Chapter 4

Derivation of a Least Squares Finite Element Space - Time Discretisation

4.1 Construction of an Equivalent Minimisation Problem

write somehwere explicitly $div(\sigma)$ is only space derivative
In this chapter we will tie the beforementioned concepts together to derive a discretised problem formulation that can subsequently be turned into an algorithm capable of approximating reaction diffusion equations. In order to do so let us firstly set the ground for the overall framework we are looking at. We consider a space-time domain

$$\Omega = \mathcal{S} \times \mathcal{T}, \quad \mathcal{T} = (0, T),\ T > 0 \text{ and } \mathcal{S} \subset \mathbb{R}^N \tag{4.1}$$

where $\mathcal{T}$ represents the time domain and $\mathcal{S}$ is the domain in space which we require to be Lipschitz regular. We may have a mixture of Dirichlet and Neumann boundary conditions on the boundary of Ω which we will denote by Γ and are labeled as $\Gamma_D, \Gamma_N \subset \Gamma$ respectively. We further assume them to be such that the problem is well posed. The class of partial differential equations introduced in the prologue that we would like to solve for then reads as the following:

$$\begin{aligned} u_t - div(D(x)\nabla u) &= f(u) && (x,t) \in \Omega \\ u &= g_D && (x,t) \in \Gamma_D \\ \nabla u \cdot n &= g_N && (x,t) \in \Gamma_N \end{aligned} \tag{4.2}$$

It describes a parabolic partial differential equation with a non-linear right hand side. Typically we will have that $u(x, 0) = u_0 = g_D$ for all $x \in \mathcal{S}$ and Neumann boundary conditions on the boundary of $\mathcal{S}$ for $t \in (0, T)$.

The next step will be to derive an equivalent optimisation problem whose solution therefore then coincides with the solution of (4.2) at least in a weak sense. Since we are entirely working with finding solutions in Sobolev spaces in the least squares setting we can generally only require equivalence to a primal weak formulation of (4.3) and (4.2) or equivalence with respect to the solution space of the variational formulation, for a further discussion we refer to [source]. Generally when working with least squares finite element formulations choosing a suitable solution space U and data space Z is often non trivial as there are a number of difficulties that can arise. One usually faces a trade off between constructing a mathematically well-defined problem and allowing for a relatively simple, efficient, robust while still accurate implementation.Therefore to make the methodology of LSFEMs competitive compared to other approaches like Galerkin approximations further considerations need to be taken into account. We saw in the previous

section that it is possible to derive least squares formulations that recover the properties of the Rayleigh- Ritz setting, however one hindrance one encounters in this setting as well as in many others is the higher order operator arising in (3.25), that would require a function space X of higher regularity. When considering a simple Poisson equation with Dirichlet boundary conditions this would for example imply that we would require the solution u to be from H_0^2, instead of H_0^1 [7]. This does not only heavily limit the set of admissible solutions but it is additionally much harder to construct appropriate finite dimensional subspaces for and therefore impractical to use. In order to succomb this obstacle we will recast (4.2) as a system of coupled of equations only containing first order derivatives to apply the methodologies introduced in section (3.4) at the price of introducing an additional variable.

$$\begin{aligned} u_t - div(\sigma) &= f(u) && (x,t) \in \Omega \\ \sigma &= D(x)\nabla u && (x,t) \in \Omega \\ u &= g_D && (x,t) \in \Gamma_D \\ \nabla u \cdot n &= g_N && (x,t) \in \Gamma_N \end{aligned} \tag{4.3}$$

Rearranging this equation onto a vector form we obtain in Ω

$$\begin{pmatrix} I & -D(x)\nabla \\ -div & \frac{\partial}{\partial t} \end{pmatrix} \begin{pmatrix} \sigma \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ f(u) \end{pmatrix} \tag{4.4}$$

which can be shortened to $\mathcal{A}([\sigma, u]) = \tilde{f}(u)$ and where $\mathcal{A}$ denotes the differential operator and $\tilde{f}(u)$ the right hand side.

Show that this operator is linear and bounded? but then norm ... spaces ... bla bla bla ...

