Analysis of a PDE to model Dislocation Motion

Luca Bollini

Contents

1 Introduction

1.1 Sketch

In February 2023, renowned strength sports athlete Vispy Kharadi broke the world record for most iron rods bent over the head in one minute, achieving an impressive total of 24. You can watch an overly dramatic video of this feat here: Most Iron Bars Bent. While one might first wonder what would possess a person to take on such a challenge, or if bending metal rods over your head is medically advisable (it's not), as mathematicians, we ask a different question: "How do metals bend?" A physicist might quickly produce an answer such as: "Well, the man's head is exerting a shear stress force on the rod, causing defects in the crystalline structure of the metal, called *dislocations*, to move. This results in the atoms rearranging themselves into the shape of a bent iron rod." We will go one step further. In this project, we formally introduce a model for the movement of a line dislocation, analysing the existence of solutions to the following linearised PDE. We then finish with a discussion about extending this model to include regions of different materials, simulating an alloy or a composite. Before we embark, let's give a summary of what dislocations are, and how they were discovered.

1.2 Geometry

Throughout this chapter, there will be numerous references to "Introduction to Dislocations" by Hull and Bacon [8], which can be regarded as the definitive source of background information regarding this topic. Their precise formulation of crystallographic defects' geometric structure and rich exploration of observational techniques are a fantastic way to immerse yourself in this theory.

Dislocations are the linear defects in crystalline structures responsible for plastic deformation. All crystalline materials can have these defects, but it is in metals that most of the interesting behaviour occurs. There are two types of dislocation: *edge* and *screw*. Let's outline how they look on the atomic level.

Imagine a simple cubic structure of atoms, as in figure (??), where we think of the vertical and horizontal lines as bonds between atoms, while the atoms themselves are placed at each intersection. Let's assume we may model the bonds as flexible springs between adjacent atoms, thus avoiding the complexity of how bonding works in real solids. What follows is a sequence of operations to describe how a dislocation can be formed from a perfect crystal:

1. Break all the bonds intersecting the half-plane defined by $ABCD$. CD is the *leading-edge*, where our dislocation is to be positioned, and the half-plane extends upwards in the direction of $\overrightarrow{CB}$.

2. For an edge dislocation, insert a half-plane of atoms where the bonds have just been broken. This is shown in figure (??)[1].

   For a screw dislocation, shift all the atoms on one side of $ABCD$ by one bond length in the direction $\overrightarrow{AB}$. This can result in one of two chiral structures depending on which direction the shift is in; a shift by $+\overrightarrow{AB}$ is the mirror image of a shift by $-\overrightarrow{AB}$. The "left-handed" version is illustrated in figure (??).

Note that both types of dislocation distort the bonds close to the leading-edge CD, and that this distortion decreases with distance. This will be relevant in §2.2 when discussing line tension.

Central to dislocation geometry is the concept of the *Burgers circuit* and *Burgers vector*. *"A* Burgers circuit *in a crystal containing dislocations is an atom-to-atom path which forms a closed loop"* [8, §1.4]. Crucially, if the same path is made in a crystal containing no dislocation, and the path does not close, then the circuit must contain at least one dislocation. You can see this in figure (??), where (a) shows a Burgers circuit encapsulating an edge dislocation, while (b) shows the same path superimposed onto a dislocation-free crystal. The vector required to close the loop is labelled as the *Burgers vector*. Furthermore, we can see from figure (??)that the Burgers vector of any edge dislocation is perpendicular to its dislocation line (the line is going into the page). For screw dislocations, the Burgers vector is parallel to the dislocation line.

It is also possible to see how these structures allow for the easy rearrangement of atomic bonds under shear. Figure (??)shows two layers of the cross section of a perfect crystal on the left, and an edge dislocation on the right. If we apply a shear force to the perfect crystal, we will need to break every bond along EF before shifting the top layer one to the left and re-forming the bonds. Such an action requires immense force, much more than is observed in practice. In contrast, applying the same to the crystal with an edge dislocation is relatively easy. We can just break and reform bonds one at a time to "fill the gap" caused by the dislocation. This requires several orders of magnitude less force. This phenomenon leads us to understand how dislocations were discovered.

1.3 Discovery

Modern dislocation theory first began being developed in the 1930s following calculations to determine the theoretical critical stress of materials. By

[1](reference diagram style as being similar to [8, §1.4])

critical stress, we mean the maximum stress a material can withstand before deforming plastically. As laid out by Hull and Bacon [8, §1.4], this was done in 1926 by Frenkel, who showed the shearing force required to move one row of atoms across another is given by the equation

$$\tau = \frac{Gb}{2\pi a} \sin \frac{2\pi x}{b}$$

where b is the spacing of atoms in the direction of shear, a is the spacing between rows of atoms, x is the displacement of the two rows from the stable position, and G is the shear modulus, or elastic shear stiffness, of the material. Figure (??)illustrates this setup. Realistic calculations for the maximum shearing force yield theoretical critical stresses around $\tau_{th} \approx \frac{G}{30}$. This is strikingly different from observational data, which indicates that critical stresses are generally between $10^{-8}G$ and $10^{-4}G$.

The mystery behind such contrast was solved in 1934, when three scientists, Orowan, Talyor and Polyani, independently theorised the presence of dislocations and reasoned that they could account for the difference between prediction and experiment. Their work induced an explosion of research into dislocations throughout the 1940s and into the early 1950s, which is when the famous Peach-Koehler equation [10] was first presented. We will make use of this in §2.2. It wasn't until 1956 that direct observation of dislocation movement was made by electron microscopy (figure (??)).

1.4 Motion

In this project, we shall be modelling the movement of an arbitrary dislocation (either edge or screw) under conservative motion, called *glide*. Dislocations capable of glide are called *glissile*, and the resulting process is known as *slip*; two planes of atoms slide over one another, while the only moving part is the dislocation itself. Think of this like using a squeegee to remove an air bubble trapped in a sticker. The dislocation is the air bubble and the squeegee is the shear force inducing slip. The sticker creeps forward as the air bubble moves, which you can imagine as the planes of atoms sliding over each other while the dislocation leads the way.

In figure (??), we can see that regions of crystal either side of the so-called *slip plane* remain undisturbed, and that the Burgers vector is depicted as being parallel to the direction of motion. This is because neighbouring atoms across the slip plane move relative to each other by precisely the Burgers vector. You can find many detailed diagrams of this in Hull and Bacon's book. It is for this reason that the Burgers vector is so important in dislocation motion.

Finally, we will discuss slip planes in more depth. For edge dislocations, note the Burgers vector and dislocation line are enough to uniquely specify the slip plane, while for screw dislocations this is not the case. We will always impose, however, that motion is confined to a single slip plane and further use that the Burgers vector lies in this plane.

We now aim to formally write down a model for the motion of a dislocation undergoing glide, then prove existence and uniqueness results for the derived equation.

2 Modelling

2.1 Setup

Let's first consider a static dislocation, figure (??). Curve γ represents the dislocation line in $\mathbb{R}^3$. The defect has fluctuations away from the x-axis entirely contained in the xy-plane, consistent with the observation of real dislocations in figure (??). We shall only consider the dislocation for $0 \leq x \leq L$ where we can think of L as the typical length scale along the line. This matches the choice of boundary conditions in this project. Moreover, we will assume control of L, making it as large as we wish compared to the fluctuations of γ. This will be a prominent feature in this model as we readily choose to send

$$\varepsilon = \frac{\lambda}{L} \to 0,$$

where $\lambda > 0$ is the maximum deviation of γ away from the x-axis.

Another important aspect of this model will be our assumption that γ can be parametrised by $\boldsymbol{\varphi}(x) = (x, f(x), 0)^T$ for some $f : [0, L] \to \mathbb{R}$. We say that γ is of *graphical* form — the main novelty in this project. This is in contrast to other dislocation models such as the Frank-Read source (c.f. Hudson et al.), where geometric calculus overlooks the model's analysis; we will instead get to work in just one dimension.

Next, we prescribe *periodic boundary conditions* to γ. That is, $\boldsymbol{\varphi}(0) = \boldsymbol{\varphi}(L)$ and $\boldsymbol{\varphi}'(0) = \boldsymbol{\varphi}'(L)$. For this reason, we will shift to viewing the problem on the torus in §3.1, but only consider $x \in (0, L)$ while deriving the model.

Remark 2.1. It is worth noting that while it seems reasonable to assume γ is continuously differentiable, we do not make the same assumption about f. Later on, we will see how our model makes sense even when f is defined almost everywhere, and this is in fact very much desired. We may want to simulate dislocations such as figure (??), where $f'(x) \to \infty$ as $x \to x_0$ and we make no attempt to define f at x_0. This will appear naturally once we define weak solutions in §3.3.

