flux operator $\phi(h) = -h^3 \partial_x^3 h$. When solving the full thin film equation, we discretise to a set of non-linear equations which are solved using a root finding algorithm. This is the same principle as that used in [25].

3.3.1 Finite Difference

Using the same notation as in Sec. (3.1.1), we will again set our discrete differential operator to be

$$\mathbf{L}(\mathbf{h}) = \mathbf{D}_x^c \boldsymbol{\phi}(\mathbf{h}). \tag{72}$$

Except this time the discrete flux operator $\boldsymbol{\phi}$ takes the form

$$\phi_i(\mathbf{h}) = (h_i)^3 \left(\mathbf{D}_x^{3,c}\mathbf{h}\right)_i. \tag{73}$$

Our discrete problem has the same form as before. We are searching for a sequence $\left(\mathbf{h}^{(j)}\right) \in \mathbb{R}^N$ such that

$$\begin{cases} \dfrac{\mathbf{h}^{(j+1)} - \mathbf{h}^{(j)}}{\Delta t} = \mathbf{L}\left(\mathbf{h}^{(j+1)}\right), & (74) \\ \mathbf{h}^0 = \mathbf{f}, & (75) \end{cases}$$

except this is now a nonlinear system due to $\mathbf{L}$ now being a nonlinear operator.

We solve this nonlinear system in the following way. Given $\mathbf{h}^{(j)}$, define the residual function $\mathbf{R} : \mathbb{R}^N \to \mathbb{R}^N$ by

$$\mathbf{R}(\mathbf{v}) = \mathbf{v} - \mathbf{h}^{(j)} - \Delta t\mathbf{L}(\mathbf{v}). \tag{76}$$

Observe that if $\mathbf{h}^{(j+1)}$ satisfies Eq. (74), then $\mathbf{R}\left(\mathbf{h}^{(j+1)}\right) = \mathbf{0}$. So evolving our system corresponds to finding a root of $\mathbf{R}$. This is done using the Newton-Kantorovich method [20]. We explain this briefly below.

Given an initial guess $\mathbf{v}_0$, iterate with $\mathbf{v}_{n+1} = \mathbf{v}_n - \left(\mathbf{DR}^{-1}(\mathbf{v}_n)\right)\mathbf{R}(\mathbf{v}_n)$, where $\mathbf{DR}$ is the Jacobean of $\mathbf{R}$. If $\mathbf{v}_0$ is sufficiently close to a root $\mathbf{v}^\star$ of $\mathbf{R}$ we should expect quadratic convergence to the root in almost all cases. A good initial choice for $\mathbf{v}_0$ is the explicit step $\mathbf{v}_0 = \mathbf{h}^{(j)} + \Delta t\mathbf{L}(\mathbf{h}^{(j)})$.

The Jacobean of $\mathbf{R}$ in Eq. (76) is given by $\mathbf{DR} = \mathbf{I} - \Delta t(\mathbf{DL})$, where

$$\begin{aligned} (\mathbf{DL}(\mathbf{h}))_{ij} &= \partial_j L_i(\mathbf{h}) & (77) \\ &= \frac{-\partial_j\phi_{i-1}(\mathbf{h}) + \partial_j\phi_{i+1}(\mathbf{h})}{2\Delta x}, & (78) \end{aligned}$$

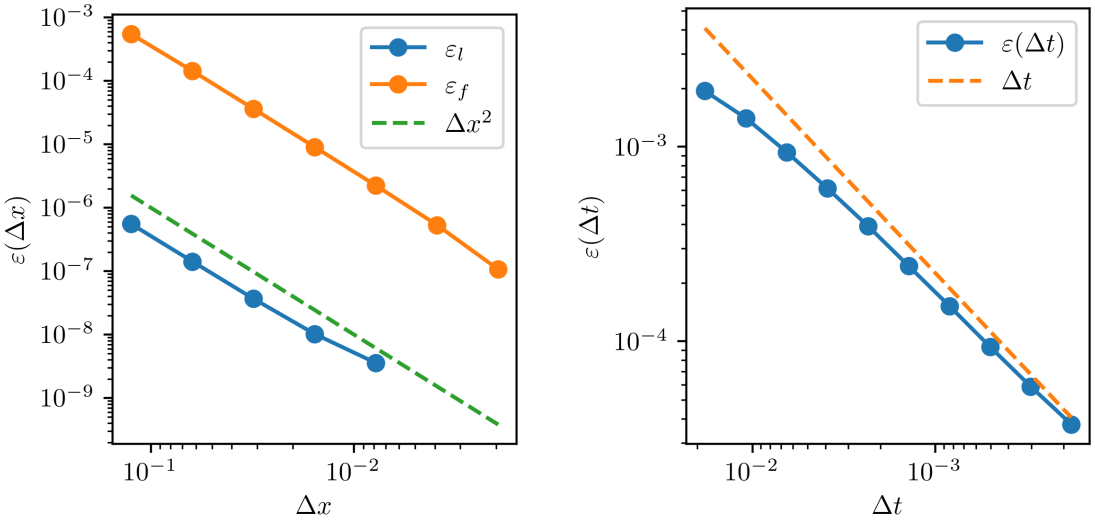
with

$$\partial_j\phi_i(\mathbf{h}) = \begin{cases} -h_i^3/\left[2\Delta x^3\right] \text{ if } j = i-2, \\ 2h_i^3/\left[2\Delta x^3\right] \text{ if } j = i-1, \\ 3h_i^2\left[-h_{i-2} + 2h_{i-1} - 2h_{i+1} + h_{i+2}\right]/[2\Delta x^3] \text{ if } j = i, \\ -2h_i^3/\left[2\Delta x^3\right] \text{ if } j = i+1, \\ h_i^3/\left[2\Delta x^3\right] \text{ if } j = i+2, \\ 0 \text{ otherwise.} \end{cases} \tag{79}$$

Note again that all taken modulo N. Notice that $\mathbf{DR}$ is almost banded in the same sense as described in Sec. (3.1.2). So we may use the coordinate transform mentioned earlier to solve the system more efficiently.

3.3.2 Stability and Convergence

A stability criterion for the explicit method on the nonlinear system is difficult to find analytically, as are results for consistency. However we may find experimental convergence results shown in Figure (3). As we would expect from the central difference approximation for spatial derivatives and Euler time integration scheme, we see quadratic convergence in space and linear convergence in time.



(a) (b)

Figure 3: Convergence of method (74) . (1a) Convergence in space with $\Delta t = 1 \times 10^{-7}$ where ε_l gives convergence to (37) with $\alpha = 1 \times 10^{-4}$ and ε_f giving convergence to the computational solution on the finest grid with $\alpha = 1 \times 10^{-1}$ and . (1a) Convergence in time to (37) $\alpha = 1 \times 10^{-2}$ and with $\Delta x = 1 \times 10^{-3}$.

3.3.3 Fourier Pseudo-spectral

Applying the pseudo-spectral method to the full system is much more complex than with the linear system. We begin by finding an expression for the evolution of the Fourier coefficients of h. Applying Eq. (32) and Eq. (32) to the PDE in Eq. (70) we obtain

$$\partial_t \widehat{h}_n = ikn\widehat{\phi(h)}_n \text{ for each } n \in \mathbb{Z}. \tag{80}$$

We would like to end up with a numerical scheme of the form $\mathbf{D}_t\widehat{\mathbf{h}} = \mathbf{L}\left(\widehat{\mathbf{h}}\right)$, so the problem becomes one of finding an expression for $\widehat{\phi(h)}_n$ in terms of the Fourier coefficients of h. Computationally we only ever use finitely many Fourier coefficients of h, so we immediately make the approximation $h \approx h_N = \sum_{n=-N/2}^{N/2-1} \widehat{h}_n E_n$, i.e. approximating h as

its truncated Fourier series. We begin by writing out the Fourier expansion of $\phi\left(h_N\right)$,

$$\phi\left(h_N\right)(x) = -\left(\sum_{n=-N/2}^{N/2-1} \widehat{h}_n e^{iknx}\right)^3 \left(\sum_{n=-N/2}^{N/2-1} (ikn)^3 \widehat{h}_n e^{iknx}\right) \tag{81}$$

$$= -\sum_{a,b,c,d\in\widehat{\Theta}_N} \widehat{h}_a\widehat{h}_b\widehat{h}_c(ikd)^3\widehat{h}_d e^{ik(a+b+c+d)x}. \tag{82}$$

By orthogonality of the Fourier basis, we see that the Fourier coefficients of $\phi\left(h_N\right)$ have the form

$$\widehat{\phi\left(h_N\right)}_n = \sum_{(a,b,c,d)\in\Lambda_n} \widehat{h}_a\widehat{h}_b\widehat{h}_c(ikd)^3\widehat{h}_d, \tag{83}$$

where $\Lambda_n = \left\{(a,b,c,d) \in (\Theta_N)^4 : a+b+c+d = n\right\}$. However, computing this sum is computationally undesirable since it is guaranteed to require $O(N^3)$ operations. Instead of this approach, we opt to compute the product $h^3\partial_x^3 h$ in physical space and then take its Fourier transform to obtain $\widehat{\phi(h)}$. This departure from doing all of our calculations in Fourier space is known as the pseudo-spectral method [9].

To perform our calculation in physical space, we must choose a number of collocation points to use, M. To maximise computational efficiency, we must choose the smallest possible M such that our computed Fourier coefficients of $\phi(h_N)$ are equal to the true values given by Eq. (83). Using M collocation points we will let $x_j = j\Delta x$ denote the x coordinate of each collocation point where $j \in \Theta_M$ and $\Delta x = L/M$. Taking the discrete Fourier transform of $\phi(h_N)$ with M collocation points gives

$$\mathcal{F}_M(\phi(h_N))_n = -\sum_{j\in\Theta_M} \phi(h)(x_j)e^{-iknx_j}\Delta x \tag{84}$$

$$= -\sum_{j\in\Theta_M}\sum_{a,b,c,d\in\widehat{\Theta}_N} \widehat{h}_a\widehat{h}_b\widehat{h}_c(ikd)^3\widehat{h}_d e^{ik(a+b+c+d)x_j}e^{-iknx_j}\Delta x \tag{85}$$

$$= -\sum_{a,b,c,d\in\widehat{\Theta}_N} \widehat{h}_a\widehat{h}_b\widehat{h}_c(ikd)^3\widehat{h}_d \sum_{j\in\Theta_M} e^{ik(a+b+c+d-n)x_j}\Delta x. \tag{86}$$

Note that that the expression

$$\sum_{j\in\Theta_M} e^{ik(a+b+c+d-n)x_j}\Delta x = \begin{cases} L \text{ if } (a+b+c+d-n) \equiv 0 \mod M, \\ 0 \text{ otherwise.} \end{cases} \tag{87}$$

So we can see that we avoid aliasing error for the n^{th} Fourier coefficient when all elements $(a,b,c,d) \in \left(\widehat{\Theta}_N\right)^4$ satisfy $|a+b+c+d-n| < M$. Noting that a, b, c, d and n all take values between $-N/2$ and $N/2-1$, we deduce that there will be no aliasing error as long as $5N/2 - 1 < M$. This is an extension to the famous '3/2 rule' [4, Chapter 3] [9]. It is not hard to generalise this to a '$(k+1)/2$ rule' using an identical method.