Remark. It is not clear yet though in which space we will be looking for a solution. In general when working with strong and weak formulations there is a a number of function spaces involved, problem formulations that are similar but not the same and it is difficult to represent them in clear but short notation.This can easily lead to confusion when one is not very familiar with this difficulty. In the current case the situation is even more complicated since we are trying to derive a weak formulation in a least squares setting, additionally there obviously exist the regular or so-called primal weak formulations of (4.2) which we will sometimes refer to. Therefore we will first derive a least functional J with a corresponding variational formulation, postponing the definition of appropriate spaces but simply assuming them to be given for now.

So instead of looking for a solution of the strong formulation, let us now turn to the derivation of an optimisation problem that we would like to satisfy in a weak sense. We want a solution of (4.3) to be a global minimum of the optimisation problem, independently of the choice of the spaces that we are using and its associated norm, hence we additionally want the least squares functional J to be zero for the solution of (4.3). On the other hand if $J(\sigma, u) = 0$ then the orginial problem also has to be satisfied if only in a weak sense with respect to the spaces U and Z. All three properties hold for the following functional

$$\tilde{J}(\sigma, u) = ||u_t - div(\sigma) - f(u)||_Z^2 + ||\sigma - D(x)\nabla u||_Z^2 \tag{4.5}$$

Additionally it has shown to be practical to be able to weight the two terms with coefficients [source] which does not affect the overall solution but grants us the possibility to numerically give more importance to one term than the other. The additional scalars of 0.5 are introduced to simplify further computations of the derivatives as we will see in the following section. Therefore the the minimisation problem then reads

$$\min_{(\sigma,u)\in U} J(\sigma, u) = \frac{1}{2}c_1||u_t - div(\sigma) - f(u)||_Z^2 + \frac{1}{2}c_2||\sigma - D(x)\nabla u||_Z^2. \tag{4.6}$$

If we now consider the functional for $f = 0$ we obtain

$$J([\sigma, u], 0) = \langle u_t - div(\sigma), u_t - div(\sigma)\rangle_Z + \langle \sigma - D(x)\nabla u, \sigma - D(x)\nabla u\rangle_Z. \tag{4.7}$$

which more specifically gives rise to the following bilinear form

$$\mathcal{B}([\sigma, u], [\tau, v]) = \langle \begin{pmatrix} I & -D\nabla \\ -div & \frac{\partial}{\partial t} \end{pmatrix} \begin{pmatrix} \sigma \\ u \end{pmatrix}, \begin{pmatrix} I & -D\nabla \\ -div & \frac{\partial}{\partial t} \end{pmatrix} \begin{pmatrix} \tau \\ v \end{pmatrix} \rangle_Z \tag{4.8}$$

The subsequent candidate for a variational formulation can be derived as usual [source] that is as the *sum* of the directional derivatives of J, where the *directions* become the testfunctions.

$$\begin{aligned} \text{Find } (\sigma, u) \in U \text{ such that } \mathcal{B}([\sigma, u], [\tau, v]) = L_{\tilde{f}}(u)(v) \quad \forall (\tau, v) \in W \\ \text{with } \mathcal{B}([\sigma, u], [\tau, v]) = (\mathcal{A}[\sigma, u]^T, \mathcal{A}[\tau, v]^T) \text{ and } \mathcal{L}_{\tilde{f}}(u)([\tau, v]) = (\mathcal{A}[\tau, v], \tilde{f}(u))_Z \end{aligned} \tag{4.9}$$

Remark: $L_f(u)(v)$ is a linear operator in v but non-linear in u. What remains to show now is the declaration of the function spaces involved U, W, Z that potentially also leads us to a norm equivalence in order for *Theorem 3* to hold. As mentioned previously this is not necessearily straightforward and will be the matter of discussion in the subsequent sections.