Now let's introduce motion. We let $\gamma(t)$ be a family of regular curves for $t \in [0, T]$ which lie entirely in the xy-plane. In the language of section 1.4, the xy-plane is the *active slip plane*, where the motion of our dislocation will be described. We will hereafter project the problem onto this plane wherever possible. Again assuming graphical form, we parametrise the dislocation at time t by

$$\boldsymbol{\varphi}(x, t) = (x, u(x, t), 0)^T \tag{1}$$

with $u : [0, L] \times [0, T] \to \mathbb{R}$ differentiable in time. We will soon show how a non-linear PDE in u can be formed under the assumption of *Quasi-static evolution*, meaning forces acting on at any point must balance at all times. The forces in question are friction, line tension, and applied shear stress.

To summarise, here are the key features of this model, which are also illustrated in figure (??):

Key features

- The dislocation curves are assumed to have graphical form (1).
- Periodic boundary conditions hold at all times t:

$$u(0,t) = u(L,t) \text{ and } \partial_t u(0,t) = \partial_t u(L,t).$$

- We may choose L to be considerably larger than λ, where

$$\lambda = \sup_{[0,L]} |u(x,0)|$$

  is the initial maximum deviation of the dislocation from the x-axis, provided $\int_0^L u(x,0)\,\mathrm{d}x = 0$. We will make use of this by sending $\varepsilon = \frac{\lambda}{L}$ to 0.

- Motion is via Quasi-static evolution, balancing forces at all times.

2.2 Force Balance

To balance any forces and derive a PDE, we must first establish how we can write down each force in question: friction, line tension and applied stress. Friction is the simplest. For any object in motion, we can write down the elementary relation

$$\boldsymbol{f}_{\text{friction}} = -\mu \boldsymbol{v},$$

where $\boldsymbol{v}$ is the velocity of the object, and μ the coefficient of friction. This also applies to dislocations, however we must be precise; $\boldsymbol{f}_{\text{friction}}$ is the *force per unit length* acting on the curve. Therefore, we should multiply by the length of curve in consideration to obtain the frictional force we use in the balance argument.

Next, let's look at line tension. After studying the geometry of dislocations, we noted how the bonds between atoms are distorted near the dislocation site, and that distortion decreases with distance. Anderson et al. [2, §6.5] describe how this distortion can be interpreted as locally exerting a line tension force on the dislocation line [2], which acts by "straightening the curve" as if it were elastic. By letting E_0 be the coefficient of dislocation

[2] However, Anderson et al. also note that *"the analogy is not exact, and the meaning of a line tension for a dislocation is somewhat nebulous"*. Our use of line tension is merely a local approximation of a much more complex picture.

stiffness [3], the line tension force that point A exerts on a neighbouring point B is expressed as

$$\boldsymbol{f}_{\text{tension}} = E_0 \boldsymbol{t}.$$

Here, $\boldsymbol{t}$ is the unit tangent vector to the curve at point A, directed away from point B.

Now for applied stress per unit length, we can use the Peach-Koehler force mentioned in §1.3 [10]

$$\boldsymbol{f}_{\text{stress}} = (\boldsymbol{\sigma}\boldsymbol{b}) \times \boldsymbol{t},$$

where $\boldsymbol{\sigma}$ is the external stress field, $\boldsymbol{b}$ is the Burgers vector, and $\boldsymbol{t}$ is the oriented unit tangent to the dislocation line. We will assume the stress field is uniform with components σ_{ij}. Stress fields inducing conservative motion are necessarily symmetric by conservation of angular momentum, so we are free to assume $\sigma_{ij} = \sigma_{ji}$. Let's restate that we do nott make any assumption about the Burgers vector other than lying in the glide plane. The Burgers vector is therefore expressed as $\boldsymbol{b} = (b_1, b_2, 0)^T$.

To derive an explicit expression for $\boldsymbol{f}_{\text{stress}}$, we compute

$$\boldsymbol{\sigma}\boldsymbol{b} = \begin{pmatrix} \sigma_{11}b_1 + \sigma_{12}b_2 \\ \sigma_{12}b_1 + \sigma_{22}b_2 \\ \sigma_{13}b_1 + \sigma_{23}b_2 \end{pmatrix} =: \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix},$$

$$\boldsymbol{t} = \frac{1}{\sqrt{1 + (\partial_x u)^2}} (1, \partial_x u, 0)^T$$

and so

$$(\boldsymbol{\sigma}\boldsymbol{b}) \times \boldsymbol{t} = \frac{1}{\sqrt{1 + (\partial_x u)^2}} \begin{pmatrix} -\alpha_3 \partial_x u \\ \alpha_3 \\ \alpha_1 \partial_x u - \alpha_2 \end{pmatrix}.$$

Appealing to the discussion in §1.4, we assume the dislocation cannot escape the glide plane, so only the projection of this force into the glide plane will have an effect on dislocation motion. We therefore conclude

$$\boldsymbol{f}_{\text{stress}} = \tilde{\sigma}\boldsymbol{n},$$

where we have set $\tilde{\sigma} = \sigma_{13}b_1 + \sigma_{23}b_2$, and $\boldsymbol{n}$ is the unit normal vector,

$$\boldsymbol{n} = \frac{1}{\sqrt{1 + (\partial_x u)^2}} (-\partial_x u, 1, 0)^T.$$

For the rest of this subsection, we fix a $t \in [0, T]$, dropping the explicit t dependence, and aim to balance the forces acting on a small section of the dislocation curve $\gamma(t)$. Let's also momentarily assume u is as differentiable as required to justify the following computations.

[3] Strictly speaking, E_0 may depend on the type of dislocation and orientation of the unit tangent: $E_0(\boldsymbol{b} \cdot \boldsymbol{t})$. This would introduce further non-linearity to the model, but not severe enough to warrant diverting our attention.

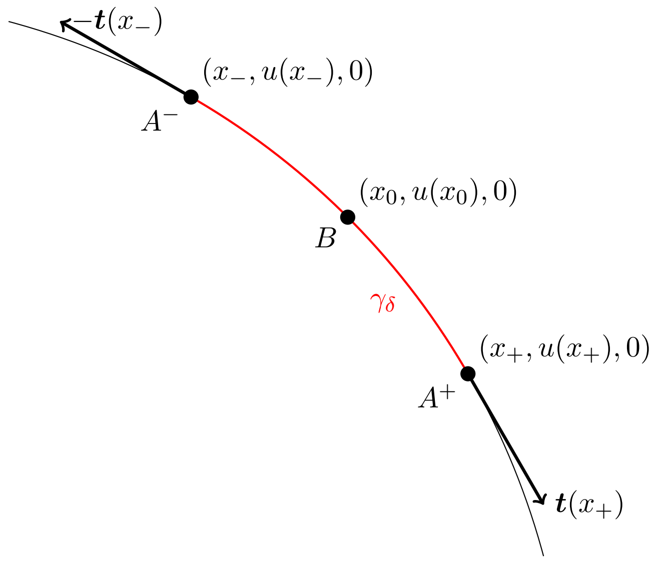

Figure 1: A small section of curve $\gamma(t)$, where the segment highlighted red is being considered for force balance.

Zooming in, as in figure 1, let γ_δ be the section of curve between $x_- = x_0 - \frac{\delta}{2}$ and $x_+ = x_0 + \frac{\delta}{2}$, for some $x_0 \in (0, L)$ and small $\delta > 0$. We label points on the curve

$$B = (x_0, u(x_0), 0)^T,$$
$$A^- = (x_-, u(x_-), 0)^T,$$
$$A^+ = (x_+, u(x_+), 0)^T$$

and note that the length of γ_δ is linearly approximated by

$$|\gamma_\delta| \approx \sqrt{\delta^2 + (u(x_+) - u(x_-))^2}.$$

This becomes a reasonable estimate when we send δ to 0. In particular,

$$\frac{|\gamma_\delta|}{\delta} \to \sqrt{1 + (\partial_x u(x_0))^2} \quad \text{as} \quad \delta \to 0. \tag{2}$$

As laid out above, we can easily see that the forces acting on B are

$$\boldsymbol{f}_{\text{tension}} = E_0 \boldsymbol{t}(x_+) - E_0 \boldsymbol{t}(x_-),$$
$$\boldsymbol{f}_{\text{friction}} = -|\gamma_\delta| \mu \boldsymbol{v}(x_0),$$
$$\boldsymbol{f}_{\text{stress}} = |\gamma_\delta| \tilde{\sigma} \boldsymbol{n}(x_0),$$

where the total line tension force is the sum of the forces acted on B by A^- and A^+. Therefore, we can conclude from our assumption of Quasi-static evolution

$$E_0 \boldsymbol{t}(x_+) - E_0 \boldsymbol{t}(x_-) + |\gamma_\delta| \left(-\mu \boldsymbol{v}(x_0) + \tilde{\sigma} \boldsymbol{n}(x_0)\right) = 0. \tag{3}$$