This gives us an efficient method for computing $\widehat{\phi(h_N)}$ in terms of the Fourier coefficients of h_N. We will write this in a more compact form. Letting $\widehat{\mathbf{h}}$ denote the vector of the Fourier coefficients of h_N whose index lies in $\widehat{\Theta}_N$, and $\widehat{\boldsymbol{\phi}} : \mathbb{C}^N \to \mathbb{C}^N$ denote the discrete reciprocal flux operator which takes the values $\widehat{\boldsymbol{\phi}}\left(\widehat{\mathbf{h}}\right)_n = \widehat{\phi\left(h_N\right)}_n$, we may write

$$\widehat{\boldsymbol{\phi}}\left(\widehat{\mathbf{h}}\right) = -\mathcal{F}_{\frac{5N}{2}}\left(-\mathcal{F}^{-1}_{\frac{5N}{2}}\left(\widehat{\mathbf{h}}\right)^3 \mathcal{F}^{-1}_{\frac{5N}{2}}\left(\boldsymbol{k}^3\widehat{\mathbf{h}}\right)\right), \tag{88}$$

where as usual we interpret multiplication of vectors as component-wise multiplication. Note that the right hand side is a $\mathbb{C}^{\frac{5N}{2}}$ vector while the left hand side is a $\mathbb{C}^N$ vector. In fact, we truncate the right hand side to the terms whose index lies in $\widehat{\Theta}_N$ before assigning it to the right hand side. We choose this implicit truncation over the use of a truncation operator for simplicity only.

We now have the discrete problem of finding $\mathbf{h}^{(j)}$ so that

$$\begin{cases} \frac{\widehat{\mathbf{h}}^{(j+1)} - \widehat{\mathbf{h}}^{(j)}}{\Delta t} = \mathbf{L}\left(\widehat{\mathbf{h}}^{(j+1)}\right), \\ \widehat{\mathbf{h}}^{(0)} = \widehat{\mathbf{f}}, \end{cases} \tag{89}$$

where

$$\mathbf{L}\left(\widehat{\mathbf{h}}\right) = i\boldsymbol{k}\widehat{\boldsymbol{\phi}}\left(\widehat{\mathbf{h}}\right). \tag{90}$$

As before, we solve this using the Newton-Kantorovich method. We define the residual function $\mathbf{R}(\mathbf{v}) = \mathbf{v} - \widehat{\mathbf{h}}^{(j)} - \Delta t\mathbf{L}(\mathbf{v})$. The Jacobean of $\mathbf{R}$ is given by

$$\mathbf{DR}(\mathbf{v}) = \mathbf{I} - \Delta t\mathbf{DL}(\mathbf{v}), \tag{91}$$

where $\mathbf{DL}(\mathbf{v}) = i\boldsymbol{k}\mathbf{D}\widehat{\boldsymbol{\phi}}(\mathbf{v})$. To find the Jacobean of $\widehat{\boldsymbol{\phi}}$ recall that

$$\widehat{\boldsymbol{\phi}}\left(\widehat{\mathbf{h}}\right)_m = \sum_{(a,b,c,d)\in\Lambda_m} \widehat{h}_a\widehat{h}_b\widehat{h}_c(ikd)^3\widehat{h}_d. \tag{92}$$

Finding the n^{th} partial derivative of $\widehat{\boldsymbol{\phi}}\left(\widehat{\mathbf{h}}\right)_m$ yields

$$\left(\mathbf{D}\widehat{\boldsymbol{\phi}}\left(\widehat{\mathbf{h}}\right)\right)_{nm} = \partial_{nm}\widehat{\boldsymbol{\phi}}(\widehat{\mathbf{h}}) = (ikn)^3 \sum_{(a,b,c)\in\Gamma_{m-n}} \widehat{h}_a\widehat{h}_b\widehat{h}_c + 3\sum_{(a,b,c)\in\Gamma_{m-n}} \widehat{h}_a\widehat{h}_b(ikc)^3\widehat{h}_c, \tag{93}$$

where $\Gamma_d = \left\{(a,b,c) \in \left(\widehat{\Theta}_N\right)^3 : a+b+c = d\right\}$. Computing each of these terms naively takes $O(N^2)$ time for each coefficient. Instead of doing this, we may apply a similar

method to that used to obtain Eq. (88) to see that

$$\sum_{(a,b,c)\in\Gamma_{m-n}} \widehat{h}_a\widehat{h}_b\widehat{h}_c = \mathcal{F}_{3N}\left(\mathcal{F}_{3N}^{-1}\left(\widehat{\mathbf{h}}\right)^3\right)_{m-n}, \tag{94}$$

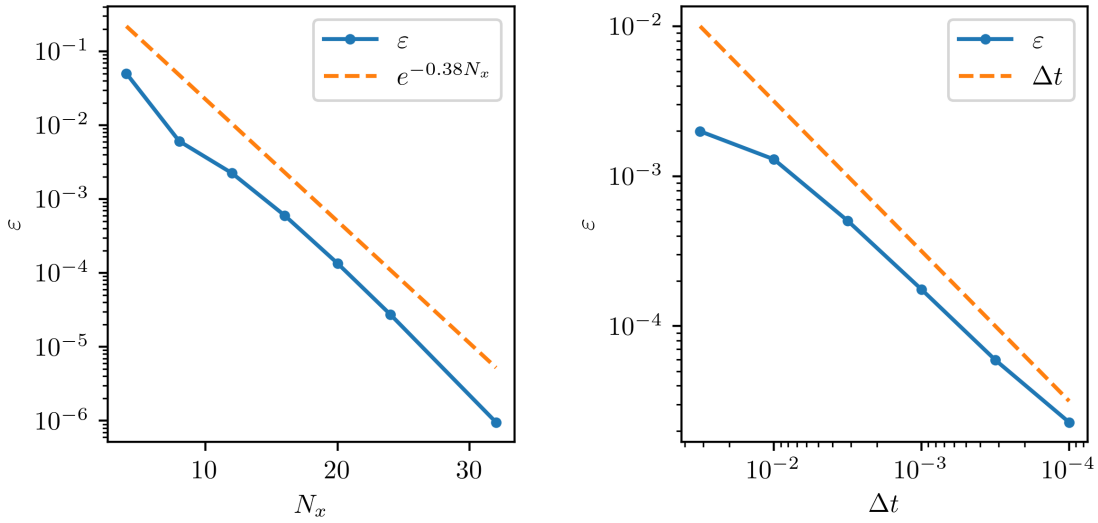
and that

$$\sum_{(a,b,c)\in\Gamma_{m-n}} \widehat{h}_a\widehat{h}_b(ikc)^3\widehat{h}_c = \mathcal{F}_{3N}\left(\mathcal{F}_{3N}^{-1}\left(\widehat{\mathbf{h}}\right)^2\mathcal{F}_{3N}^{-1}\left((i\boldsymbol{k})^3\widehat{\mathbf{h}}\right)\right)_{m-n}. \tag{95}$$

These identities are useful because we may exploit the 'Fast Fourier Transform' to compute them. Unfortunately the Jacobean of this residual function is a dense matrix, so we cannot exploit the same coordinate transform as with the finite difference method. This means that the time complexity of each step is $O(N^2)$ rather than $O(N\log(N))$ as would be the case in an explicit pseudo-spectral method.

3.3.4 Convergence

In Figure (4) we see that the error of the numerical scheme decreases exponentially with the number of gridpoints used, and linearly with decreasing timestep size.



(a) (b)

Figure 4: Convergence of method (89) with (a) number of Fourier nodes applied to the initial value problem with $u_0(x) = 1 + \frac{7}{10}\cos(\pi x)$ on domain $\Omega = [-1, 1]$ and constant $\Delta t = 1 \times 10^{-7}$. Convergence is measured experimentally against a solution using 256 Fourier nodes. (b) Convergence in timestep size to linearised solution in Eqn. (37).

3.4 Summary

We have now implemented numerical methods to solve the linearised and full thin film equations. The spectral method is clearly superior for solving the linearised thin film equation. Our finite difference scheme for the linearised equation attains second order convergence in space and first order convergence in time, while our spectral scheme is essentially analytic, with no error beyond that arising from the truncation of the Fourier series. Additionally, the spectral scheme is much easier to implement than the finite difference scheme, the independence of each of the Fourier coefficients means that an implicit or analytical scheme can be found by simple algebraic rearrangement or finding the solution to a simple ODE. When using the finite difference method, the evolution matrix A must be constructed, a process which is vulnerable to error, adding to the time taken for implementation. We note that the time complexity of the finite difference scheme is linear in the number of gridpoints used, while the time complexity of the spectral method is only $O(N \log(N))$. However, this is counteracted by the remarkable convergence rate of the spectral scheme, which requires significantly fewer grid points for the same level of accuracy.

In solving the full thin film equation, the pseudo-spectral method is only slightly superior to the finite difference method. One could argue that the finite difference method is more practical, since deriving expressions for its residual function and Jacobean, and then implementing the code to solve the nonlinear system may be done with relative ease. The implementation of the pseudo-spectral scheme is a much more involved process due to the need for anit-aliasing and the fact that the Jacobean is inherently complicated. It is also worth noting that if one wishes to use an interpreted language such as Python to implement the pseudo-spectral scheme, the use of vectorised operations when filling our the dense Jacobean matrix is crucial for performance. This is not such an issue for the finite difference scheme since the Jacobean is sparse. The reason we conclude that the pseudo-spectal method is better than the finite difference method is due to its exponential spatial convergence, in comparison to only quadratic convergence of the finite difference method. As with the linear case, this convergence rate leads to the pseudo-spectral method having better performance overall. The quadratic computational complexity in the number of gridpoints is compensated for by its rate of convergence.

Now that we have confidence in the efficacy of these methods for approximating solutions to the deterministic thin film equation, we are ready to incorporate the effects of thermal noise into our computational models. This will be done in the following section in which we find computational solutions to the stochastic thin film equation.

4 Numerical Methods for the Stochastic Thin Film Equation

The numerical schemes derived in Sec. (3) to solve the deterministic thin film equation will now be adapted to solve the stochastic thin film equation. In the linear case, the noise is additive, so adding the noise before our implicit step turns out to be sufficient. Although when using the Fourier spectral method, a better scheme that exploits the analytical solution Eq. (52) may be used as well.

In the full STFE, the noise is multiplicative. When using the finite difference method, this may be treated implicitly by incorporating the noise into the residual function. For the Fourier spectral method, the aliasing error caused by the fractional power of h is more difficult to remedy than that for quartic non-linearities seen in the deterministic equation. For this reason, the noise term is treated explicitly instead.

4.1 Linear system

We will solve the linear system using finite difference and Fourier spectral methods. The stochastic PDE problem is given as follows. We wish to find h such that

$$\begin{cases} \partial_t h(x,t) = \mathcal{L}(h)(x,t) + \sqrt{2\psi}\sigma(\mathcal{N})(x,t) \text{ for all } (x,t) \in \Omega \times I, \\ h(x,0) = f(x) \text{ for all } x \in \Omega, \\ \partial_x^\alpha h(0,t) = \partial_x^\alpha h(1,t) \text{ for all } \alpha \in \{0,\dots,4\}, t \in I, \end{cases} \tag{96}$$

where $\mathcal{L}$ is as defined in Eq. (55) and σ is the linear noise operator defined by $\sigma(\mathcal{N})(x,t) = \partial_x \mathcal{N}(x,t)$.