4.2 Function Spaces

In order to define U and Z let us consider the optimisation problem (4.9) and its variational formulation (4.12) again. We would like to allow for the broadest class of solutions possible while still ensuring that all terms are well-defined, while staying away from Sobolev spaces of negative or fractional powers due to the beforementioned practicality reasons. Additionally to guarantee the existence of all terms involved we have to make sure that the present weak derivatives of σ and u exist in the induced inner product of Z. Let us therefore consider the following spaces

$$H^1_{div}(\mathcal{S}) = \{\sigma \in (L^2(\mathcal{S}))^n : div(\sigma) \in L^2(\mathcal{S})\} \tag{4.10}$$

$$H^1_{DIV}(\Omega) = H^1_{div}(\mathcal{S}) \times L_2(\mathcal{T}) \tag{4.11}$$

where and when to use $\times$, overall notation

$$U = H^1_{DIV}(\Omega) \times H^1(\Omega) \tag{4.12}$$

$$Z = L^2(\Omega) \tag{4.13}$$

that is we require that

$$\sigma \in H^1_{DIV}(\Omega) \text{ and } u \in H^1(\Omega) \tag{4.14}$$

we can then check that all above terms in $J([\sigma, u], 0)$ are well defined. If we additionally assume $f(u) \in L^2(\Omega)$ for all u, the problem remains to be well-posed. Theoretically one could potentially only require $f \in H^{-1}(\Omega)$ but for the scope of this thesis, we will restrict ourselves to the former. Don't we need derivatives of f later? Higher regularity to solve?
$\mathcal{B}$ induces an inner product on U and subsequently a norm. It is clearly a symmetric bilinear form. We also have that $\mathcal{B}([\sigma, u], [\sigma, u]) \geq 0$ for all $[\sigma, u] \in U$, it only remains to show that $\mathcal{B}([\sigma, u], [\sigma, u]) = 0 \Rightarrow [\sigma, u] = 0$.

in previous section would have the bounds from X, how to do this here? how does this even

make sense, hilbert space, strong form?

Another point that has not been a mentioned so far but will have to be taken into account is the way of how to treat the boundary conditions in least-squares formulations. One possibility is to also include them in the functional as an additional term while another one would be to directly include them in the discretised system of the space. The former one entails the additional definition of an appropriate norm on the boundary while also requiring the treatment of the additional term. Therefore we will assume here that the boundary is sufficiently regular and that the appropriate conditions can directly be imposed as part of the discretised system which will be discussed in more detail in the implementation section.

4.3 Norm Equivalence

$\mathcal{B}([\sigma, u], [\tau, v])$ induces a norm on U, that is $\mathcal{B}$ defines an inner product. It is clearly a symmetric bilinear form.
Show positive definiteness :
We clearly have $\mathcal{B}([\sigma, u], [\sigma, u]) \geq 0$ for all $(\sigma, u) \in U$ since $J([\sigma, u], 0)$ is the sum of two squared L_2-norms. It remains to show that

$$\mathcal{B}([\sigma, u], [\sigma, u]) = 0 \iff (\sigma, u) = 0 \tag{4.15}$$

Can't be right, the point is that J is zero at the solution, so unless the solution is zero ...

4.4 A Finite Element Space-Time Formulation

Now let us turn towards deriving a finite element discretisation of the problem. We want to consider conforming subspaces of U and W, hence we will be considering finite dimensional spaces $U_h \subset U$ and $W_h \subset W$ that are also defined on the entire space-time domain.

We will choose $U_h = W_h$, that is the test and solution space are the same using a Galerkin approach. U_h contains the solution space for σ and u. They can be chosen independently from each other which can potentially be advantageous due to the different occurences of their partial derivatives. Hence let us assume that $\sigma \in \tilde{V}_h$ and $u \in V_h$, where $U_h = \tilde{V}_h \times V_h$.

Suppose we have $\tilde{n} = dim(\tilde{V}_h)$ and let $\{\tilde{\phi}_1, ..., \tilde{\phi}_{\tilde{n}}\}$ be a basis of $\tilde{V}_h$ and similarly $n = dim(V_h)$ and $span\{\phi_1, ..., \phi_n\} = V_h$. Since we are in a space-time setting we have $\tilde{\phi} = \tilde{\phi}(x, t)$ and $\phi = \phi(x, t)$.