Dividing through by δ and taking the dot product with $\boldsymbol{n}(x_0)$, we find

$$\frac{E_0\boldsymbol{t}(x_+) - E_0\boldsymbol{t}(x_-)}{\delta} \cdot \boldsymbol{n}(x_0) + \frac{|\gamma_\delta|}{\delta}\left(-\mu\boldsymbol{v}(x_0) \cdot \boldsymbol{n}(x_0) + \tilde{\sigma}\right) = 0.$$

If we take $\delta \to 0$, we can see from the definitions of x_- and x_+ that

$$\frac{E_0\boldsymbol{t}(x_+) - E_0\boldsymbol{t}(x_-)}{\delta} \to E_0\boldsymbol{t}'(x_0)$$

and so together with the limit (2),

$$E_0\frac{\boldsymbol{t}'(x_0) \cdot \boldsymbol{n}(x_0)}{\sqrt{1 + (\partial_x u(x_0))^2}} - \mu\boldsymbol{v}(x_0) \cdot \boldsymbol{n}(x_0) + \tilde{\sigma} = 0. \tag{4}$$

Now referring back to the parametrisation (1), it is straightforward to compute

$$\boldsymbol{t}'(x) = \left(\partial_x\left(\frac{1}{\sqrt{1 + (\partial_x u)^2}}\right), \partial_x\left(\frac{\partial_x u}{\sqrt{1 + (\partial_x u)^2}}\right), 0\right)^T \tag{5a}$$

$$\boldsymbol{n}(x) = \frac{1}{\sqrt{1 + (\partial_x u)^2}}(-\partial_x u, 1, 0)^T \tag{5b}$$

$$\boldsymbol{t}'(x) \cdot \boldsymbol{n}(x) = \sqrt{1 + (\partial_x u)^2}\,\partial_x\left(\frac{\partial_x u}{\sqrt{1 + (\partial_x u)^2}}\right), \tag{5c}$$

while in order to compute $\boldsymbol{v}(x)$ we must think carefully about how points on the dislocation move. If, in equation (3), we instead take the dot product with $\boldsymbol{t}(x_0)$ and let $\delta \to 0$, we observe (since both $\boldsymbol{t}'$ and $\boldsymbol{n}$ are orthogonal to $\boldsymbol{t}$)

$$-\mu\boldsymbol{v}(x_0) \cdot \boldsymbol{t}(x_0) = 0.$$

Together with the fact that the dislocation's motion is confined to the glide plane, we deduce that $\boldsymbol{v}(x_0)$ is in the normal direction, $\boldsymbol{n}(x_0)$. Hence, the velocity vector is

$$\boldsymbol{v}(x_0) = (\partial_t\boldsymbol{\varphi}(x_0) \cdot \boldsymbol{n}(x_0))\boldsymbol{n}(x_0) = \frac{\partial_t u(x_0)}{\sqrt{1 + (\partial_x u(x_0))^2}}\boldsymbol{n}(x_0). \tag{6}$$

Finally, putting equations (4), (5) and (6) together we have derived the PDE

$$\mu\frac{\partial_t u}{\sqrt{1 + (\partial_x u)^2}} - E_0\partial_x\left(\frac{\partial_x u}{\sqrt{1 + (\partial_x u)^2}}\right) = \tilde{\sigma}. \tag{7}$$

2.3 Challenges

At this point, we must make a few observations pertaining to the difficulty of tackling this equation. First of all, this is clearly a non-linear PDE. Especially so, since the bulk of non-linearity appears in the highest order derivative. However, as we shall see, this alone will not deter us from our endeavour. The primary challenge posed by equation (7) is the $\sqrt{1+(\partial_x u)^2}$ factor dividing the $\partial_t u$ term. Multiplying through, we obtain

$$\mu\partial_t u - E_0\sqrt{1+(\partial_x u)^2}\partial_x\left(\frac{\partial_x u}{\sqrt{1+(\partial_x u)^2}}\right) = \tilde{\sigma}\sqrt{1+(\partial_x u)^2}.$$

Two challenges now emerge:

1. We see that the problematic factor is now affixed to the forcing term $\tilde{\sigma}$, so the equation is *advection–diffusion*, rather than pure diffusion. Whilst this may not appear like a trivial increase in complexity at first, it actually won't put up too much resistance against general quasi-linear PDE theory.

2. The more concerning issue is that we now see equation (7) is not of *divergence form* (more on this terminology later). This is a significant problem, as theory for non-divergence form PDEs is considerably less approachable than the theory for divergence form PDEs.

This second challenge prompts us to reconsider jumping straight into tackling such a problem head-on. Instead, we will first make a rather egregious simplification. By ignoring the $\sqrt{1+(\partial_x u)^2}$ factor dividing the $\partial_t u$ term in equation (7), we have a (still non-linear) divergence form parabolic PDE which is much more approachable with functional analytic existence theory.

Contrary to first impressions, this simplification is not entirely dim-witted. With the right Nondimensionalization, it is possible to reason that each $\sqrt{1+(\partial_x u)^2}$ factor in equation (7) approaches 1 as some small parameter $\varepsilon \to 0$. As such, understanding this simplified case may lead us to explore the right concepts for understanding the full problem. Therefore, for at least the next two chapters of this project, we will instead work towards proving theoretical results for the following equation:

$$\mu\partial_t u - E_0\partial_x\left(\frac{\partial_x u}{\sqrt{1+(\partial_x u)^2}}\right) = \tilde{\sigma}. \tag{8}$$

Once we have a good grasp of how solutions to this PDE behave (provided they exist), we shall briefly discuss how one might extend the theory to the full equation (7).

2.4 Nondimensionalization

This subsection consists of three steps which reduce equation (8) to its simplest form:

1. Adding a correction term to u that removes constant forcing.
2. Nondimensionalization; using characteristic length scales of vertical fluctuations λ, and horizontal length L to reframe the problem with dimensionless quantities.
3. Redefining constants to write the equation in its simplest form.

Step one is to replace the function u with some U that corrects for the forcing term $\tilde{\sigma}$ on the right hand side of (8). This is equivalent to changing coordinate frame into one that moves at a constant speed of $\frac{\tilde{\sigma}}{\mu}t$ in the y-direction, thus following the dislocation as it propagates. This is achieved by defining

$$U(x,t) = u(x,t) - \frac{\tilde{\sigma}}{\mu}t,$$

which transforms (8) into

$$\mu\partial_t U - E_0\partial_x\left(\frac{\partial_x U}{\sqrt{1+(\partial_x U)^2}}\right) = 0.$$

The advantage of this formulation is that we impose a sort of "zero boundary values" simply by doing nothing; the integral $\int_0^L U(x,t)\,\mathrm{d}x$ will forever remain small if it starts at 0.

For the second step, we can define new dimensionless quantities $\tilde{t}$, $\tilde{x}$ and $\tilde{u}$ by

$$T\tilde{t} = t, \qquad L\tilde{x} = x, \quad \text{and} \quad \lambda\tilde{u}(\tilde{x},t) = U(x,t),$$

where $\lambda = \sup_{[0,L]}|U(x,0)|$ is the initial maximum deviation of the dislocation from the x-axis. A basic computation reveals

$$\frac{\mu\lambda}{T}\partial_{\tilde{t}}\tilde{u} - \frac{E_0\lambda}{L^2}\partial_{\tilde{x}}\left(\frac{\partial_{\tilde{x}}\tilde{u}}{\sqrt{1+\left(\frac{\lambda}{L}\partial_{\tilde{x}}\tilde{u}\right)^2}}\right) = 0.$$

We complete step three by setting $\varepsilon = \frac{\lambda}{L}$ and $\kappa = \frac{TE_0}{\mu L^2}$. Dropping the tilde, this boils down to

$$\partial_t u - \kappa\partial_x\left(\frac{\partial_x u}{\sqrt{1+(\varepsilon\partial_x u)^2}}\right) = 0. \tag{9}$$

2.5 The Non-linear Equation

We are now ready to formulate the problem we will be analysing for the rest of this project. Briefly note that the previous subsection has normalised the equation so that $u : [0, 1] \times [0, T] \to \mathbb{R}$ and $|u(x, 0)| \leq 1$ for every $x \in [0, 1]$. As discussed in §2.1, we intend to assign periodic boundary conditions

$$u(0, t) = u(1, t), \quad \partial_t u(0, t) = \partial_t u(1, t) \tag{10}$$

to the model. An easy way to implement this is to redefine the problem on the one-dimensional torus, $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. This way, if a solution to equation (9) is continuously differentiable on $\mathbb{T}$, it necessarily satisfies periodic boundary conditions (10) when viewed as a function on $[0, 1]$.