The noise may be treated either explicitly, by simply adding it after the implicit integration step, or implicitly, by adding it before the implicit integration. Since the noise is additive, this choice does not effect the convergence or stability properties of the method, and both should be equivalent in the limit $\Delta t \to 0$. In this section we will explore methods that use each of these approaches.

4.1.1 Finite Difference

Before we outline the finite difference scheme we will address the discretisation of the noise term, $\mathcal{N}$. On the uniform mesh $\mathbf{\Omega}$, we approximate the Dirac-delta function with the scaled Kronecker delta, that is $\delta(x_i - x_j) \approx \frac{1}{\Delta x}\delta_{ij}$. Similarly, in the discrete time grid, we use the approximation $\delta(t_m - t_n) \approx \frac{1}{\Delta t}\delta_{mn}$. So at time t_j, we approximate the noise

term $\mathcal{N}(\cdot, t_j)$ with the Gaussian vector $\mathbf{N}^{(j)} \in \mathbb{R}^N$ with mean and auto-correlation

$$E\left[N_i^{(j)}\right] = 0, \tag{97}$$

$$E\left[N_i^{(j)} N_{i'}^{(j')}\right] = \frac{1}{\Delta x \Delta t}\delta_{ii'}\delta_{jj'}, \tag{98}$$

for each $i, i', j, j' \in \Theta_N$.

To approximate the noise operator, σ, we use a central difference to approximate the spatial derivative. This gives us the discrete noise operator, $\boldsymbol{\sigma} : \mathbb{R}^N \to \mathbb{R}^N$ defined by $\boldsymbol{\sigma}(\mathbf{N}) = \mathbf{D}_x^c \mathbf{N}$. Writing out our system with these finite difference operators gives us the discrete problem, to find a sequence $\left(\mathbf{h}^{(j)}\right)_{j=1}^{\infty}$ such that

$$\begin{cases} \frac{\mathbf{h}^{(j+1)} - \mathbf{h}^{(j)}}{\Delta t} = \mathbf{L}(\mathbf{h}^{(j)}) + \sqrt{2\psi}\boldsymbol{\sigma}(\mathbf{N}^{(j)}), \\ \mathbf{h}^{(0)} = \mathbf{f}, \end{cases} \tag{99}$$

where $\mathbf{L}$ is defined as in Eq. (58).

This reduces to the linear system of equations given by

$$\mathbf{A}\mathbf{h}^{(j+1)} = \mathbf{h}^{(j)} + \Delta t\sqrt{2\psi}\boldsymbol{\sigma}(\mathbf{N}^{(j)}), \tag{100}$$

with $\mathbf{A}$ defined as in Eq. (60). As before, this system may be solved using row reduction.

4.1.2 Fourier Spectral

Recall that in the deterministic case, the Fourier spectral method was tantamount to evolving the system of ODEs given by Eq. (34) and applying Eq. (33). The analogous procedure in the stochastic setting is to evolve the system of ODEs in Eq. (44). We may discretise this system of ODEs as follows,

$$\begin{cases} \frac{\widehat{\mathbf{h}}^{(j+1)} - \widehat{\mathbf{h}}^{(j)}}{\Delta t} = -(i\boldsymbol{k})^4 \widehat{\mathbf{h}}_n^{(j+1)} + i\boldsymbol{k}\sqrt{2\psi}\widehat{\mathbf{N}}^{(j)}, \\ \widehat{\mathbf{h}}_n^{(0)} = \widehat{\mathbf{f}}. \end{cases} \tag{101}$$

Here $\widehat{\mathbf{N}}$ denotes the discrete Fourier transform of the noise vector, $\mathcal{F}_N(\mathbf{N})$.

One might presume it is better to make use of the analytical solution to the deterministic system, Eq. (36). Using an operator splitting approach, we can solve the deterministic part of the evolution analytically and add the noise explicitly at the end. This would give us the scheme

$$\begin{cases} \widehat{\mathbf{h}}^{(j+\frac{1}{2})} = e^{-(i\boldsymbol{k})^4 \Delta t}\widehat{\mathbf{h}}^{(j)}, \\ \widehat{\mathbf{h}}^{(j+1)} = \widehat{\mathbf{h}}^{(j+1/2)} + i\boldsymbol{k}\Delta t\sqrt{2\psi}\widehat{\mathbf{N}}^{(j)}, \\ \widehat{\mathbf{h}}_n^{(0)} = \widehat{\mathbf{f}}. \end{cases} \tag{102}$$

However, as will be discussed in Sec. (4.3) the efficacy with which this method captures the spectrum is no better than our first Fourier method, Eq. (101). This is because the analytical integration step fails to capture the effects of the noise between times t and $t + \Delta t$. The most efficacious method exploits the exact solution to the linearised STFE in Eq. (50) and Eq. (51). If we know the value of $\widehat{\delta h}(t)$, then the value of $\widehat{\delta h}(t + \Delta t)$ is a random variable with mean $e^{-(ikn)^4 t}\widehat{\delta h}(t)$ and variance $\frac{\psi}{(ikn)^2}\left[e^{-2(ikn)^4\Delta t} - 1\right]$. Explicitly we write our method as

$$\begin{cases} \widehat{\mathbf{h}}^{(j+1)} = e^{-(i\boldsymbol{k})^4\Delta t}\widehat{\mathbf{h}}^{(j)} + \sqrt{\frac{\psi}{(i\boldsymbol{k})^2}\left(1 - e^{-2(i\boldsymbol{k})^4\Delta t}\right)}\mathbf{N}^{(j)}, \\ \widehat{\mathbf{h}}^{(0)} = \widehat{\mathbf{f}}, \end{cases} \tag{103}$$

where in this case, $\mathbf{N}$ is a Gaussian vector with mean 0 and correlation $E\left[N_n^{(j)} N_{n'}^{(j')}\right] = \frac{\delta_{ii'}\delta_{jj'}}{\Delta x}$. Here, all of the operations such as raising to powers, exponentiation and taking the square root are taken to be component-wise operations. This method is used in [15] to explore ruptures in thin films governed by a constant mobility thin film equation.

4.2 Nonlinear System

We formulate the full stochastic thin film problem as

$$\begin{cases} \partial_t h(x,t) = \mathcal{L}(h)(x,t) + \sqrt{2\psi}\sigma(h,\mathcal{N})(x,t) \text{ for all } (x,t) \in \Omega \times I, \\ h(x,0) = h_0(x) \text{ for all } x \in \Omega, \\ \partial_x^\alpha h(0,t) = \partial_x^\alpha h(1,t) \text{ for all } \alpha \in \{0,\ldots,4\}, t \in I, \end{cases} \tag{104}$$

where σ is our noise operator defined by $\sigma(h,\mathcal{N})(x,t) = \partial_x\left(h(x,t)^{\frac{3}{2}}\mathcal{N}(x,t)\right)$, and $\mathcal{L}$ is defined in Eq. (71).

4.2.1 Finite Difference

The following method is a special case of the adaptive grid scheme developed in [25]. As in the deterministic case, we will only be using a uniform spatial grid.

We approximate our noise operator with the discrete noise operator $\boldsymbol{\sigma}$ defined by

$$\boldsymbol{\sigma}(\mathbf{h},\mathbf{N}) = \mathbf{D}_x^c\mathbf{F}(\mathbf{h},\mathbf{N}), \tag{105}$$

where $(\mathbf{F}(\mathbf{h},\mathbf{N}))_i = h_i^{\frac{3}{2}}N_i$. Discretising we get the following nonlinear system of equations

$$\begin{cases} \frac{\mathbf{h}^{(j+1)} - \mathbf{h}^{(j)}}{\Delta t} = \mathbf{L}(\mathbf{h}^{(j+1)}) + \sqrt{2\psi}\boldsymbol{\sigma}\left(\mathbf{h}^{(j+1)}, \mathbf{N}^{(j+1)}\right), \\ \mathbf{h}^{(0)} = \mathbf{f}, \end{cases} \tag{106}$$

where $\mathbf{L}$ is defined in Eq. (72). We intend to use the Newton-Kantorovich method to

find an appropriate $\mathbf{h}^{(j+1)}$ given $\mathbf{h}^{(j)}$. To do this we require a residual function and its Jacobean. For this system, the residual function is given by

$$\mathbf{R}(\mathbf{v}) = \mathbf{v} - \widehat{\mathbf{h}}^{(j)} - \Delta t \left(\mathbf{L}(\mathbf{v}) + \sqrt{2\psi}\boldsymbol{\sigma}\left(\mathbf{v}, \mathbf{N}^{(j+1)}\right)\right). \tag{107}$$

The Jacobean of this residual function is given by $\mathbf{DR} = \mathbf{I} - \Delta t \left(\mathbf{DL} + \sqrt{2\psi}\mathbf{D}\boldsymbol{\sigma}\right)$, where $\mathbf{L}$ is given in Eq. (77).

Note that the Jacobean $\mathbf{D}\boldsymbol{\sigma}$ refers only to the Jacobean with respect to $\mathbf{h}$. This is given by

$$\partial_j \sigma_i(\mathbf{h}, \hat{\mathbf{n}}) = \begin{cases} -3h_j^{\frac{1}{2}} N_j/[4\Delta x] \text{ if } j = i-1, \\ 3h_j^{\frac{1}{2}} N_j/[4\Delta x] \text{ if } j = i+1, \\ 0 \text{ otherwise.} \end{cases} \tag{108}$$

4.2.2 Fourier Pseudo-spectral

If we wish to apply the Fourier pseudo-spectral method to the problem in Eq. (104) we encounter an additional difficulty. Due to the $h^{\frac{3}{2}}$ term, we will always encounter aliasing error. To the author's knowledge, there is no efficient method for anti-aliasing when raising to non-integer powers. This also means that solving the PDE fully implicitly is more difficult, since writing down the Jacobean of our residual function becomes a lot more difficult when we cannot easily find the exact expressions for the Fourier coefficients of $h^{\frac{3}{2}}$. Instead of doing this we invoke operator splitting. We solve the deterministic part of the equation implicitly, and add the noise explicitly at the end. Our discrete noise operator, $\boldsymbol{\sigma}$ takes the form

$$\boldsymbol{\sigma}\left(\widehat{\mathbf{h}}, \widehat{\mathbf{N}}\right) = i\boldsymbol{k}\mathcal{F}_{\frac{5N}{2}}\left(\mathcal{F}^{-1}_{\frac{5N}{2}}\left(\widehat{\mathbf{h}}\right)^{\frac{3}{2}}\mathcal{F}^{-1}_{\frac{5N}{2}}\left(\widehat{\mathbf{N}}\right)\right). \tag{109}$$

Note that the choice to use $\frac{5N}{2}$ collocation points in physical space is somewhat arbitrary. It is chosen to be the largest possible number of collocation points that does not exceed the amount used in the implicit time integration step. This gives us the scheme

$$\begin{cases} \frac{\widehat{\mathbf{h}}^{(j+1/2)} - \widehat{\mathbf{h}}^{(j)}}{\Delta t} = \mathbf{L}\left(\widehat{\mathbf{h}}^{(j+1/2)}\right), \\ \widehat{\mathbf{h}}^{(j+1)} = \widehat{\mathbf{h}}^{(j+1/2)} + \sqrt{2\psi}\Delta t\boldsymbol{\sigma}\left(\widehat{\mathbf{h}}, \widehat{\mathbf{N}}\right), \\ \widehat{\mathbf{h}}^{0} = \mathbf{f}, \end{cases} \tag{110}$$

where $\mathbf{L}$ is defined as in Eq. (90).