We then represent σ and u as a linear combination of basis functions in U_h, that is

$$\sigma_h(x, t) = \sum_{i=1}^{\tilde{n}} \sigma_i \tilde{\phi}_i(x, t) \qquad u_h(x, t) = \sum_{i=1}^{n} u_i \phi_i(x, t) \tag{4.16}$$

The functional J then looks as follows

$$J(\sigma_h, u_h) = || \sum_{i=1}^{n} u_i \ (\phi_i)_t - \sum_{i=1}^{\tilde{n}} \sigma_i \ div(\tilde{\phi}_i) - f(\sum_{i=1}^{n} u_i \phi_i)||_Z^2 + || \sum_{i=1}^{\tilde{n}} \sigma_i \tilde{\phi}_i - D(x) \nabla (\sum_{i=1}^{n} u_i \phi_i)||_Z^2 \tag{4.17}$$

which will give rise to a vector-valued system once we discretise the domain.

Let us now consider the arising variational formulation that we will attempt to solve. That is we introduce a set of test functions consisting of the basis vectors of $\tilde{V}_h$ and V_h. *general enough? or separate set?* That is the discretised problem reads

$$\text{Find } (\sigma_h, u_h) \in U_h \text{ such that } B([\sigma_h, u_h], [\tau_h, v_h]) = L_{\tilde{f}}(u_h)(v_h) \quad \forall (\tau_h, v_h) \in W_h \tag{4.18}$$

where $B \in \mathbb{R}^{m \times m}$, where $m = \tilde{n}+n$, be the matrix arising from the discretised bilinear operator, then we obtain for the i-th row and j-th column
how do we construct this thing
block structure,

$$B = \begin{bmatrix} B_{\sigma\sigma} & B_{\sigma u} \\ B_{u\sigma} & B_{uu} \end{bmatrix} \qquad L_{\tilde{f}}(u_h) = \begin{bmatrix} (L_{\tilde{f}}(u_h))_\sigma \\ (L_{\tilde{f}}(u_h))_u \end{bmatrix} \tag{4.19}$$

where $B_{\sigma\sigma}$ corresponds to the indices $i, j \in \{1, ..., \tilde{n}\}$ for which we have

$$\text{For } B_{\sigma\sigma}: \qquad B_{ij} = \langle \tilde{\phi}_j, \tilde{\phi}_i \rangle_Z + \langle div(\tilde{\phi}_j), div(\tilde{\phi}_i) \rangle_Z \qquad \forall i, j \in \{1, ..., \tilde{n}\} \tag{4.20}$$

$$\text{For } B_{\sigma u}: \qquad B_{ij} = -\langle D(x)\nabla\phi_j, \tilde{\phi}_i \rangle_Z - \langle (\phi_j)_t, div(\tilde{\phi}_i) \rangle_Z \qquad \forall i \in \{1, ..., \tilde{n}\}, j \in \{\tilde{n}+1, ..., m\} \tag{4.21}$$

$$\text{For } B_{u\sigma}: \qquad B_{ij} = -\langle D(x)\nabla\phi_i, \tilde{\phi}_j \rangle_Z - \langle (\phi_i)_t, div(\tilde{\phi}_j) \rangle_Z \qquad i \in \{\tilde{n}+1, ..., m\}, \forall j \in \{1, ..., \tilde{n}\} \tag{4.22}$$

$$\text{For } B_{uu}: \qquad B_{ij} = \langle D(x)\nabla\phi_j, D(x)\nabla\phi_i \rangle_Z + \langle (\phi_j)_t, (\phi_i)_t \rangle_Z \qquad \forall i, j \in \{\tilde{n}+1, ..., m\} \tag{4.23}$$

In the case that f is independent of u, that is $f = 0$ or $f = f(x,t)$, the derivative of f with respect to u is clearly zero and therefore the right-hand side, which we will denote by $L_{\tilde{f}} = L_{\tilde{f}}(u_h)$ to underline its independence of u only contains the following terms

$$(L_{\tilde{f}})_i = 0 \qquad i \in \{1, ..., \tilde{n}\} \tag{4.24}$$

$$(L_{\tilde{f}})_i = \langle div(\tilde{\phi}_i) - (\phi_i)_t, f \rangle_Z \qquad i \in \{\tilde{n}+1, ..., m\} \tag{4.25}$$