It is also worth pointing out that equation (9) looks strikingly similar to the heat equation:

$$\partial_t u - \kappa \partial_{xx} u = 0. \tag{11}$$

In fact, we should intuitively expect this to be the case. With the forcing removed, dislocation motion reduces to a form of *curve-shortening flow*, where sharp corners in the initial data are instantly made smooth — exactly as described by the heat equation. On closer inspection, setting $\varepsilon = 0$ in equation (9) exactly recovers the heat equation, and it might be reasonable to expect solutions of (9) (provided they exist) to converge to solutions of the heat equation in some appropriate function space as $\varepsilon \to 0$. This will now be the main focus of the project.

The aims of the project are formally laid out in the following:

The ε-Problem

Define the ε-initial value problem

$$\begin{cases} \partial_t u^\varepsilon - \kappa \partial_x \left(\frac{\partial_x u^\varepsilon}{\sqrt{1+(\varepsilon \partial_x u^\varepsilon)^2}} \right) & = 0 \quad \text{in } \mathbb{T} \times (0, T] \\ \qquad\qquad\qquad\qquad u^\varepsilon & = u_0^\varepsilon \quad \text{on } \mathbb{T} \times \{0\} \end{cases} \tag{$\star$}$$

where $u_0^\varepsilon : \mathbb{T} \to \mathbb{R}$ is a given function with $|u_0^\varepsilon| \leq 1$, and $u^\varepsilon : \mathbb{T} \times [0, T] \to \mathbb{R}$ is the unknown.

Project aims

- Establish functional analytic results which demonstrate existence theory for uniformly parabolic PDEs.
- Show a unique solution to ($\star$) exists in an appropriate function space.
- Show solutions to ($\star$) converge to solutions of the heat equation (11) in an appropriate function space as $\varepsilon \to 0$.

3 Linear Existence Theory

This chapter explores the well-developed existence theory for linear elliptic and parabolic PDEs. The main reference throughout is *Partial Differential Equations* by Evans [5]. From here we state several major theorems, mostly without proof, and comment on the elegant ideas they contain. One crucial contextual difference between our problem and Evans is that we want to work on the torus in order to capture periodicity. This motivates us to first consider alternative characterisations of Sobolev spaces — the function spaces we will be working over. Let's start there.

3.1 Sobolev Spaces

To introduce this properly, let's first say some words about Fourier series. For any integrable function $u : \mathbb{T} \to \mathbb{C}$, its *Fourier coefficients* are given by

$$\hat{u}(n) = \int_{\mathbb{T}} u(x) e^{-2\pi i n x} \, \mathrm{d}x.$$

A central question in harmonic analysis is if the *Fourier series*

$$\sum_{-\infty}^{\infty} \hat{u}(n) e^{2\pi i n x}$$

converges to u, and in what sense. In the 1960s, Carleson [4] proved that Fourier series of L^2 functions converge almost everywhere. This was then generalised to L^p by Hunt [9] for $p \in (1, \infty)$. The convergence of Fourier series in the L^p norms (again $p \in (1, \infty)$) has been known since the 1930s.

Now, while the following definition may at first appear somewhat contrived, we will soon explain why it is completely natural.

Definition 3.1. Let $\mathcal{I}$ be the vector space of all integrable functions $\mathbb{T} \to \mathbb{R}$. For every $s \in \mathbb{R}$ and $p \in (1, \infty)$ we define the *inhomogeneous Sobolev spaces*, $L^p_s(\mathbb{T})$ via the norm

$$\|u\|_{L^p_s(\mathbb{T})} := \left\| \sum_{-\infty}^{\infty} \langle n \rangle^s \hat{u}(n) e^{2\pi i n x} \right\|_{L^p(\mathbb{T})},$$

where $\langle n \rangle = \sqrt{1 + 4\pi^2 |n|^2}$.

$$L^p_s(\mathbb{T}) := \{ u \in \mathcal{I} \colon \|u\|_{L^p_s(\mathbb{T})} < \infty \}.$$

We will make a similar definition for the *homogeneous Sobolev spaces* $\dot{L}^p_s(\mathbb{T})$, but first need to take care of a technicality. The homogeneous spaces are only defined up to addition by a constant, so we let $\stackrel{+}{\sim}$ be the equivalence relation

$$u \stackrel{+}{\sim} v \iff u(x) = v(x) + c \quad \text{for some } c \in \mathbb{R}.$$

The space $\mathcal{I}/\overset{+}{\sim}$ is then the space of equivalence classes of integrable functions under this relation. Observe that if we take any two functions, u and v from the same equivalence class, they will have all but one Fourier coefficient the same: $\hat{u}(n) = \hat{v}(n)$ for all $n \neq 0$. This is encoded in the following definition.

Definition 3.2. For every $s \in \mathbb{R}$ and $p \in (1, \infty)$ we define the *homogeneous Sobolev spaces*, $\dot{L}^p_s(\mathbb{T})$ with the norm on $\mathcal{I}/\overset{+}{\sim}$

$$\|[u]\|_{\dot{L}^p_s(\mathbb{T})} := \left\| \sum_{-\infty}^{\infty} |n|^s \hat{u}(n) e^{2\pi i n x} \right\|_{L^p(\mathbb{T})},$$

$$\dot{L}^p_s(\mathbb{T}) := \{[u] \in \mathcal{I}/\overset{+}{\sim} \colon \|[u]\|_{\dot{L}^p_s(\mathbb{T})} < \infty\}.$$

Here, we identify the equivalence class $[u] \in \mathcal{I}/\overset{+}{\sim}$ with any of its representatives[4]; they all yield the same norm. A crucial observation is that we are free to assume this representative has mean-zero

$$(u) = \int_{\mathbb{T}} u \, \mathrm{d}x = 0.$$

This will play an important role later. Definitions 3.1 and 3.2 are in line with those stated by Bényi and Oh [3] in their short paper *The Sobolev inequality on the torus revisited.*

One can easily check that $L^p_s(\mathbb{T})$ and $\dot{L}^p_s(\mathbb{T})$ are Banach spaces by using the completeness of L^p. A mollification argument also yields that smooth functions, $C^\infty(\mathbb{T})$, are dense in both types of spaces; one of their most appealing attributes. Let's now explain why this is a suitable setting for working with PDEs.

Let $\phi \in C^\infty(\mathbb{T})$ and $s \in \mathbb{N}$. Then we want to reason that

$$\|\phi\|_{\dot{L}^p_s(\mathbb{T})} \sim \|\partial^s \phi\|_{L^p(\mathbb{T})}.$$

We can see this by computing the Fourier coefficients of $\partial^s \phi$. Keep in mind that whenever we integrate by parts over the torus we incur no boundary terms.

$$\begin{aligned}
\widehat{\partial^s \phi}(n) &= \int_{\mathbb{T}} \partial^s \phi(x) e^{-2\pi i n x} \, \mathrm{d}x \\
&= \int_{\mathbb{T}} (2\pi i n)^s \phi(x) e^{-2\pi i n x} \, \mathrm{d}x \\
&= (2\pi i n)^s \hat{\phi}(n) \\
&= \left(i \frac{n}{|n|} \right)^s (2\pi |n|)^s \hat{\phi}(n).
\end{aligned}$$

[4]We will use the shorthand $u \in \dot{L}^p_s(\mathbb{T})$ to mean: pick any u (function) in $[u]$ (equivalence class) in $\dot{L}^p_s(\mathbb{T})$. We will emphasise when to assume u is mean-zero.

When we put these into a Fourier series, the $\left(i\frac{n}{|n|}\right)^s$ factors just correspond to phase changes and do not affect convergence. Therefore,

$$\sum_{-\infty}^{\infty} |n|^s \hat{\phi}(n) e^{2\pi inx} \in L^p(\mathbb{T}) \quad \text{if and only if} \quad \partial^s \phi \in L^p(\mathbb{T}),$$

whence we conclude the equivalence of $\|\cdot\|_{\dot{L}^p_s(\mathbb{T})}$ with the L^p norm of the order s derivative (at least for C^∞ functions). One may infer by density that the same should be true for *weak derivatives* of integrable functions.