4.3 Capillary Wave Evolution

To test our numerical methods we compare the Fourier transform of their numerical solution to Eq. (53). For numerical solutions to the linear system we expect this to be followed

exactly for all noise levels, while in solutions to the nonlinear system we only expect the spectrum to evolve in this way while the perturbation size from the initial conditions is small. For this reason we test the systems with only small values of ψ.

For brevity, we only include plots of the spectrum for the Fourier spectral method, Eq. (101), for the linearised STFE, and the finite difference scheme, Eq. (106), for the full STFE in the main body of the text, plots of the spectrum evolution for other methods are all located in Appendix. (B). We see in Figure. (5) and Figure. (6) both capture the spectrum effectively under small timestep sizes (in this case around 1×10^{-11}). It is important to note however that the plot in Figure. (5) showing the spectrum evolution for the non-linear system is cropped to the first 800 wavenumbers. In Figure. (9), Appendix. (B), we see that the method does not capture the evolution of the largest wavenumbers so well. While we do not expect the linear system to be full representative of the nonlinear system, the rapid decrease to zero in the larger wavenumbers has the appearance of being an artifact of the method rather than a physical phenomenon. The Spectral scheme, Eq. (110), also suffers from these artifacts.

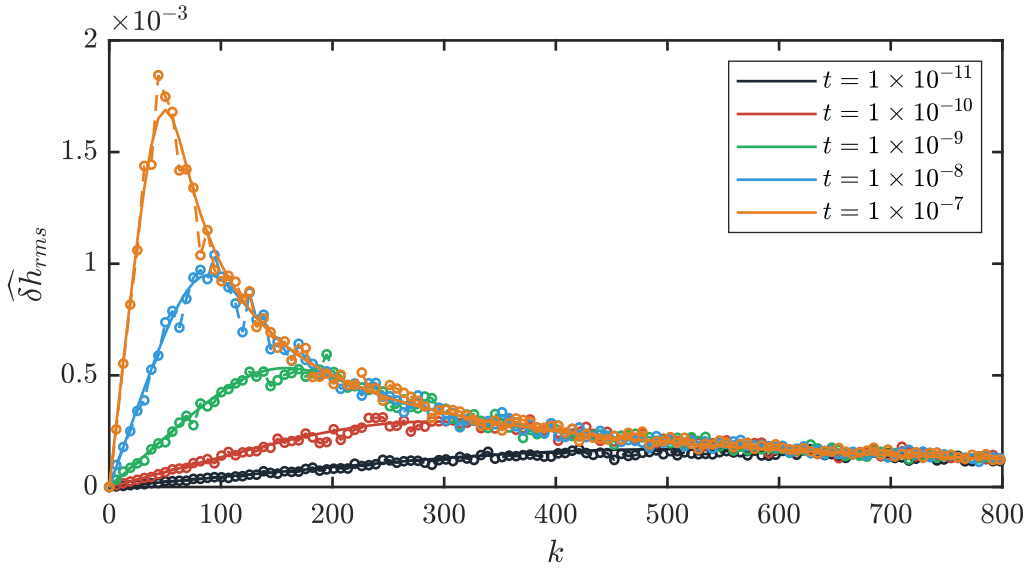


Figure 5: Spectrum evolution in Eq. (53) (solid lines) and RMS of 50 numerical realisations (dashed lines) from the scheme in Eq. (106) with $\Delta t = 1 \times 10^{-11}$ and 512 spatial grid points.

This is similar in the linear case. While Figure. (6) shows the full spectrum, illustrating that for this small timestep (1×10^{-11}), the simple spectral approach, Eq. (101), for the linear system captures the full spectrum effectively, we see in Figure. (10) that it fails to capture the full spectrum for larger timestep values. Integrating the deterministic part of the equation analytically, Eq. (102), does not remedy this, in fact, it is seen in Figure. (11) that it is less effective than simple Euler integration, failing to capture the latter part of the spectrum completely for larger timestep values. Performance of the nonlinear methods under larger timestep values is found to be similar¿ The evolution of smaller wavenumbers

is captured effectively while there are artifacts in the evolution of larger wavenumbers. Qualitatively, the reason for failure of the methods to capture the large wavenumbers is because the deterministic part of the equation acts as a relaxation of order 4 in the wavenumber for each of the Fourier modes, while the effects of the noise are only order 1 in the wavenumber. So if the timestep is too large, the modes for the larger wavenumbers are instantly relaxed, and the cumulative effects of the noise are unable to develop. The only method that does not suffer from this problem is the analytical method, Eq. (103). We see in Figure. (12) that this method successfully captures all parts of the spectrum evolution for our larger timestep value. This is not surprising since it is essentially an analytical solution to the system. However, this illustrates that the approximation of the stochastic integral being used can be improved substantially. This is a potential area for future development in this work.

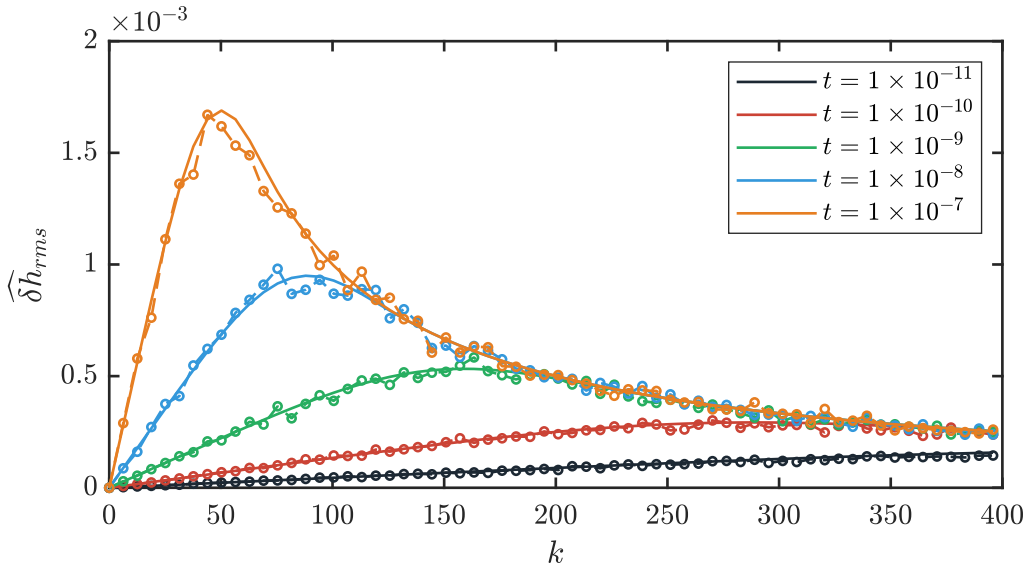


Figure 6: Spectrum evolution in Eq. (53) (solid lines) and RMS of 100 numerical realisations (dashed lines) from the scheme in Eq. (101) with $\Delta t = 1 \times 10^{-11}$ and 512 spatial grid points.

4.4 Summary

In this section we derived implicit Finite difference schemes for the linearised and full thin film equation, along with an implicit spectral scheme for the linearised thin film equation and a semi-implicit pseudo-spectral scheme for the full thin film equation. To verify these methods we benchmarked them against a known result for the evolution of capillary waves due to thermal fluctuations. Now that we are confident that all of these schemes work, we briefly evaluate each of the schemes.

For the linear system, we again find the spectral scheme to be superior. Specifically the spectral scheme based off the analytical solution to the linear system. It is not surprising that this method captures the evolution of the spectrum most effectively. We find the

spectral schemes that do not use this analytical solution and the finite difference scheme are roughly as effective as each other.

For the full thin film equation, the finite difference method is superior for many reasons. The first is that it is a fully implicit method, in comparison to the pseudo-spectral method which is only semi-implicit. This leads to the finite difference scheme having slightly better stability properties than the pseudo-spectral scheme. This will be particularly important in the next section where we study ruptures in this films. We also choose to use the finite difference scheme for practical reasons. Since deriving finite difference schemes is easier in the presence of non-linearities, the finite difference scheme is more adaptable. For example we may wish to consider the effects of disjoining pressure as means for rupture in addition to thermal fluctuations, and adding this term is much easier when using the finite difference method. Furthermore, since both the finite difference scheme and the pseudo-spectral scheme fail to capture the evolution of the larger wavenumbers at practical timesteps, the claim of 'exponential convergence' for the spectral method is less convincing than in the deterministic case, even if the meaning of this in the context of stochastic PDEs is less clear. For these reasons we will be using the implicit finite difference scheme derived in Sec. (4.2.1) when exploring ruptures in thin liquid films in the next section.

5 Ruptures in Thin Fluid Films

Spontaneous rupture is a frequently observed phenomenon in thin liquid films. Its mechanisms are of interest due to its wide reaching applications. In the deterministic setting, the film is made linearly unstable through the addition of Van der Waals forces, otherwise known as 'disjoining pressure' [6]. Recently, thermal fluctuations have been investigated as a mechanism for the destabilisation of linearly stable films to the point of a Van der Waals induced rupture [15]. Here, we investigate the viability of thermal fluctuations alone as a mechanism for ruptures in thin liquid films.

We cover the analytical results for the rupture profile and the expected time taken for the film to rupture derived in [15], and attempt to verify these results computationally. However, rupture in thin fluid films is a rare event and using Monte-Carlo methods for computing waiting times for rupture is either computationally infeasible or has very high variance for small noise values. We suggest the 'importance splitting' method for sampling instances of the rare event and estimating its waiting time more efficiently than standard Monte-Carlo. We also propose the use of Markov-Chain approximation, a method for efficiently capturing macroscopic properties of the system.

5.1 Rare Event Theory for Ruptures in Thin Films

We would first like to derive the expected rupture profile for a rupture caused entirely by thermal fluctuations. To do this we follow the method used in [15]. We begin by recalling the stochastic gradient flow form of h given in Eq. (42). This characterises the evolution h as begin forced down the direction of greatest descent of $\mathcal{H}(h)$ subject a to mass conservation and mass conserving thermal fluctuations. The maximum likelyhood rupture profile of h is therefore the profile that minimises the energy functional $\mathcal{H}$ subject to mass conservation. In the limit $\psi \to 0$, the expected rupture profile converges to this as well. We derive the optimal rupture profile using Lagrangian optimisation as follows.

We seek the profile $h : [0,1] \to \mathbb{R}$ with $h\left(\frac{1}{2}\right) = 0$ that minimises the energy functional $\mathcal{H}$ given by

$$\mathcal{H}(h) = \int_0^1 \frac{1}{2}(\partial_x h)^2 \mathrm{d}x, \tag{111}$$

subject to mass conservation and periodicity of derivatives up to order 4. That is

$$\begin{cases} M(h) = \int_0^1 (h - h_0)\mathrm{d}x = 0, \\ \partial_x^\alpha h(0,t) = \partial_x^\alpha h(1,t) \text{ for all } \alpha \in \{0,\dots,4\}. \end{cases} \tag{112}$$

It is clear that such a profile h will be symmetric about $x = \frac{1}{2}$. We use this to simplify the problem, writing $\mathcal{H}(h) = 2\int_0^{\frac{1}{2}} \frac{1}{2}(\partial_x h)^2 \mathrm{d}x$ and $M(h) = 2\int_0^{\frac{1}{2}} (h - h_0)\mathrm{d}x$. With this assumption we can also conclude that $\partial^\alpha h(0,t) = 0$ for each $\alpha \in \{0,\dots,4\}$. The solution

to this minimisation problem will be a critical point of the Lagrangian

$$\mathcal{L}(h) = 2\int_0^{\frac{1}{2}} \frac{1}{2}(\partial_x h)^2 \mathrm{d}x - 2\lambda \int_0^{\frac{1}{2}} (h - h_0)\mathrm{d}x. \tag{113}$$

Note that the meaning of $\mathcal{L}$ has changed from its original use as the spatial differential operator.