These terms we can then all compute directly and assemble to one large linear system of equations

$$B \begin{pmatrix} \sigma \\ u \end{pmatrix} = L_{\tilde{f}} \tag{4.26}$$

and can then be solved immediately for $[\sigma, u]^T$ using for example a multigrid method. In the case of $f = f(u)$ the situation is more complicated, since we also have to take the derivatives of f with respect to u into account, and then construct a nonlinear iteration scheme, to solve for u unless we have that $f(u)$ is a linear function. In which case we would obtain additional terms that can be added to B_{uu}.

4.5 Iteration Scheme

As discussed in the previous chapter [see section 3.4] the general approach to finding a minimiser of the functional J, is to search for a tuple $[\sigma, u]$ for which $\nabla J([\sigma, u]) = 0$, and $\nabla^2 J([\sigma, u])$ is positive definite. We start with an initial guess $s_0 = s_{\text{init}} = [\sigma_{\text{init}}, u_{\text{init}}]$ and then successively try to decrease energy. In case of a Newton step by finding a quadratic approximation of J at the value of the current iterate s_k, where the updated solution s_{k+1} is the minimiser of the quadratic approximation. However if J is not convex, and hence $\nabla^2 J([\sigma, u])$ positive definite, the extremum of the quadratic approximation actually gives us a maximum or in the numerically

very unlikely indefinite case, no change at all. Therefore if we want to use a Newton iteration and have a decrease of energy in every step, that is $J(s_0) \geq J(s_1) \geq ... \geq J(s_k) \geq ...$ we need to be checking for convexity. The second option that was discussed previously to obtain a reduction in energy is to use gradient descent method where we simply take an iteration step into the steepest descent direction. In order to perform a Newton step we have to compute the Hessian of J, as well as its gradient in each iterate. For a steepest descent method we only need to evaluate the gradient, which is computationally much cheaper but only leads to linear convergence rates, instead of quadratic ones in the convex Newton case.

There are two ways to derive a discretised formulation in f. We can either first derive all terms using the continuous formulation and then discretise. Or we can linearise the problem first and then differentiate the discretised problem.
There has been discussion bla bla bla. No better or worse.
Here we will assume that we discretise the problem first which is also how we implemented the problem. In this case a Newton step would like as following, where $s_k = [\sigma_k, u_k]$ is the We want $\nabla J = 0$
discussion different derivatives
We have the functions $f, f', f'' : U \to$?? with $u \mapsto f(u), f'(u), f''(u)$ that we can approximate in the same way as the solution u. In order for the derivatives of f to exist, even if it is only in a weak sense we *officially* need f to be in H^2 in what? in U?

$$(f_h)_i = \sum_{i=1}^{n} f_i \phi_i, \qquad (f_h')_i = \sum_{i=1}^{n} f_i' \phi_i, \qquad (f_h'')_i = \sum_{i=1}^{n} f_i'' \phi_i \tag{4.27}$$

these denote piecewise linear functions approximating the continuous functions, whereas other one simply coefficient vector where in this case we already know the coefficients as they are determined pointwise through the current solution u_h, that is we have $f_h = f(u_h), f_h' = f'(u_h), f_h'' = f''(u_h)$. These are the piecewise linear functions, hence inner products well defined. i-th entry

$$(L_{\tilde{f}}(u_h))_\sigma = \langle div(\sigma_h), f_h' \rangle_Z \tag{4.28}$$

$$(L_{\tilde{f}}(u_h))_u = -\langle u_t, f'(u) \rangle_Z - \langle v_t, f(u) \rangle_Z + \langle f(u), f'(u) \rangle_Z \tag{4.29}$$

$$\nabla J_k = \nabla J(s_k) = B \begin{pmatrix} \sigma_k \\ u_k \end{pmatrix} - L_{\tilde{f}}(u_k) \tag{4.30}$$

For one step of gradient descent we therefore compute the update by

$$s_{k+1} = s_k + \alpha(-\nabla J_k) \tag{4.31}$$

where $\alpha > 0$ is a scaling parameter that can be chosen in different ways, for example a line search algorithm and which will be discussed in more detail in the implementation chapter.