Definition 3.3. Let $u \in \mathcal{I}$. u has weak derivative $\partial^k u$ ($k \in \mathbb{N}$) if there exists a function $v \in \mathcal{I}$ such that

$$\int_{\mathbb{T}} v\phi \,\mathrm{d}x = (-1)^k \int_{\mathbb{T}} u \partial^k \phi \,\mathrm{d}x$$

for all $\phi \in C^\infty(\mathbb{T})$. We denote $v = \partial^k u$.

The Fundamental Lemma of the Calculus of Variations immediately gives us the uniqueness of weak derivatives.

Proposition 3.4. *For any $s \in \mathbb{N}$, $p \in (1, \infty)$, the homogeneous Sobolev spaces $\dot{L}^p_s(\mathbb{T})$ consist of equivalence classes of functions u with weak derivative $\partial^s u \in L^p$.*

Proof. Take any $u \in \dot{L}^p_s(\mathbb{T})$. Then $\sum_{-\infty}^{\infty} |n|^s \hat{u}(n) e^{2\pi inx} \in L^p$, and by our computation above,

$$g := \sum_{-\infty}^{\infty} (2\pi in)^s \hat{u}(n) e^{2\pi inx} \in L^p.$$

If we integrate this against some $\phi \in C^\infty(\mathbb{T})$ we find

$$\int_{\mathbb{T}} g\phi \,\mathrm{d}x = \int_{\mathbb{T}} \int_{\mathbb{T}} \sum_{-\infty}^{\infty} (2\pi in)^s u(y) \phi(x) e^{2\pi in(x-y)} \,\mathrm{d}y\mathrm{d}x$$

where we have used the dominated convergence theorem to justify the exchange of sum and integral. Fubini now permits the exchange of integrals, and so

$$\begin{aligned}\int_{\mathbb{T}} g\phi \,\mathrm{d}x &= (-1)^s \int_{\mathbb{T}} u(y) \sum_{-\infty}^{\infty} (-2\pi in)^s \hat{\phi}(-n) e^{-2\pi iny} \,\mathrm{d}y \\ &= (-1)^s \int_{\mathbb{T}} u(y) \partial^s \phi(y) \,\mathrm{d}y.\end{aligned}$$

Precisely what it means for g to be the weak derivative $\partial^s u$.

Now observe that each equality is also true in reverse. So if we know $\partial^s u \in L^p$, then we also know it must be equal to g by the uniqueness of weak derivatives, so $u \in \dot{L}^p_s(\mathbb{T})$. □

A similar result holds for the inhomogeneous spaces:

Proposition 3.5. *For any $s \in \mathbb{N}$, $p \in (1, \infty)$, the inhomogeneous Sobolev spaces $L^p_s(\mathbb{T})$ consist of all functions $u \in L^p$ with weak derivatives $\partial^k u \in L^p$ for each $k = 1, 2, \ldots, s$.*

Proof. We will show that for any $u \in \mathcal{I}$,

$$\|u\|_{L^p_s(\mathbb{T})} \sim \|u\|_{L^p(\mathbb{T})} + \sum_{k=1}^{s} \|u\|_{\dot{L}^p_k(\mathbb{T})}$$

which combined with the previous proposition concludes the result.

First consider, for each $n \in \mathbb{Z}$,

$$m(n) := \frac{(1 + 4\pi^2 |n|^2)^{\frac{s}{2}}}{1 + (4\pi^2 |n|^2)^{\frac{s}{2}}}.$$

$m(0) = 1$ and otherwise $m(n) \leq 2^{\frac{s}{2}}$. Thus multiplying Fourier coefficients by m will not change convergence of the series, meaning

$$\sum_{-\infty}^{\infty} \langle n \rangle^s \hat{u}(n) e^{2\pi i n x} = \sum_{-\infty}^{\infty} m(n) \left(1 + (4\pi^2 |n|^2)^{\frac{s}{2}}\right) \hat{u}(n) e^{2\pi i n x}$$

is in $L^p(\mathbb{T})$ if $u \in L^p(\mathbb{T}) \cap \dot{L}^p_s(\mathbb{T})$. So

$$\|u\|_{L^p_s(\mathbb{T})} \lesssim \|u\|_{L^p(\mathbb{T})} + \|u\|_{\dot{L}^p_s(\mathbb{T})}.$$

For the other direction, consider instead

$$m_0(n) := \frac{1}{(1 + 4\pi^2 |n|^2)^{\frac{s}{2}}} \quad \text{and} \quad m_k(n) := \frac{(4\pi^2 |n|^2)^{\frac{k}{2}}}{(1 + 4\pi^2 |n|^2)^{\frac{s}{2}}}$$

for each $k = 1, 2, \ldots, s$. Clearly $m_0(n) \leq 1$ for every $n \in \mathbb{Z}$, while $m_k(0) = 1$ and $m_k(n) \leq (4\pi^2 |n|^2)^{\frac{k-s}{2}}$ otherwise. Clearly, $\frac{k-s}{2} \leq 0$, so $m_k(n) \leq 1$ for every $n \in \mathbb{Z}$, too. With similar reasoning as above, we conclude

$$\|u\|_{L^p_s(\mathbb{T})} \gtrsim \|u\|_{L^p(\mathbb{T})} + \sum_{k=1}^{s} \|u\|_{\dot{L}^p_k(\mathbb{T})}$$

which proves the proposition. □

Intuition

These two propositions have shown us that definitions 3.1 and 3.2 of Sobolev spaces agree with the classical notion described by Evans involving weak derivatives. We therefore think of Sobolev spaces in precisely this manner. The advantage of this formulation, however, is that we have also defined negative and even fractional weak derivative spaces, the reason for which becomes clear on a deeper delve into the theory of distributions.

We end this subsection by introducing special notion for the case where $p = 2$:

$$H^s := L^p_s(\mathbb{T}), \qquad \dot{H}^s := \dot{L}^p_s(\mathbb{T}).$$

The space H^1 will be of particular interest in this project, as its Hilbert space structure is paramount in the forthcoming existence theorems.

3.2 Important Inequalities

Let's repurpose one of the most important inequalities in the study of Sobolev maps: the Sobolev embedding theorem. We will state this for the one-dimensional torus, but note that it applies to considerably more general settings.

Theorem 3.6 (Sobolev Embedding)**.** *Let u be a function on $\mathbb{T}$ with mean-zero. Suppose $s > 0$ and $1 < p < q < \infty$ satisfy*

$$s \geq \frac{1}{p} - \frac{1}{q}.$$

Then we have

$$\|u\|_{L^q(\mathbb{T})} \lesssim \|u\|_{\dot{L}^p_s(\mathbb{T})}.$$

A proof which emphasises the periodicity in this setting can be found in *The Sobolev inequality on the torus revisited* [3]. Let's make a few observations.

First, this result clearly implies $\|u\|_{L^q(\mathbb{T})} \lesssim \|u\|_{L^p_s(\mathbb{T})}$ by Proposition 3.5. Second, the assumption that u has mean-zero is characteristic of working on a compact manifold with no boundary. For a more complete view of Sobolev spaces on manifolds, see *Sobolev spaces on Riemannian manifolds* by Hebey [6].

We now emphasise a special case of Theorem 3.6 famously known as Poincaré's inequality.

Theorem 3.7 (Poincaré's inequality)**.** *For any $p \in (1, \infty)$, and every $u \in L^p_1(\mathbb{T})$,*

$$\|u - (u)\|_{L^p(\mathbb{T})} \lesssim \|\partial u\|_{L^p(\mathbb{T})}.$$

Here, (u) denotes the average value of u

$$(u) = \int_{\mathbb{T}} u \, \mathrm{d}x.$$

Let's demonstrate a straightforward proof for the case $p = 2$ in order to understand this result a little better. We can also see the power of using Fourier series to define derivatives.

Proof. Let $u \in H^1$. Notice that $\hat{u}(0) = \int_{\mathbb{T}} u \, \mathrm{d}x = (u)$, so

$$u(x) - (u) = \sum_{n \neq 0} \hat{u}(n) e^{2\pi i n x}.$$

We therefore compute

$$\begin{aligned} \|u - (u)\|_{L^2(\mathbb{T})}^2 &= \int_{\mathbb{T}} \left| \sum_{n \neq 0} \hat{u}(n) e^{2\pi i n x} \right|^2 \mathrm{d}x \\ &= \int_{\mathbb{T}} \left| \sum_{n \neq 0} \frac{1}{2\pi i n} \widehat{\partial u}(n) e^{2\pi i n x} \right|^2 \mathrm{d}x \\ &\leq \frac{1}{4\pi^2} \int_{\mathbb{T}} \left| \sum_{n \neq 0} i \frac{n}{|n|} \widehat{\partial u}(n) e^{2\pi i n x} \right|^2 \mathrm{d}x \\ &\lesssim \|\partial u\|_{L^2(\mathbb{T})}^2 \end{aligned}$$

□

Remark 3.1. We will also make use of this result later in the following form:

$$\|u\|_{L^2(\mathbb{T})}^2 \lesssim \|\partial u\|_{L^2(\mathbb{T})}^2 + \left(\int_{\mathbb{T}} u \, \mathrm{d}x \right)^2.$$

We have now completed the groundwork for discussing the existence of solutions to PDEs.