Let δh be a periodic perturbation that is symmetric about $x = \frac{1}{2}$ with $\delta h\left(\frac{1}{2}\right) = 0$. Ignoring higher order terms of δh we see that

$$\mathcal{L}(h + \delta h) - \mathcal{L}(h) = 2\int_0^{\frac{1}{2}} \frac{1}{2}\left(2\partial_x \delta h \partial_x h\right)\mathrm{d}x - 2\lambda \int_0^{\frac{1}{2}} (\delta h)\mathrm{d}x \tag{114}$$

$$= 2\int_0^{\frac{1}{2}} -\partial_x^2 h \delta h - \lambda \delta h \mathrm{d}x, \tag{115}$$

where to have used integration by parts to deal with the first term of the right hand side. This gives the first variation of $\mathcal{L}$ to be

$$\frac{\delta \mathcal{L}}{\delta h} = -2\partial_x^2 h - 2\lambda. \tag{116}$$

Setting this equal to zero and solving for h we obtain that $h(x) = 6h_0(x + \frac{1}{2})(\frac{1}{2} - x)$ for $x \in [0, 1/2]$. Reflecting about $x = \frac{1}{2}$, we obtain the expected rupture profile

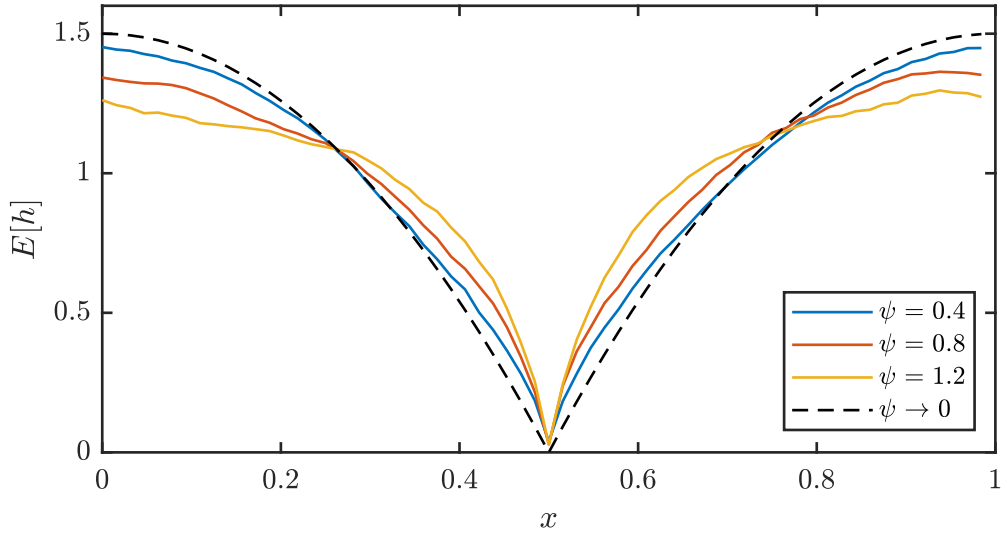
$$E[h(x)] = 6h_0 \left|x - 1/2\right| \left(1 - \left|x - 1/2\right|\right). \tag{117}$$

We also note the result for expected rupture time found in [15]. It is found that the expectation of the rupture time, τ, has the form

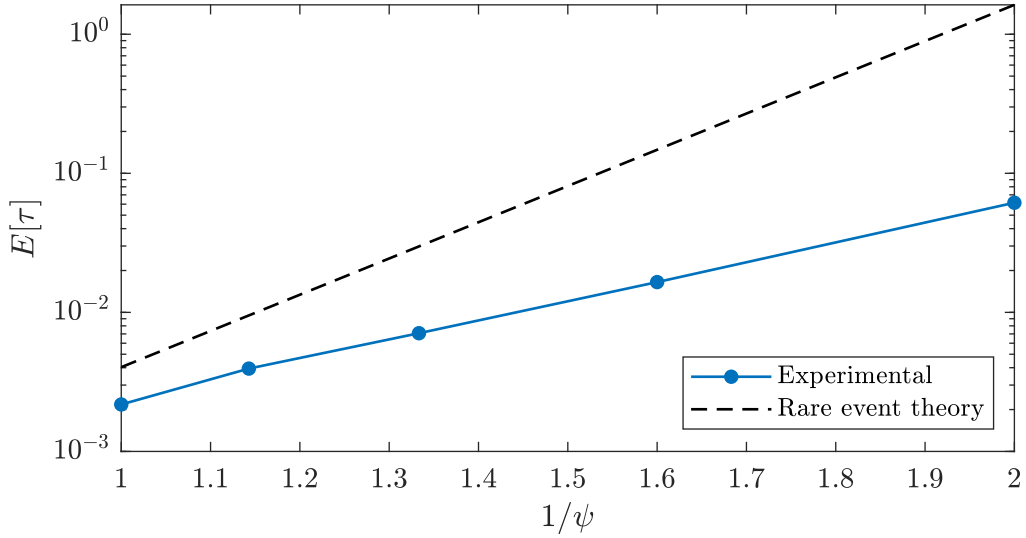
$$E[\tau] = Ce^{\frac{6h_0^2}{\psi}}. \tag{118}$$

We see in Figure (7a) that the our computational estimate for the rupture profile approaches the analytical solution as the noise strength is decreased. However, the experimental expected rupture time seen in Figure (7b) does not fit the theoretical result so well. While the relationship between expected rupture time and $\frac{1}{\psi}$ is found to be exponential, the exponent is not as high as that predicted by the rare event theory. It is possible that this is because the value of ψ is not small enough such that we are in the rare events regime. Testing smaller values of ψ with a Monte-Carlo approach is not feasible without additional computing power. For this reason we propose the use of sampling methods to allow us to observe the rare event more efficiently for small values of ψ.

An established method used in the sampling of rare events is importance splitting [2]. This is where the probability of a rare event is estimated by defining a nested sequence of less



(a) Mean rupture profiles from 100 Monte Carlo realisations of the full STFE.



(b) Mean rupture times from 100 Monte Carlo realisations of the full STFE.

Figure 7

rare events and computing the conditional probability of each event given the previous one. In our case, the rare event in question is the rupture of the tin film by time T. We do not address the implementation of importance splitting in this report for brevity. We point the reader to [16] and [2] for mathematical formulation, analysis and implementation of importance splitting. When implemented appropriately, importance splitting gives an unbiased estimate for the probability of the rare event with lower variance than a Monte-Carlo approach with an equal number of samples. We choose instead to develop the idea of Markov Chain approximation.

5.2 Markov Chain Approximation

While importance splitting allows us to sample rupture profiles and estimate the probability of rupture by a certain time, there are other properties we might like to capture about the film on its way to rupture. For example we might wish to estimate a probability distribution for the rupture time rather than just an expected rupture time. We might also like to find an expected path of $h_{\min}$ over time. The proposed method, Markov chain approximation, allows these things to be done. Using Markov chains to approximate stochastic systems is not a concept original to this work. However, to the author's knowledge, the application of the to systems with absorbing states is novel. As before, we aim to approximate the stochastic process $(h_{\min}(t) : t \geq 0)$ over time. While $h_{\min}$ is a continuous time stochastic process, we will instead aim to approximate the discrete time stochastic process $\left(H^{(j)} : j \in \mathbb{N}\right)$ defined by $H^{(j)} = h_{\min}\left(\tilde{t_j}\right)$ for uniform sample times $\tilde{t_j} = j\tilde{\Delta}t$, with $\tilde{\Delta}t$ being the sample interval, not necessarily equal to the timestep Δt used in our numerical methods.

Assuming the film is initially in equilibrium, $h(x, 0) = 1$ for each $x \in [0, 1]$, we see that $H^{(j)}$ takes values in $[0, 1]$, where we always have $H^{(0)} = 1$. The state $H^{(j)} = 0$, corresponding to the rupture of the film, is referred to as an absorbing state. An absorbing state of a stochastic process, for our purposes, is a subset A of the state space I such that $\mathbb{P}(H^{(i)} \in A | H^{(j)} \in A) = 1$ for all $i \geq j$. In or case we see that $H^{(j)} = 0$ is absorbing since it corresponds to the end of the simulation.

If we know the values $H^{(j)} = y^{(j)}$ for $j = 1, \ldots, n-1$, then $H^{(j)}$ is a random variable completely determined by a probability density function g, in the sense that for any $0 \leq a < b \leq 1$ we have

$$\mathbb{P}\left(H^{(n)} \in [a, b] \middle| H^{(j)} = y^{(j)} \text{ for } j = 1, \ldots, n-1\right) = \int_a^b g(x)\mathrm{d}x, \tag{119}$$

The first approximation we make on $H^{(j)}$ is to replace it with a discrete state stochastic process $\left(\tilde{H}^{(j)} : j \in \mathbb{N}\right)$. To do this we choose any $\delta > 0$ and partition the interval $[-\delta, 1+\delta]$ into $N + 2$ intervals $(I_n)_{n=0}^{N+1}$ defined by the sequence $0 = x_0 < x_1 < \cdots < x_N = 1$, where $I_0 = [-\delta, 0]$, $I_{N+1} = [1, 1 + \delta]$, $I_n = (x_{n-1}, x_n]$ for $n = 1, \ldots, N-1$ and $I_N = (x_{N-1}, x_N)$. The extension by δ is only so that the initial state $H^{(0)}$ and the absorbing rupture state have their own intervals. We will refer to these intervals as 'bins'.