From the gradient and previous section we can determine the discretised Hessian which is needed for a Newton step. We obtain

$$H_k = \nabla^2 J(s_k) = B+ \tag{4.32}$$

$$s_{k+1} = s_k - H_k^{-1}(\nabla J_k) \tag{4.33}$$

that is we would like to solve the linear system of equations

$$e_{k+1} = -H_k^{-1}(\nabla J_k) \quad \text{with} \quad s_{k+1} = s_k + e_{k+1} \tag{4.34}$$

But since it is generally expensive to compute the inverse of a large matrix, even if H_k is sparse and symmetric, because this property does in general not translate to the inverse we apply a multigrid method to solve (4.33). That is we need to solve a linear system of equations in each Newton step.
put derivative to zero therefore if we assume local convexity must be an extremum, positive definite, means minimum, therefore variational formulation makes sense

As mentioned in the prologue we consider an iterative approach to solve this problem and will therefore consider linearisations of the problem. But before going into more detail about that, let us consider the variational formulation of the coupled reaction-diffusion system (4.3)
A solution to (4.3) will always be a minimiser of the functional J independent on the choice of Z, however the sequence of iterates is dependent on the choice of the norm.
QUESTION OF WHERE TO LINEARISE ...
So then we do actually end up with a system where we have, for $x = (\sigma, u)$, $w = (\tau, v)$: $\mathcal{A}(x, w) = \mathcal{F}_k(w)$ variational equation in each step?!
"In particular, for linear PDEs, residual minimization can lead to unconstrained optimization problems for convex quadratic functionals even if the original equations were not at all associated with optimiza- tion. If the PDE problem is nonlinear, then properly executed residual minimization leads to unconstrained minimization problems whose linearization22 gives rise to unconstrained minimization problems with convex quadratic functionals." (LSFEM book p.50)
Hence we have that a minimiser to the above formulation is at the same time a solution to our original problem. One can easily see that we have $J(\sigma, u) \geq 0$ for all $(\sigma, u) \in X_1 \times X_2$. Hence if $J(\sigma, u) = 0$ we must be at a minimum. The general strategy for solving these type of optimisation problems is more broad though, the idea is to find a pair $(\sigma, u) \in X_1 \times X_2$ for which $\nabla J(\sigma, u) = 0$ and $\nabla^2 J(\sigma, u)$ is positive definite which must consequently mean that (σ, u) is a minimiser.
Something like: We have seen above that every solution to the original problem is a solution to the minimisation problem and therefore if the solution to the original problem exists and is unique this one must be as well. Subsequently we proceed by determining the gradient and hessian of J and as we will see later are at the same time deriving a weak formulation that we will attempt to solve.
In the following I will denote $||\cdot||_{L^2(\Omega\times(0,T))}$ by $||\cdot||_2$ for brevity. choice of norms, why is this equivalent

4.6 Derivation of the Derivatives

In this section we will derive the first and second order directional partial derivatives of J with respect to its two variables in order to then set them to zero. The problem at hand is more complicated than in the standard case because we also have to take into account that the right hand side f is dependent on u and will therefore also appear in the differentiation. We can generally assume that the first and second order derivatives of J exist and are continuous almost everywhere since $||\cdot||_2$ is "continuouly differentiable" and $\sigma \in H^1_{div}(\Omega)$ and $u \in H^1(\Omega)$.
Something no second order derivatives. In the following we will be determining the Gateaux derivatives of J, where we will be splitting the functional in three different terms that will be considered separately for greater clarity dividing it into the linear and nonlinear terms which is possible due to the linearity of the inner product.
The following notation will be used subsequently $||x||_2^2 = \langle x, x\rangle$ and $x = (\sigma, u)$, $h = (\tau, v)$,