3.3 Uniformly Elliptic Equations

Before we get to the types of equation which will allow us to solve ($\star$), we first have to discuss elliptic problems. In Evans' *Partial Differential Equations*, elliptic problems with Dirichlet boundary conditions on bounded open subsets of $\mathbb{R}^d$ are defined as:

Definition 3.8 (Elliptic problem on $U \subset \mathbb{R}^d$). Let U be an open, bounded subset of $\mathbb{R}^d$. The elliptic problem on U is

$$\begin{cases} \mathcal{L}u = f & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases}$$

where $f : U \to \mathbb{R}$ is given and $u : \overline{U} \to \mathbb{R}$ is unknown. $\mathcal{L}$ denotes a second-order *uniformly elliptic* partial differential operator, which can take one of two forms.

- Divergence form:

$$\mathcal{L}u = -\sum_{i,j=1}^{d} \partial_j (a^{ij}(x) \partial_i u) + \sum_{i=1}^{d} b^i(x) \partial_i u + c(x) u$$

– Non-divergence form:

$$\mathcal{L}u = -\sum_{i,j=1}^{d} a^{ij}(x)\partial_{ij}u + \sum_{i=1}^{d} b^i(x)\partial_i u + c(x)u$$

where a^{ij}, b^i, c are given coefficient functions in $L^\infty(U)$.

Uniform ellipticity means there exists $\theta > 0$ such that

$$\sum_{i,j=1}^{d} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2$$

for almost all $x \in U$ and all $\xi \in \mathbb{R}^d$.

We make a similar definition on d-dimensional tori, although we must make a few modifications in order to navigate around a couple of intricacies. These are mainly avoided by restricting the operator $\mathcal{L}$ to just the highest order term in divergence form.

$$\mathcal{L}u = -\sum_{i,j=1}^{d} \partial_j(a^{ij}(x)\partial_i u) \tag{12}$$

This is the minimal case we need to consider to understand ($\star$). We can also spot that with $c \equiv 0$, solutions will only be defined up to a constant. In other words, solutions to this restricted elliptic problem are equivalence classes under $\overset{+}{\sim}$. We will therefore choose (without loss of generality) to represent each equivalence class by its unique mean-zero element. Furthermore, integrating over the whole space reveals the *compatibility condition*

$$\int_{\mathbb{T}^d} f \,\mathrm{d}x = 0, \tag{13}$$

motivating us to define

Definition 3.9 (Elliptic problem on $\mathbb{T}^d$)**.**

$$\mathcal{L}u = f \qquad \text{on } \mathbb{T}^d \tag{14}$$

where $f\colon \mathbb{T}^d \to \mathbb{R}$ is a given function satisfying the compatibility condition (13) and $u\colon \mathbb{T}^d \to \mathbb{R}$ is unknown. $\mathcal{L}$ denotes a second-order *uniformly elliptic* partial differential operator, with the form (12) and a^{ij} are given coefficient functions in $L^\infty(\mathbb{T}^d)$.

Let's address the fact that we have defined this problem on d-dimensional tori, rather than simply $\mathbb{T}$. This is to ensure we are actually solving a PDE.

If we reduce the above definition (in divergence form) to the case where $d = 1$, we get the ODE

$$-(a(x)u')' = f.$$

While more routine analysis of this problem may indeed be enough for this project, it would fail to capture the ideas needed later on in chapter §(??). Fortunately, the work of §3.1 - §3.2 naturally extends to any multidimensional torus, and we will proceed with this in mind.

The way forward now is to use energy estimates to show a unique solution of (14) exists. We will first weaken the notion of what it means to be a solution of (14), and employ powerful results from functional analysis to effortlessly conclude such solutions exist. We could then go about showing solutions have higher regularity under certain conditions.

What's the point of weakening?

Introducing weak derivatives and weak solutions might seem obscure, but the reason is clear once you realise that classical derivative spaces do not lend themselves to approximation. Convergence in the sup-norm is not an easy thing to prove, and spaces such as $C^2(\mathbb{T})$ are not complete with respect to more convenient norms such as L^p. To put it simply: completeness matters a lot. Sobolev spaces give us an excellent foundation in that they are complete with respect to norms we can actually use, and densely contain all smooth functions. This subsection aims to illustrate how quickly results drop out after such judicious preparation.

The motivation for the definition of weak solution comes from integration by parts. By this we mean:

Definition 3.10. The *Bilinear form* associated with the divergence form elliptic operator $\mathcal{L}$ is

$$B[u, v] := \int_{\mathbb{T}^d} \sum_{i,j=1}^{d} a^{ij}(x)\partial_i u \partial_j v \, \mathrm{d}x \tag{15}$$

for $u, v \in H^1(\mathbb{T}^d)$.

We then naturally rewrite (14) in the form

$$B[u, v] = (f, v)_{L^2(\mathbb{T}^d)} \tag{16}$$

with $f \in L^2(\mathbb{T}^d)$; a weak solution of (14) should satisfy (16) for every $v \in H^1(\mathbb{T}^d)$. However, we can actually go one step more general here. By considering the dual space of H^1, called H^{-1}, one can show that $H^1 \subset L^2 \subset H^{-1}$ and that elements of H^{-1} are all of the form

$$\langle f, v\rangle = \int f^0 v - \sum_{i=1}^{d} f^i \partial_i v \, \mathrm{d}x$$

(sometimes written $f = f^0 - \sum_{i=1}^{d} \partial_i f^i$) for some $f^0, f^1, \dots, f^d \in L^2$. And so it makes sense to define:

Does something need to be said about $f \in H^{-1}(\mathbb{T}^d)$ and the compatibility condition?

Definition 3.11. $u \in H^1(\mathbb{T}^d)$ is a weak solution of (14) if u is mean-zero and

$$B[u, v] = \langle f, v\rangle$$

for all $v \in H^1(\mathbb{T}^d)$, where $\langle f, v\rangle = \int_{\mathbb{T}^d} f^0 v - \sum_{i=1}^{d} f^i \partial_i v \, \mathrm{d}x$, and $\langle \, , \, \rangle$ is the pairing of $H^{-1}(\mathbb{T}^d)$ with $H^1(\mathbb{T}^d)$.

Note that requiring u to be mean-zero is not restrictive. Weak solutions to (14) are really equivalence classes in $\dot{H}^1(\mathbb{T}^d)$, and representatives of the same class may have different $H^1(\mathbb{T}^d)$ norms. Choosing the mean-zero representative naturally allows us to consider our solutions as being in $H^1(\mathbb{T}^d)$ through the Sobolev Embedding Theorem 3.6. We can now immediately state the abstract theorem which is the engine for elliptic existence theory.

Theorem 3.12 (Lax-Milgram)**.** *Let H be a real Hilbert space with H^* its dual. Let $\| \ \|$ denote the norm on H, $(\, , \,)$ the inner product, and $\langle \, , \, \rangle$ the pairing of H^* with H.*

Given a bounded, coercive *bilinear form $B : H \times H \to \mathbb{R}$, then for all $f \in H^*$, there exists a unique element $u \in H$ such that*

$$B[u, v] = \langle f, v\rangle$$

for all $v \in H$.

Coercivity means that there exists some $\beta > 0$ such that

$$\beta \|u\|^2 \leq B[u, u]$$

for every $u \in H$, which looks very similar to uniform ellipticity. In our case, $H = H^1(\mathbb{T}^d)$; a Hilbert space with the inner product

$$(u, v)_{H^1(\mathbb{T}^d)} = \int_{\mathbb{T}^d} uv + Du \cdot Dv \, \mathrm{d}x$$

($Du = (\partial_1 u, \cdots, \partial_d u)^T$). So it remains to show that the bilinear form B we defined in (15) is both bounded and coercive.

Theorem 3.13. *We have the following energy estimates on the bilinear form* (15).

(i) B is a bounded bilinear form; there exists $\alpha > 0$ such that

$$|B[u,v]| \leq \alpha \|u\|_{H^1(\mathbb{T}^d)} \|v\|_{H^1(\mathbb{T}^d)}$$

for all $u, v \in H^1(\mathbb{T}^d)$.