We then define the approximating discrete state stochastic process $\left(\tilde{H}^{(j)}\right)$ by $\tilde{H}^{(j)} = \sum_{k=0}^{N+1} k \mathbb{1}_{I_k}$, i.e. the value of $\tilde{H}^{(j)}$ is the index of the bin containing $H^{(j)}$. In the same sense as with the continuous state process, if we know the values $\tilde{H}^{(j)} = y^{(j)}$ for $j = 1, \ldots, n-1$, then $\tilde{H}^{(j)}$ is a random variable completely determined by a probability mass function

$\tilde{g} : \{0, \ldots, N+1\} \to \mathbb{R}$ in the sense that

$$\mathbb{P}\left(\tilde{H}^{(j)} = k \middle| \tilde{H}^{(i)} = y^{(i)} \text{ for } i = 1, \ldots, n-1\right) = \tilde{g}(k). \tag{120}$$

Note that up to this point, we have not made any approximations on H, we have recast it into a discrete form $\tilde{H}$, but this is constructed so that it completely captures the behaviour H in its discrete setting. The main approximation we make is known as the Markov projection. To define this we quote almost directly from [10]. To begin with, at timestep j we define the first order transition matrix $\mathbf{P}^{(j)}$ by

$$\left(\mathbf{P}^{(j)}\right)_{mn} = \mathbb{P}\left(\tilde{H}^{(j)} = n \middle| \tilde{H}^{(j-1)} = m\right). \tag{121}$$

Then Markov projection of $\tilde{H}$ is the Markov process $\bar{H}$ defined by

$$\begin{cases} \bar{H}^{(0)} = \tilde{H}^{(0)}, \\ \mathbb{P}\left(\bar{H}^{(j)} = n \middle| \bar{H}^{(j-1)} = m\right) = \left(\mathbf{P}^{(j)}\right)_{mn} \text{ for } j \in \mathbb{N}_{>0} \text{ and } m, n \in \{0, \ldots, N+1\}, \end{cases} \tag{122}$$

i.e. the Markov projection $\bar{H}$ is Markov process with the same transition probabilities and initial distribution as $\bar{H}$. Since a Markov chain is uniquely defined by its initial distribution and transition probabilities, the Markov projection is unique. Ignoring time dependence past the first order in $\tilde{H}$ makes finding the evolution of the distribution of $\bar{H}$ very easy. Given an initial distribution $\bar{\pi}^{(0)}$, the distribution $\bar{\pi}^{(j)}$ of $\bar{H}^{(j)}$ is given by

$$\bar{\pi}^{(j)} = \Pi_{k=1}^{j} \mathbf{P}^{(k)} \pi^{(0)} \text{ for every } j \in \mathbb{N}. \tag{123}$$

Now we must ask whether the Markov projection $\bar{H}$ accurately captures the behaviour of $\tilde{H}$. In fact $\bar{H}$ is first order equivalent to $\tilde{H}$ [10]. This means that given the same initial distribution $\bar{\pi}^{(0)} = \tilde{\pi}^{(0)} = \pi^{(0)}$, we have

$$\tilde{\pi}^{(j)} = \Pi_{k=1}^{j} \mathbf{P}^{(k)} \pi^{(0)} = \bar{\pi}^{(j)} \text{ for every } j \in \mathbb{N}. \tag{124}$$

So $\bar{H}$ accurately captures the evolution of the distribution of $\tilde{H}$. This is sufficient for the purposes of estimating waiting times for rupture and even in the expected path of $\tilde{H}$ over time, since these are only dependant on the evolution of its distribution.

Now that we have argued for the validity of the method, we will show how it can be used to estimate waiting times for rupture.

5.2.1 Estimating Rupture Times

If we have the transition matrices $\mathbf{P}^{(j)}$ for every $j \in \mathbb{N}$, we can compute the expected rupture time directly as

$$E[\tau] = \sum_{j=1}^{\infty} \tilde{t}_j \left[\pi_{N+1}^{(j)} - \pi_{N+1}^{(j-1)} \right]. \tag{125}$$

Of course we do not have access to all of these, since we can only run the simulation for finite time. Additionally, computing the transition matrices for every $j \in \mathbb{N}$ would not save any computation time, since to compute the transition matrices for time T, a sufficient number of instances of H must be simulated from time $j = 0$ to $j = T$ to estimate the transition probabilities.

Instead of doing this, we note that it can be seen experimentally that the first order transition matrix $\mathbf{P}^{(j)}$ of $\tilde{H}^{(j)}$ converges to a steady state $\mathbf{P}^{\infty}$. That is, for sufficiently large j, it is reasonable to approximate $\mathbf{P}^{j} \approx \mathbf{P}^{\infty}$. The Markov chain with transition matrix $\mathbf{P}^{(\infty)}$ is an absorbing Markov chain [17, Chapter. 3].

Choosing the time index T as the cut off point after which we make the assumption of constant transition probabilities, for $j > T$ we can approximate the evolution of the distribution, $\pi^{(j)}$, as

$$\pi^{(j)} = (\mathbf{P}^{\infty})^{j-T} \Pi_{k=1}^{T} \mathbf{P}^{(k)} \pi^{(0)} \tag{126}$$

$$= (\mathbf{P}^{\infty})^{j-T} \pi^{(T)}. \tag{127}$$

Let $\left(G^{(j)} : j \in \mathbb{N}\right)$ be the Markov chain with transition matrix $\mathbf{P}^{\infty}$ and initial distribution $\pi^{(T)}$. The state $G^{(j)} = 0$ is an absorbing state of this Markov chain. Let $\mathbf{Q}$ denote the transient part of $\mathbf{P}^{\infty}$ obtained by removing the row and column of $\mathbf{P}^{\infty}$ associated with the absorbing state. The fundamental matrix of the Markov chain is given by $\mathbf{F} = (\mathbf{I} - \mathbf{Q})^{-1}$. It is a known result that the sum of the elements on the i^{th} row of $\mathcal{F}$ gives the expected number of timesteps before reaching the absorbing state given the initial state i [17, Chapter. 3]. That is

$$E\left[\tau_G | G^{(0)} = i\right] = \sum_{j=1}^{N+1} F_{ij}, \tag{128}$$

where τ_G is measured in number of timesteps. Since the initial distribution of G is $\pi^{(T)}$, we see that the expected absorption time of G is

$$E\left[\tau_G\right] = \sum_{i=1}^{N+1} \pi_i^{(T)} \sum_{j=1}^{N+1} F_{ij}, \tag{129}$$

To find the expected absorption time of $\bar{H}$, we must also add on the contribution from absorptions that happen before time T. Adding in this contribution gives us our expected

rupture time

$$E[\tau] = \sum_{j=1}^{T} \tilde{t}_j \left[\pi_{N+1}^{(j)} - \pi_{N+1}^{(j-1)}\right] + E\left[\tau_G\right] \tilde{\Delta} t. \tag{130}$$

This gives a convincing mathematical argument for the validity of approximating the stochastic process $\left(\tilde{H}^{(j)} : j \in \mathbb{N}\right)$ with its Markov projection $\left(\bar{H}^{(j)} : j \in \mathbb{N}\right)$. We are left with the problem of accurately and efficiently finding the Markov projection of $\left(\tilde{H}^{(j)} : j \in \mathbb{N}\right)$.

5.2.2 Finding the Markov Projection

Estimating the Markov projection of $\tilde{H}$ reduces to the problem of finding the first order transition matrix $\mathbf{P}^{(j)}$ for $j = 1, \ldots, T$. At sample time j, we can estimate the probabilities $\mathbb{P}\left(\tilde{H}^{(j)} = n | \tilde{H}^{(j-1)} = m\right)$ by taking M i.i.d. samples $\tilde{H}_1^{(j)}, \ldots, \tilde{H}_M^{(j)}$ from the distribution $\mathcal{D}\left(\tilde{H} | \tilde{H}^{(j-1)} = m\right)$ and use the estimator

$$\mathbb{P}\left(\tilde{H}^{(j)} = n | \tilde{H}^{(j-1)} = m\right) \approx \frac{1}{M} \sum_{i=1}^{M} \delta_{n \tilde{H}_i^{(j)}}, \tag{131}$$

This leaves the question of obtaining the samples. The simplest way that this can be done is through running simulations from time $t = 0$ to $t = \tilde{t}_j$ until we have M samples for which $\tilde{H}^{(j-1)} = m$. However, this is inefficient, in fact, it renders the method much slower than even our Monte-Carlo approach for obtaining information about the process.

A faster method is to dynamically reuse data from the old simulations. This will be described by example starting from time t_0. We know that $H^{(0)} = 1$ ($\tilde{H}^{(0)} = N + 1$), since the film always begins in equilibrium. Computing the transition matrix $\mathbf{P}^{(1)}$ is simple. We run M simulations until time $\tilde{t}_1$ and estimate $\mathbf{P}$ with Eq. (131). Since our initial distribution is $\pi^{(0)} = (0, \ldots, 0, 1)$, entries of $\mathbf{P}^{(1)}$ outside of its last column are not important. Along with an estimate for $\mathbf{P}^{(1)}$, we also have M_i samples from the distribution $\mathcal{D}(\tilde{H} | \tilde{H}^{(1)} = i)$ for each $i \in \{0, \ldots, N+1\}$ from these simulations, where $\sum_{i=0}^{N+1} M_i = M$. Note that when we refer to 'samples' from the distribution $\mathcal{D}(\tilde{H} | \tilde{H}^{(1)} = i)$, we refer not only to the value of $\tilde{H}^{(1)}$, but all of the information included with it, i.e. the full height profile of the film that gave this value of $\tilde{H}^{(1)}$.

Now to approximate $\mathbf{P}^{(2)}$. For simplicity of this explanation, assume that we have at least 1 sample in each bin at time t_1. In practise this will not be the case, but we will ignore bins with no samples in at this point since if they have no samples it is very unlikely for them to be entered this early anyway and they will not have a large effect on the distribution of H.

To estimate $\left(\mathbf{P}^{(2)}\right)_{mn}$, we would like to run M simulations drawn from the distribution $\mathcal{D}(H | H^{(1)} = m)$. However, we only have M_m samples from this distribution to draw from.

For the sake of performance we sacrifice the independence of these samples. The most simple approach is to draw only from this collection of M_m samples, accepting that some will be duplicates. We can improve this slightly by artificially adding to the samples by also adding samples from nearby timesteps. That is, if $\tilde{\Delta}t \gg \Delta t$, it may be reasonable to also add samples from timesteps shortly before $\tilde{t_1}$. If these strategies fail, it may be necessary to simply perform more than M simulations starting from t_0.

Estimating $\mathbf{P}^{(3)}$ should be similar to the general case of estimating $\mathbf{P}^{(j)}$ for any $j > 2$. When estimating $\mathbf{P}^{(2)}$, we ran M simulations drawn from the distribution $\mathcal{D}(H|H^{(1)} = m)$ for each $m \in \{0, \ldots, N+1\}$. This means that some bins for $\tilde{t_2}$ will contain more than M samples, and some will contain less. Those that contain less may artificially have samples added as explained earlier. The key thing to note is that when we ran M samples from each bin to estimate $\mathbf{P}^{(2)}$, the experimental distributions in each bin are biased towards those bins that had more artificial samples added. This is counteracted by recording which bin each sample in the experimental distribution came from in the previous timestep and weighting the distribution of drawn samples according to $\pi^{(2)}$.

5.2.3 Simple SDE Test Case

Due to time constraints, the Markov chain approximation method was not applied to the stochastic thin film system. However, as a proof of concept, we will apply the method to estimate the transition times for the simple SDE

$$\mathrm{d}X = (X - X^3)\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}W, \tag{132}$$

where $\epsilon > 0$. The deterministic part of this SDE is a potential well pushing the value of X towards ± 1. As ϵ is decreased to 0, transition from $X = -1$ to $X = 1$ and vice versa becomes a rare event. We compute the transition times for this SDE experimentally, using the Euler-Maruyama method for time integration.