$k = (\rho, w)$. So let us consider the following

$$\begin{aligned} J_1(\sigma, u) &= \frac{1}{2} c_1 \langle u_t - div(\sigma), u_t - div(\sigma) \rangle \\ J_2(\sigma, u) &= \frac{1}{2} c_1 \langle 2u_t - 2div(\sigma) - f(u), -f(u) \rangle \\ J_3(\sigma, u) &= \frac{1}{2} c_2 \langle \sigma - \beta \nabla u, \sigma - \beta \nabla u \rangle \end{aligned} \tag{4.35}$$

where we can see that $J(\sigma, u) = J_1(\sigma, u) + J_2(\sigma, u) + J_3(\sigma, u)$. Now taking the partial directional derivatives we obtain, again taking the linearity of the inner product as well as its symmetry into account that

$$\begin{aligned} \frac{\partial J_1}{\partial \sigma} &= \lim_{\epsilon \to 0} \frac{J_1(\sigma + \epsilon\tau, u) - J_1(\sigma, u)}{\epsilon} \\ &= \lim_{\epsilon \to 0} \frac{c_1}{2\epsilon} (\langle u_t - div(\sigma + \epsilon\tau), u_t - div(\sigma + \epsilon\tau) \rangle - \langle u_t - div(\sigma), u_t - div(\sigma) \rangle) \\ &= \lim_{\epsilon \to 0} \frac{c_1}{2\epsilon} (\langle u_t, u_t \rangle - \langle u_t, div(\sigma) \rangle - \epsilon \langle u_t, div(\tau) \rangle - \langle div(\sigma), u_t \rangle + \langle div(\sigma), div(\sigma) \rangle \\ &+ \epsilon \langle div(\sigma), div(\tau) \rangle - \epsilon \langle div(\tau), u_t \rangle + \epsilon \langle div(\tau), div(\sigma) \rangle + \epsilon^2 \langle div(\tau), div(\tau) \rangle \\ &- \langle u_t, u_t \rangle + \langle u_t, div(\sigma) \rangle + \langle div(\sigma), u_t \rangle - \langle div(\sigma), div(\sigma) \rangle) \\ &= \lim_{\epsilon \to 0} \frac{c_1}{2\epsilon} (-2\epsilon \langle u_t, div(\tau) \rangle + 2\epsilon \langle div(\sigma), div(\tau) \rangle + \epsilon^2 \langle div(\tau), div(\tau) \rangle) \\ &= -c_1 \langle u_t, div(\tau) \rangle + c_1 \langle div(\sigma), div(\tau) \rangle \end{aligned} \tag{4.36}$$

We can see that the terms only containing σ or u cancel. We end up with a number of mixed terms as well as the terms containing purely τ and v. Due to the factor of $\frac{1}{2}$ in front of the inner products in J and the symmetry of the inner product, the mixed terms add up 1 or -1 respectively. Again because of the bilinearity of the inner product we can write ϵ in front of the individual terms, often they will cancel with the factor of $\frac{1}{\epsilon}$ in front. If we now take the limit with respect to ϵ going to zero all terms with an ϵ in both arguments will tend to zero which gives us the remaining result. By proceeding analogously for equation J_2 and J_3 we obtain in these cases:

$$\begin{aligned} \frac{\partial J_2}{\partial \sigma} &= c_1 \langle div(\tau), f(u) \rangle \\ \frac{\partial J_3}{\partial \sigma} &= c_2 \langle \sigma, \tau \rangle - c_2 \beta \langle \tau, \nabla u \rangle \end{aligned} \tag{4.37}$$

Let us now turn to the partial derivatives with respect to u. Here we obtain the following for J_1 and J_3:

$$\begin{aligned} \frac{\partial J_1}{\partial u} &= \lim_{\epsilon \to 0} \frac{J_1(\sigma, u + \epsilon v) - J_1(\sigma, u)}{\epsilon} \\ &= \lim_{\epsilon \to 0} \frac{c_1}{2\epsilon} (\langle (u + \epsilon v)_t - div(\sigma), (u + \epsilon v)_t - div(\sigma) \rangle - \langle u_t - div(\sigma), u_t - div(\sigma) \rangle) \\ &= c_1 \langle u_t, v_t \rangle - c_1 \langle v_t, div(\sigma) \rangle \end{aligned} \tag{4.38}$$

$$\frac{\partial J_3}{\partial u} = -c_2 \beta \langle \sigma, \nabla v \rangle + c_2 \beta^2 \langle \nabla u, \nabla v \rangle$$