(ii) B is a coercive bilinear form; there exists a constant $\beta > 0$ such that

$$\beta \|u\|^2_{H^1(\mathbb{T}^d)} \leq B[u,u] \tag{17}$$

for all $u \in H^1(\mathbb{T}^d)$.

Proof. (i). It is easy to check

$$\begin{aligned} |B[u,v]| &\leq \sum_{i,j=1}^{d} \|a^{ij}\|_{L^\infty(\mathbb{T}^d)} \int_{\mathbb{T}^d} |Du||Dv| \,\mathrm{d}x \\ &\leq \alpha \|u\|_{L^2(\mathbb{T}^d)} \|v\|_{L^2(\mathbb{T}^d)} \\ &\leq \alpha \|u\|_{H^1(\mathbb{T}^d)} \|v\|_{H^1(\mathbb{T}^d)} \end{aligned}$$

for an appropriate constant α.

(ii). Invoking ellipticity with $\xi = Du$ and integrating over $\mathbb{T}^d$ obtains

$$\begin{aligned} \theta \int_{\mathbb{T}^d} |Du|^2 \,\mathrm{d}x &\leq \int_{\mathbb{T}^d} \sum_{i,j=1}^{d} a^{ij} \partial_i u \partial_j u \,\mathrm{d}x \\ &= B[u,u]. \end{aligned}$$

Using Poincaré's inequality 3.7 with the fact that u has mean-zero,

$$\beta \|u\|^2_{H^1(\mathbb{T}^d)} \leq \theta \|Du\|^2_{L^2(\mathbb{T}^d)} \leq B[u,u]$$

for an appropriate constant β. □

So, with an application of Lax-Milgram, we have shown:

Theorem 3.14 (Elliptic existence theorem)**.** *For every $f \in H^{-1}(\mathbb{T}^d)$ satisfying the compatibility condition* (13)*, there exists a unique, mean-zero weak solution $u \in H^1(\mathbb{T}^d)$ of the elliptic problem* (14).

3.4 Uniformly Parabolic Equations

The goal now is to understand PDEs involving evolution. That is, equations of the form

$$\partial_t u + \mathcal{L}u = f$$

with $u : [0,T] \times \mathbb{T}^d \to \mathbb{R}$. We are still following Evans [5] throughout this section.

A more sophisticated way to view this problem is with the unknown u being a *Banach space-valued function*, $\boldsymbol{u} : [0,T] \to H^1(\mathbb{T}^d)$. This requires formulating a new type of function space.

Definition 3.15. $L^p([0,T];X)$ denotes the vector space of strongly measurable[5] functions $\boldsymbol{u} : [0,T] \to X$ such that the norm $\|\boldsymbol{u}\|_{L^p(X)}$ is finite. These norms are defined as

$$\|\boldsymbol{u}\|_{L^p(X)} = \left(\int_0^T \|\boldsymbol{u}(t)\|_X^p \,\mathrm{d}t \right)^{\frac{1}{p}}$$

for $1 \leq p < \infty$, and

$$\|\boldsymbol{u}\|_{L^\infty(X)} = \operatorname*{esssup}_{[0,T]} \|\boldsymbol{u}(t)\|_X.$$

Definition 3.16. $\boldsymbol{v} \in L^1([0,T];X)$ is the weak derivative of $\boldsymbol{u} \in L^1([0,T];X)$ if

$$\int_0^T \phi'(t)\boldsymbol{u}(t) \,\mathrm{d}t = -\int_0^T \phi(t)\boldsymbol{v}(t) \,\mathrm{d}t$$

for all $\phi \in C_c^\infty(0,T)$. We denote $\boldsymbol{u}' = \boldsymbol{v}$.

In the same way as §3.3, the following definitions set up the parabolic problem we intend to study.

Definition 3.17 (Parabolic problem on $\mathbb{T}^d$)**.**

$$\begin{cases} \partial_t u + \mathcal{L}u &= f \quad \text{in } \mathbb{T}^d \times [0,T] \\ \qquad\quad u &= g \quad \text{on } \mathbb{T}^d \times \{0\}, \end{cases} \tag{18}$$

where $f : \mathbb{T}^d \times [0,T] \to \mathbb{R}$, $g : \mathbb{T}^d \to \mathbb{R}$ are given functions, and $u : \mathbb{T}^d \times [0,T] \to \mathbb{R}$ is the unknown. $\mathcal{L}$ denotes a second order *uniformly parabolic* partial differential operator of the form

$$\mathcal{L}u = -\sum_{i,j=1}^d \partial_j(a^{ij}(x,t)\partial_i u).$$

[5]Strongly measurable means that a function $\boldsymbol{f} : [0,T] \to X$ can be approximated by simple functions of the form $\boldsymbol{s} = \sum_1^m \chi_{E_i}(t)u_i$, where the E_i are measurable subsets of $[0,T]$, and $u_i \in X$.

Uniform parabolicity means there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{d} a^{ij}(x,t)\xi_i\xi_j \geq \theta|\xi|^2$$

for almost all $(x,t) \in \mathbb{T}^d \times [0,T]$ and all $\xi \in \mathbb{R}^d$. The coefficient functions a^{ij} are assumed to be in $L^\infty(\mathbb{T}^d \times [0,T])$.

Definition 3.18. The *time-dependent bilinear form* associated with parabolic operator $\mathcal{L}$ is

$$B[u,v;t] := \int_{\mathbb{T}^d} \sum_{i,j=1}^{d} a^{ij}(x,t)\partial_i u \partial_j v \, \mathrm{d}x$$

for $u, v \in H^1(\mathbb{T}^d)$ and almost all $t \in [0,T]$.

Viewing $\boldsymbol{u} : [0,T] \to H^1(\mathbb{T}^d)$ and $\boldsymbol{f} : [0,T] \to L^2(\mathbb{T}^d)$, we can see how to define weak solutions by multiplying (18) by $v \in H^1(\mathbb{T}^d)$ and integrating by parts

$$(\boldsymbol{u}', v) + B[\boldsymbol{u}, v; t] = (\boldsymbol{f}, v).$$

But we can also see from (18) that

$$\partial_t u = f + \sum_{j=1}^{d} \partial_j g^j$$

where $g^j = \sum_{i=1}^{d} a^{ij}\partial_i u \in L^2$. Therefore, we can more generally think of $\partial_t u \in H^{-1}(\mathbb{T}^d)$, and so we have the following definition for weak solutions to (18).

Definition 3.19.

$$\boldsymbol{u} \in L^2([0,T]; H^1(\mathbb{T}^d)) \quad \text{with} \quad \boldsymbol{u}' \in L^2([0,T]; H^{-1}(\mathbb{T}^d))$$

is a weak solution of (18) if

$$\langle \boldsymbol{u}', v \rangle + B[\boldsymbol{u}, v; t] = (\boldsymbol{f}, v)$$

for all $v \in H^1(\mathbb{T}^d)$ and almost all $t \in [0,T]$, and

$$\boldsymbol{u}(0) = g.$$

Note we do not identify $\boldsymbol{u}(t)$ with any mean-zero representative here, since the initial condition constrains which member of the equivalence class to choose. We can, however, choose g to be mean zero without any loss of generality, and we shall do so because it makes our lives easier.

The approach to solving this problem is central to this project. We will use an elegant technique first introduced in 1915 by Russian mathematician Boris Galerkin.

Key concept

Galerkin's technique is to build a solution to (18) with finite-dimensional approximations. We can alternatively think of this as projecting the problem onto a finite-dimensional subspace of H^1. The idea is the following:

1. Let $(w_k)_{k=1}^{\infty}$ be an orthonormal basis of L^2 and orthogonal in H^1.

2. Write $\boldsymbol{u}_m(t) = \sum_1^m d_m^k(t) w_k$, for appropriate coefficients d_m^k, such that $\boldsymbol{u}_m$ satisfies the projected problem.

3. Using an energy estimate on $\boldsymbol{u}_m$, take a limit as m, the number of dimensions, $\to \infty$. With any luck, $\lim_{m\to\infty} \boldsymbol{u}_m$ will be the unique solution to (18).

So, let's begin.

Proposition 3.20. *There exists an orthonormal basis, $(w_k)_{k=1}^{\infty}$, of $L^2(\mathbb{T}^d)$ which is orthogonal in $H^1(\mathbb{T}^d)$.*

Proof. We can choose $(w_k)_{k=1}^{\infty}$ to be an orthonormal basis of eigenfunctions for the Laplace operator, $-\Delta$. That is, $-\Delta w_k = \lambda_k w_k$ for eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$. See Evans [5, §6.5.1] for details on why such eigenfunction exist. Furthermore, periodicity on $\mathbb{T}^d$ means the w_k are necessarily mean-zero; a fact we will leverage in the following theorems.