When using the Markov chain approximation method, we use 40 uniformly sized bins for the interval $[-1, 1]$. We aim to study transition from $X = -1$ to $X = 1$, treating this as an absorbing state. We note that the values of X are not confined to this interval, or any finite interval, however, for small ϵ, the event of X taking values outside the interval $[-L, L]$ for sufficiently large L becomes extremely unlikely and may be ignored. The values $\epsilon \in [0.3, 1.3]$ and added the 'endcap' bins $[-10, -1]$ and $[1, 10]$ to contain the non-transitional behaviour of X. We know that this stochastic process has no long term dependencies, so we assume that the first order transition matrix is constant, i.e. $P^{(t)} = p^{\infty}$ for all $t \in \mathbb{N}$. To draw samples from the transitional bins we assume that X is uniformly distributed throughout each bin. When drawing from the 'endcap' bin $[-10, -1]$, we approximate $\mathcal{D}(X|X < -1)$ by saving each of the visited X values in this bin and drawing randomly from this list. We see that even with these somewhat arbitrarily

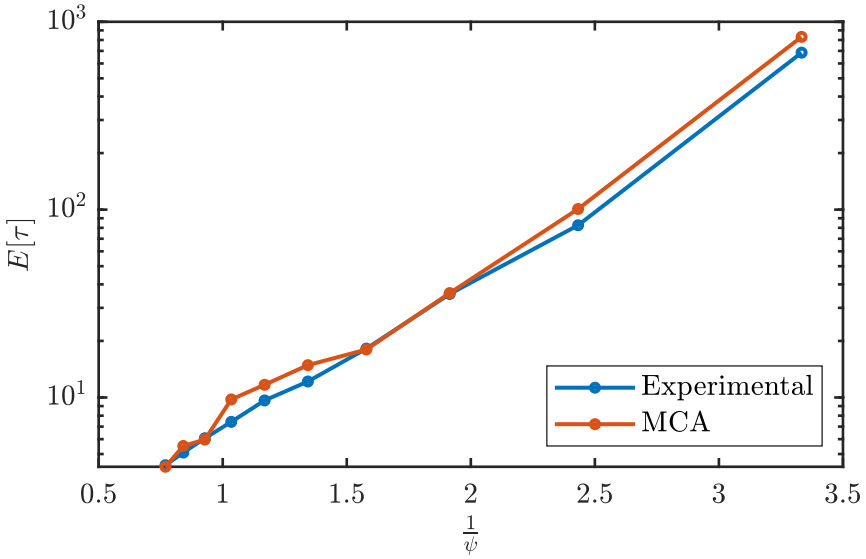


Figure 8: Experimental transition times of SDE given by equation 132.

chosen heuristics, we see in Figure (8) that the Markov chain approximation captures the transition times of the SDE very well, although further testing will be required to verify this method in general, and there is no guarantee for its efficacy when applied to more complicated stochastic processes such as the thin film system.

6 Conclusion

6.1 Summary

To begin with, we introduced the theory of thin liquid films in the deterministic setting, covering some of the known results for thin liquid films covered by [7] [19]. This included applying the lubrication approximation to the Navier-Stokes equations, we obtained a nonlinear fourth order evolution equation for the height profile of the film. Linearising this equation, we found that under small perturbations we should expect exponential decay in each of the Fourier modes of the height profile. We then moved to the stochastic setting, where a multiplicative noise term is added to the thin film equation. We then proceeded as in, the long term evolution of the spectrum took the form of a power law in the frequencies of the Fourier modes rather than the exponential decay seen in the deterministic case.

We then developed numerical schemes to simulate the deterministic evolution of the film. Schemes using the finite difference method and the Fourier spectral method were constructed to solve the linearised thin film equation and the full thin film equation. In the linear case this came down to solving a linear system of equations for both approaches. For the nonlinear system, our discretisation resulted in a non-linear system of equations which the Newton-Kantorovich [20] method had to be employed to solve. In the pseudo-spectral approach to solving the non-linear equation, we found that anti-aliasing was absolutely necessary when computing the nonlinear terms, since without it, computing the Jacobean of the residual function is not possible. Comparing the methods, we found the spectral approach to be much better for simulating the linearised system due the fact that it was essentially an analytical solution. We found the spectral method to be only slightly more effective for the nonlinear system, since despite having much better convergence properties, it was much less adaptable, harder to implement and has a higher computational complexity with the number of gridpoints than the finite difference scheme.

These methods were then adapted to find approximate solutions to the stochastic thin film equation. The analytical result for the evolution of the spectrum of the film was used as a test to verify the validity of the methods, along with being a benchmark to test their effectiveness. In the finite difference linear case, the noise was treated using an operator splitting approach, while for the spectral method it was possible to exploit an analytical solution to the stochastic thin film equation, as done in [15]. Unsurprisingly, this analytical approach was much more effective at capturing the spectrum evolution at a wide range of timesteps. While both methods were effective for very small timestep values, the finite difference approach was unable to capture the spectrum evolution for larger timestep values where the analytical spectral approach succeeded. For the nonlinear system, no such analytical solution exists, so we do not see the same advantage when using the spectral method. In fact, the multiplicative noise term must be added explicitly after an implicit time integration step when using the spectral method, since no fast

method for anti-aliasing terms involving factional powers could be found. In the finite difference approach, it was possible to treat the noise term implicitly by including it in the residual function and applying the Newton-Kantorovich method. For this reason, and for its superior adaptability, the finite difference approach was found to be superior for simulating the stochastic thin film equation.

After verifying these methods, we chose to study ruptures in thin fluid films. We began by discussing the established rare event theory for rupture, using the gradient flow form of the evolution of h we repeated the strategies in [15] to derive an expected rupture profile in the limit of small noise strength, along with making note of a result for the expected waiting time of rupture. We attempted to verify these results via a Monte-Carlo approach. The convergence of the rupture profile with decreasing ψ was promising, however the experimentally observed waiting times did not fit the the theory. We hypothesise that this is due to the value of the noise strength being too large for the rare event theory to apply. The Monte-Carlo approach becomes computationally infeasible in the limit of vanishing noise strength due to the exponential relationship between rupture time and the reciprocal of the noise strength. To study thin film rupture under these smaller noise strength values, we proposed the use of importance splitting [16] to observe the rare event. We also developed the idea of applying Markov chain approximation to the system [10] as a means for estimating the waiting times. We demonstrated the viability of the method on a simple stochastic system on which it was seen to well approximate its transition times.

6.2 Future Work

The work done here leaves much to be improved upon and developed. We intend to implement the proposed sampling methods, importance splitting and Markov chain approximation, onto the thin film system to test smaller noise strength values, with the intent of reproducing the established theoretical results. This will not be without challenges. Ensuring that the sampling methods draw appropriately from the distribution in an efficient way is something that will require development in its own right. These sampling methods will allow us to observe the rare event and estimate its properties more frequently with the use of less computational resources.

Computational efficiency can be improved in other areas too. Inspired by the effectiveness of the semi-analytical scheme for the linearised thin film equation, we hope to implement higher order methods which accurately capture the development of the film for larger timestep values. This will greatly improve the rate at which rare events such as rupture may be observed.

It is also possible that the rare event theory is not effectively reproduced by our experiments for reasons other than the noise strength value being too high. One such reason could be the choice of $h = 0.05$ as the cut off point for rupture. It is possible that setting a lower cut-off point will lead to a larger value in the exponent for the relationship between

rupture time and the reciprocal of the noise strength. This was not tested in this work because, with the timestep and grid sizes used, smaller cut off points often lead to the value of $h_{\min}$ 'overshooting' to values below zero. To test smaller cut-off values, a scheme that is adaptive in the grid size and timestep must be used. A possible branch of future work would be implementing such a scheme and using it to test these smaller cut-off values.

Bibliography

References

[1] IV. on the theory of lubrication and its application to mr. beauchamp tower's experiments, including an experimental determination of the viscosity of olive oil. *Philosophical Transactions of the Royal Society of London*, 177:157–234, December 1886.

[2] Charles-Edouard Bréhier, Tony Lelièvre, and Mathias Rousset. Analysis of adaptive multilevel splitting algorithms in an idealized case. *ESAIM: Probability and Statistics*, 19:361–394, 2015.

[3] C. R. Canning. Fluid flow in the anterior chamber of a human eye. *Mathematical Medicine and Biology*, 19(1):31–60, March 2002.

[4] Claudio Canuto, M Yousuff Hussaini, Alfio Quarteroni, and Thomas A Zang. *Spectral methods: fundamentals in single domains*. Springer Science & Business Media, 2007.

[5] René Carmona, Harry Kesten, John B Walsh, and John B Walsh. *An introduction to stochastic partial differential equations*. Springer, 1986.

[6] Kai-Seng Chou and Ying-Chuen Kwong. Finite time rupture for thin films under van der waals forces. *Nonlinearity*, 20(2):299–317, January 2007.

[7] Richard V Craster and Omar K Matar. Dynamics and stability of thin liquid films. *Reviews of modern physics*, 81(3):1131, 2009.

[8] Miguel A. Durán-Olivencia, Rishabh S. Gvalani, Serafim Kalliadasis, and Grigorios A. Pavliotis. Instability, rupture and fluctuations in thin liquid films: Theory and computations. *Journal of Statistical Physics*, 174(3):579–604, January 2019.

[9] Denys Dutykh. A brief introduction to pseudo-spectral methods: application to diffusion problems. *arXiv preprint arXiv:1606.05432*, 2016.

[10] András Faragó. Approximating general discrete stochastic processes by markov chains. *Journal of Statistics and Computer Science*, 1(2):135–145, 2022.

[11] SANDIP GHOSAL. Lubrication theory for electro-osmotic flow in a microfluidic channel of slowly varying cross-section and wall charge. *Journal of Fluid Mechanics*, 459:103–128, May 2002.

[12] Daniel T. Gillespie. Exact numerical simulation of the ornstein-uhlenbeck process and its integral. *Phys. Rev. E*, 54:2084–2091, Aug 1996.

[13] Günther Grün, Klaus Mecke, and Markus Rauscher. Thin-film flow influenced by thermal noise. *Journal of Statistical Physics*, 122(6):1261–1291, March 2006.

[14] Eugene Isaacson and Herbert Bishop Keller. *Analysis of numerical methods*. Courier Corporation, 2012.

[15] Duncan A. Lockerby James E. Sprittles, Jingbang Liu and Tobias Grafke. Rogue nanowaves: A route to film rupture. *Under Review*.

[16] Cyrille Jegourel, Axel Legay, and Sean Sedwards. Importance splitting for statistical model checking rare properties. In *Computer Aided Verification: 25th International Conference, CAV 2013, Saint Petersburg, Russia, July 13-19, 2013. Proceedings 25*, pages 576–591. Springer, 2013.

[17] John G Kemeny and James Laurie Snell. *Finite Markov chains: with a new appendix" Generalization of a fundamental matrix"*. Springer, 1976.

[18] L.D. Landau, E.M. Lifshitz, and L.P. Pitaevskii. *Course of Theoretical Physics: Statistical Physics, Part 2 : by E.M. Lifshitz and L.P. Pitaevskii*. Number v. 9. 1980.

[19] Jingbang Liu, Chengxi Zhao, Duncan A. Lockerby, and James E. Sprittles. Thermal capillary waves on bounded nanoscale thin films. *Phys. Rev. E*, 107:015105, Jan 2023.

[20] Boris Teodorovich Polyak. Newton-kantorovich method and its global convergence. *Journal of Mathematical Sciences*, 133:1513–1523, 2006.

[21] Celia Reina and Johannes Zimmer. Entropy production and the geometry of dissipative evolution equations. *Phys. Rev. E*, 92:052117, Nov 2015.

[22] Christian Schoof. Marine ice sheet stability. *Journal of Fluid Mechanics*, 698:62–72, March 2012.

[23] Jie Shen, Tao Tang, and Li-Lian Wang. *Spectral methods: algorithms, analysis and applications*, volume 41. Springer Science & Business Media, 2011.

[24] Wendy W. Zhang and John R. Lister. Similarity solutions for van der waals rupture of a thin film on a solid substrate. *Physics of Fluids*, 11(9):2454–2462, September 1999.

[25] Chengxi Zhao, Jingbang Liu, Duncan A. Lockerby, and James E. Sprittles. Fluctuation-driven dynamics in nanoscale thin-film flows: Physical insights from numerical investigations. *Phys. Rev. Fluids*, 7:024203, Feb 2022.