In the case of J_2, we have to take the non-linearity of f into account. If we assume that f sufficiently smooth (what do we need exactly?!) that is $\lim_{\epsilon\to 0} f(u+\epsilon v) = f(u)$ and $\lim_{\epsilon\to 0}\langle f(u+\epsilon v), f(u+\epsilon v)\rangle - \langle f(u), f(u)\rangle = \langle f'(u)\cdot v, f(u)\rangle + \langle f(u), f'(u)\cdot v\rangle$ which can be added due to symmetry.

$$\begin{aligned}
\frac{\partial J_2}{\partial u} &= \lim_{\epsilon\to 0}\frac{J_2(\sigma, u+\epsilon v) - J_2(\sigma, u)}{\epsilon} \\
&= \lim_{\epsilon\to 0}\frac{c_1}{2\epsilon}(\langle 2(u+\epsilon v)_t - 2div(\sigma) - f(u+\epsilon v), -f(u+\epsilon v)\rangle - \langle 2u_t - 2div(\sigma) - f(u), -f(u)\rangle) \\
&= \lim_{\epsilon\to 0}\frac{c_1}{2\epsilon}(-2\langle u_t, f(u+\epsilon v)\rangle + 2\langle u_t, f(u)\rangle \\
&- 2\epsilon\langle v_t, f(u+\epsilon v)\rangle \\
&+ 2\langle div(\sigma), f(u+\epsilon v\rangle - 2\langle div(\sigma), f(u)\rangle \\
&+ \langle f(u+\epsilon v), f(u+\epsilon v)\rangle - \langle f(u), f(u)\rangle) \\
&= -c_1\langle u_t, f'(u)\cdot v\rangle - c_1\langle v_t, f(u)\rangle + c_1\langle div(\sigma), f'(u)\cdot v\rangle + c_1\langle f(u), f'(u)\cdot v\rangle
\end{aligned} \tag{4.39}$$

Hence we obtain the following partial first order directional derivatives.

$$J_\sigma[\tau] = \frac{\partial}{\partial\sigma}J(\sigma,u)[\tau] = c_2\langle\sigma,\tau\rangle + c_1\langle div(\sigma), div(\tau)\rangle - c_2\beta\langle\nabla u,\tau\rangle - c_1\langle u_t, div(\tau)\rangle - c_1\langle f(u), div(\tau)\rangle \tag{4.40}$$

$$\begin{aligned}
J_u[v] = \frac{\partial}{\partial u}J(\sigma,u)[v] =& c_1\langle u_t, v_t\rangle - c_1\langle v_t, div(\sigma)\rangle - c_2\beta\langle\sigma,\nabla v\rangle + c_2\langle\nabla u,\nabla v\rangle \\
&- c_1\langle u_t, f'(u)\cdot v\rangle - c_1\langle v_t, f(u)\rangle - c_1\langle div(\sigma), f'(u)\cdot v\rangle + c_1\langle f(u), f'(u)\cdot v\rangle
\end{aligned} \tag{4.41}$$

Following the same principles one can determine the second order partial derivatives whose derivation will only be briefly outlined here for the most difficult terms which are those including f.

$$\frac{\partial^2}{\partial\sigma^2}J[\tau][\rho] = c_2\langle\rho,\tau\rangle + c_1\langle div(\rho), div(\tau)\rangle \tag{4.42}$$

$$\frac{\partial^2}{\partial\sigma\partial u}[v][\tau] = \frac{\partial^2}{\partial u\partial\sigma}[\tau][v] = -\langle\tau,\nabla v\rangle - \langle v_t, div(\tau)\rangle - \langle div(\tau), f'(u)v\rangle \tag{4.43}$$