Such eigenfunctions form an orthonormal basis of $L^2(\mathbb{T}^d)$, and lie in $H^1(\mathbb{T}^d)$. A short calculation shows they are indeed orthogonal in $H^1(\mathbb{T}^d)$. For any $v \in H^1(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d} Dw_k \cdot Dv \, \mathrm{d}x = \int_{\mathbb{T}^d} \lambda_k w_k \, \mathrm{d}x.$$

So

$$\begin{aligned}(w_k, w_l)_{H^1} &= \int_{\mathbb{T}^d} w_k w_l + Dw_k \cdot Dw_l \, \mathrm{d}x \\ &= (1 + \lambda_k) \int_{\mathbb{T}^d} w_k w_l \, \mathrm{d}x \\ &= 0.\end{aligned}$$

□

We are now looking to select coefficients d_m^k so that

$$\boldsymbol{u}_m(t) := \sum_1^m d_m^k(t) w_k \tag{19}$$

solves the projected problem

$$\begin{cases} (\boldsymbol{u}_m', w_k) + B[\boldsymbol{u}_m, w_k; t] &= (\boldsymbol{f}, w_k) \quad \text{for each } k = 1, \dots, m, \\ \boldsymbol{u}_m(0) &= g. \end{cases} \tag{20}$$

We claim such functions exist with the following theorem.

Theorem 3.21. *For each $m = 1, 2, \dots$ there exists a unique function $\boldsymbol{u}_m$ of the form (19) satisfying (20).*

Proof. The only key idea here is applying standard ODE existence theory to a linear system involving the d_m^k. For details, see Evans [5, §7.1.2]. □

Next up is the energy estimate

Theorem 3.22. *There exists a constant $C > 0$ depending only on T and the coefficients of $\mathcal{L}$ such that for each $m = 1, 2, \dots$*

$$\begin{aligned} \max_{[0,T]} \|\boldsymbol{u}_m(t)\|_{L^2(\mathbb{T}^d)} + \|\boldsymbol{u}_m\|_{L^2(H^1(\mathbb{T}^d))} + \|\boldsymbol{u}_m'\|_{L^2(H^{-1}(\mathbb{T}^d))} \\ \leq C(\|\boldsymbol{f}\|_{L^2(L^2(\mathbb{T}^d))} + \|g\|_{L^2(\mathbb{T}^d)}). \end{aligned}$$

Proof. The main requirement for this proof is making use of uniform parabolicity through the energy estimate (17) in the previous section. We are allowed to do this since the w_k, and consequently $\boldsymbol{u}_m$, are mean-zero. The rest of the proof only requires Gronwall's inequality and the boundedness of B. For details, see Evans [5, §7.1.2]. □

Finally, we show the approximate solutions we have constructed do indeed converge to a weak solution of (18).

Theorem 3.23 (Parabolic existence theorem)**.** *There exists a weak solution to* (18)

Proof. The energy estimate in the previous theorem has shown us that the sequence $\{\boldsymbol{u}_m\}_{m=1}^\infty$ is bounded in $L^2([0,T]; H^1(\mathbb{T}^d))$ and $\{\boldsymbol{u}_m'\}_{m=1}^\infty$ is bounded in $L^2([0,T]; H^{-1}(\mathbb{T}^d))$. We know from functional analysis that a bounded sequence in a Hilbert space contains a weakly convergent subsequence, so we can deduce the existence of a subsequence $\{\boldsymbol{u}_{m_l}\}_{l=1}^\infty$ such that

$$\begin{cases} \boldsymbol{u}_{m_l} \rightharpoonup \boldsymbol{u} & \text{weakly in } L^2([0,T]; H^1(\mathbb{T}^d)) \\ \boldsymbol{u}_{m_l}' \rightharpoonup \boldsymbol{u}' & \text{weakly in } L^2([0,T]; H^{-1}(\mathbb{T}^d)), \end{cases}$$

for a function $\boldsymbol{u} \in L^2([0,T]; H^1(\mathbb{T}^d))$ with $\boldsymbol{u}' \in L^2([0,T]; H^{-1}(\mathbb{T}^d))$.

We must then check if $\boldsymbol{u}$ is a weak solution. This is done in two steps:

1. Choose $\boldsymbol{v} \in C^1([0,T]; H^1(\mathbb{T}^d))$ with the form $\boldsymbol{v}(t) = \sum_1^N d^k(t) w_k$ for a fixed N and smooth functions d^k. As long as $m \geq N$, we can deduce from (20) that

$$\int_0^T \langle \boldsymbol{u}_m', \boldsymbol{v} \rangle + B[\boldsymbol{u}_m, \boldsymbol{v}; t] \, \mathrm{d}t = \int_0^T (\boldsymbol{f}, \boldsymbol{v}) \, \mathrm{d}t. \tag{21}$$

   passing this through our subsequence, we then have

$$\int_0^T \langle \boldsymbol{u}', \boldsymbol{v} \rangle + B[\boldsymbol{u}, \boldsymbol{v}; t] \, \mathrm{d}t = \int_0^T (\boldsymbol{f}, \boldsymbol{v}) \, \mathrm{d}t. \tag{22}$$

2. We conclude the above equality also holds for any $\boldsymbol{v} \in L^2([0,T]; H^1(\mathbb{T}^d))$ since functions of the form in step 1 are dense in this space. So we have

$$\langle \boldsymbol{u}', v \rangle + B[\boldsymbol{u}, v; t^*] = (\boldsymbol{f}, v)$$

   for all $v \in H^1(\mathbb{T}^d)$ and almost all $t^* \in [0,T]$. This can be seen through a short localisation argument whilst remembering that the bilinear form B is only defined almost everywhere.

To check if $\boldsymbol{u}(0) = g$, we use both of the equalities (21), (22) for $\boldsymbol{v} \in C^1([0,T]; H^1(\mathbb{T}^d))$ with $\boldsymbol{v}(T) = 0$. Integrating by parts[6] and taking the weak limit in (21) we find

$$(\boldsymbol{u}(0), \boldsymbol{v}(0)) = (g, \boldsymbol{v}(0)).$$

As $\boldsymbol{v}(0)$ is arbitrary, $\boldsymbol{u}(0) = g$, which completes the proof. □

The uniqueness of this solution follows again from the energy estimate (17) using parabolicity.

Let's wrap up this section by commenting on the regularity of the solutions we have constructed. In general, solutions will have up to two more weak derivatives than the given function f and up to one more than the initial condition g. High enough order derivatives can even be used to conclude strong differentiability via certain embedding theorems such as *Morrey's inequality* [5, §5.6.2]. We will not go into regularity theory in this report, but we will hint at its uses in the next chapter.

[6] This is integration by parts in Sobolev-Bochner spaces.

4 Non-linear Existence Theory

5 Applications and Extensions

Some citations: [1].

References

[1] Aluminum in Space: Pushing the Limits of Aerospace Technology, 2024. URL https://elkamehr.com/en/aluminum-in-space-pushing-the-limits-of-aerospace-technology/. Accessed on 27th January 2025.

[2] P. M. Anderson, J. P. Hirth, and J. Lothe. *Theory of dislocations*. Cambridge University Press, 2017.

[3] Á. Bényi and T. Oh. The sobolev inequality on the torus revisited. *Publicationes Mathematicae Debrecen*, 83(3):359, 2013.

[4] L. Carleson. On convergence and growth of partial sums of fourier series. 1966.

[5] L. C. Evans. *Partial differential equations*, volume 19. American Mathematical Society, 2 edition, 2022.

[6] E. Hebey. *Sobolev spaces on Riemannian manifolds*, volume 1635. Springer Science & Business Media, 1996.

[7] T. Hudson, F. Rindler, and J. Rydell. A quantitative model for the frank-read dislocation source based on pinned mean curvature flow. 2024. https://arxiv.org/abs/2409.20294.

[8] D. Hull and D. J. Bacon. *Introduction to dislocations*, volume 37. 5 edition, 2011.

[9] R. A. Hunt. On the convergence of fourier series, orthogonal expansions and their continuous analogues (proc. conf., edwardsville, ill., 1967). *Illinois. Univ. Press Carbondale, Ill*, pages 235–255, 1968.

[10] M. Peach and J. Koehler. The forces exerted on dislocations and the stress fields produced by them. *Physical Review*, 80(3):436, 1950.