A Interpretation of Derivatives of Spatially Uncorrelated Variables

Consider the gaussian random variable $\mathcal{N}$ with mean $\mathbb{E}[\mathcal{N}(t)] = 0$ and correlation $\mathbb{E}[\mathcal{N}(t)\mathcal{N}(t')] = \delta(t-t')$. This may be interpreted as the derivative of the Wiener process W_t in the sense that for a given function $f \in L^2([0,T])$, we interpret the integral $\int_0^T f(t)\mathcal{N}(t)dt$ as

$$\int_0^T f\mathcal{N}dt = \int_0^T f\mathrm{d}W \tag{133}$$

where $\int_0^T f\mathrm{d}W$ is Ito's stochastic integral. Formally, we would like to interpret $\partial_t\mathcal{N}(t)$ in the same sense, so that we can write

$$\int_0^T f\partial_t\mathcal{N}dt = \int_0^T f\mathrm{d}(\partial_t W). \tag{134}$$

To do this, we must prescribe a meaning to the right hand side. One way that we can do this is to define the stochastic process W^h by $W^h(t) = \frac{W(t+h)-W(t)}{h}$ for $h > 0$. We would like to define

$$\int_0^T f\partial_t(\mathrm{d}W) = \lim_{h\to 0}\int_0^T f\mathrm{d}W^h \tag{135}$$

Now we must compute the value of $\int_0^T f\mathrm{d}W^h$. Let P_n be a sequence of partitions of $[0,T]$ with $|p_n| \to 0$ as $n \to \infty$. Then

$$\int_0^T fdW^h = \sum_{j=1}^{N_n} f(t_j)\left[\frac{(W(t_j+h)-W(t_j)) - (W(t_{j-1}+h)-W(t_{j-1}))}{h}\right] \tag{136}$$

$$= \frac{1}{h}\left[\sum_{j=1}^{N_n} f(t_j)\left[W(t_j+h)-W(t_{j-1}+h)\right] + \sum_{j=1}^{N_n} f(t_j)\left[W(t_{j-1})-W(t_j)\right]\right] \tag{137}$$

$$= \frac{1}{h}\sum_{j=1}^{N_n}\left(f(t_j-h)-f(t_j)\right)\left[W(t_j)-W(t_{j-1})\right] \tag{138}$$

$$= \int_0^T -\frac{f(t)-f(t-h)}{h}\mathrm{d}W \tag{139}$$

And taking the limit as $h \to 0$, we obtain a suitable definition

$$\int_0^T f\mathrm{d}(\partial_t W) = -\int_0^T f'(t)\mathrm{d}W. \tag{140}$$

B Efficacy of Methods for Capturing the Full Spectrum Evolution

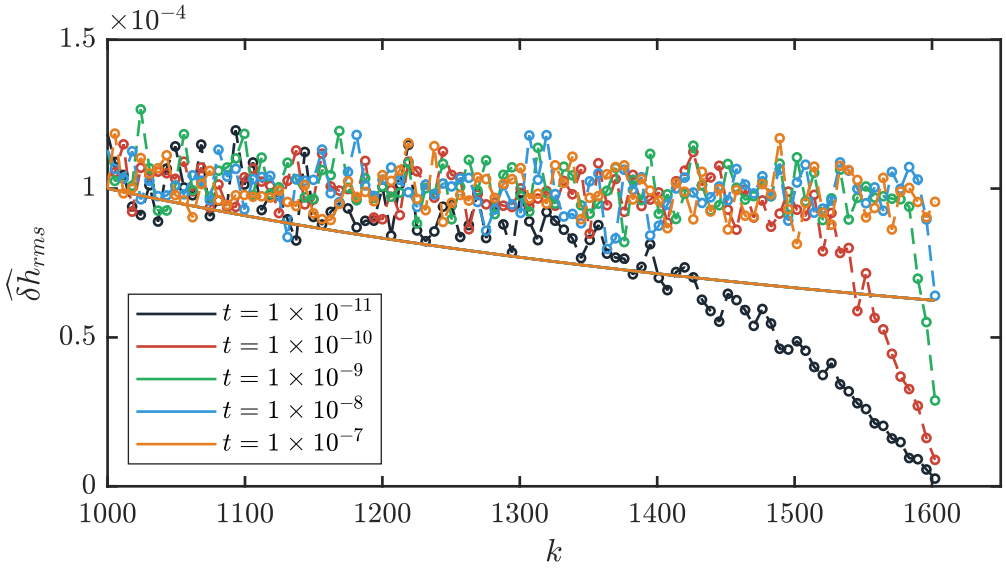


Figure 9: Spectrum evolution in Eq. (53) (solid lines) and RMS of 50 numerical realisations (dashed lines) from the scheme in Eq. (106) with $\Delta t = 1\times10^{-11}$ and 512 spatial grid points. Cropped to latter half of wavenumbers to show failure of the method when capturing the largest wavenumbers.

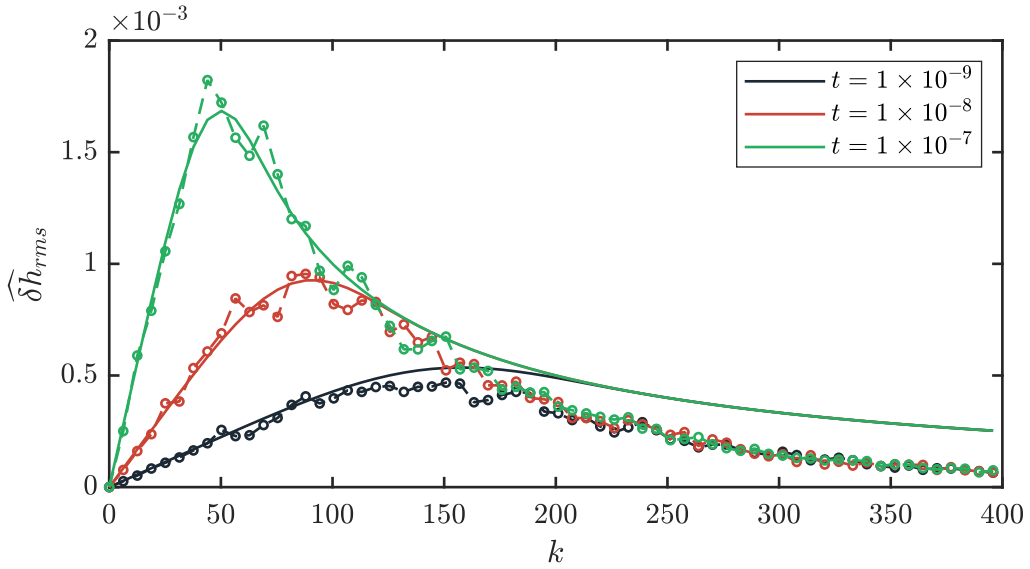


Figure 10: Spectrum evolution in Eq. (53) (solid lines) and RMS of 50 numerical realisations (dashed lines) from the scheme in Eq. (101) with $\Delta t = 1 \times 10^{-9}$ and 128 spatial grid points.

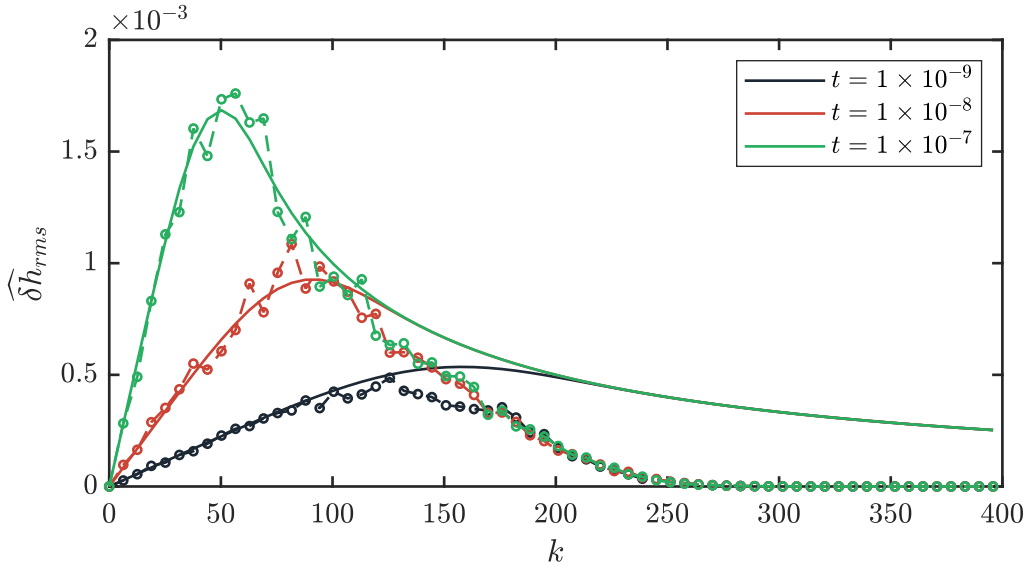


Figure 11: Spectrum evolution in Eq. (53) (solid lines) and RMS of 50 numerical realisations (dashed lines) from the scheme in Eq. (102) with $\Delta t = 1 \times 10^{-9}$ and 128 spatial grid points.

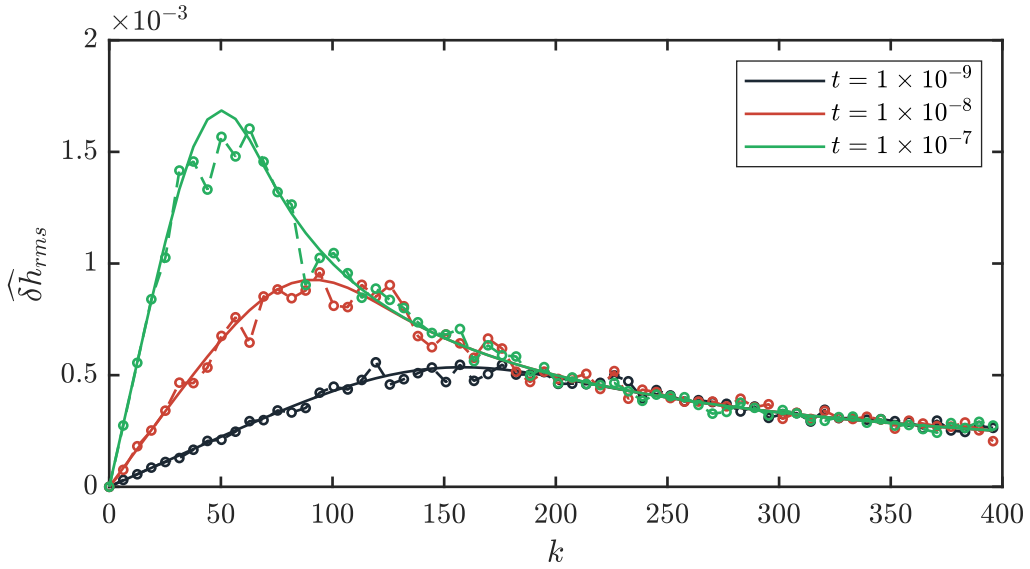


Figure 12: Spectrum evolution in Eq. (53) (solid lines) and RMS of 50 numerical realisations (dashed lines) from the scheme in Eq. (103) with $\Delta t = 1 \times 10^{-9}$ and 128 spatial grid points.