A Treatise on the Finite Element Method (working title)

Lucas Payne

May 1, 2021

- The analytic theory of variational methods. Euler-Lagrange equations.
- Try and derive and describe integration by parts intuitively (Why should it be possible? What does it do to differentiability requirements and why is this allowable, what does this imply?)
- Discussion of Dirichlet's principle and resolution of problems with Sobolev spaces, distributions, etc.
- Examples. Poisson's problem analysis.
- The finite element method, overview and application to Poisson's problem (linear), Burger's equation (non-linear).
- Galerkin methods.
- ...
- Using the finite element method to solve some very difficult problem (such as the Navier-Stokes equations in 2D).
- ...
- (possibly appendices)
- Numerics, discussion of libraries for equation solving (petsc, linpack, lapack), standard numerical tools used often in finite element solving. Possibly dicuss issues such as conditioning and precision.

Contents

1 Introduction

Modelling physical processes and problems with partial differential equations, then solving them numerically. Possibly a physically-motivated derivation of the heat equation as described in Larsson.

2 Variational calculus

2.1 Some history and motivation

The variational calculus is effectively the calculus of function spaces. Variational methods consider the local changes of functionals with respect to perturbations of their input functions. The calculus stems from techniques developed around the turn of the 18th century. A classical problem of a variational nature, which was a great motivator for the development of these methods, is a challenge posed by Swiss mathematician Johann Bernoulli. The challenge was taken up in 1697 by Sir Isaac Newton, Gottfried Leibniz, Marquis de l'Hôpital, and Jacob Bernoulli, among others. The brachistochrone problem, as stated by Bernoulli, is [1]

" Given two points A and B in a vertical plane, what is the curve traced out by a point acted on only by gravity, which starts at A and reaches B in the shortest time. "

A fundamental idea is that the optimal curve (a function $C : [0, 1] \to \mathbb{R}^2$) is the minimizer of a certain "energy" functional E, a function from the space of feasible curves to $\mathbb{R}$. A *functional derivative* is defined, leading to a theorem in complete analogy to the principle, attributed to Fermat, that a necessary condition for x^* to be a minimizer of a differentiable function $f(x)$ is that the function is stationary at x^*:

$$\nabla f(x^*) = 0.$$

One such derivative is the *Gâteaux derivative*, denoted $\frac{dE}{du}$, which will be defined below. u, in a certain sense, is from the space of possible infinitesmal deformations of the curve, which preserve required boundary conditions and differentiability. With a functional derivative defined, the analogue to Fermat's theorem is that

$$\frac{dE}{du}(u^*) = 0$$

is necessary for u^* to be a minimizer. This PDE is called the *Euler-Lagrange equation* of the minimization problem, and, in particular, gives the exact minimizer if the minimized energy $f(u)$ is strictly convex. [1]

Application to brachistochrone problem, deriving cosh solution by defining a suitable energy

[1] $f(u)$ is strictly convex if $\theta f(u) + (1 - \theta)f(v) > f(\theta u + (1 - \theta)v)$ for $0 < \theta < 1$. For $f \in C^2$ this means that the function has strictly positive curvature, and thus no non-global minimizers.

A more fundamental example is given by Poisson's equation,

$$\Delta u = g, \tag{1}$$

which is the Euler-Lagrange equation of a certain quadratic (and therefore convex) energy minimization problem, minimizing what is called the *Dirichlet energy*, $f(u) = \int_\Omega \|\nabla u\|^2 \, dx$. This idea, and its validity, was heavily investigated by mathematicians [—when and by who?], and is called *Dirichlet's principle* [2]. The minimizer of f is not necessarily a solution to (1), but it is necessarily a solution to what is called the *weak* form of (1),

$$\int_\Omega u \nabla \phi \, dx = \int_\Omega g\phi \, dx. \tag{2}$$

(—dubious)

Correct details

2.2 Deriving the Euler-Lagrange equations

The functional derivative and the Euler-Lagrange equations are the very basic tools of variational calculus, directly analagous to the gradient and Fermat's theorem of stationary points. Here they are derived for a certain class of functionals, those defined by density integrals. Some examples of functionals of this kind are:

- The Dirichlet energy, $E(u) = \int_\Omega \|\nabla u\|^2 \, dx$, which is the Euclidean square-norm of the gradient. Minimization leads to harmonic functions, solutions of $\nabla u = 0$ on the interior.
- $E(u) = \int_\Omega \frac{1}{2} \|\nabla u\|^2 + ug \, dx$ for some $g : \Omega \to \mathbb{R}$ defined on the interior of Ω. Minimization leads to solutions of Poisson's equation, $\nabla u = g$ on the interior or exterior of the closed domain. This is used for example to compute gravitational potential fields.
- The length of a curve segment, where $c : [0, 1] \to \mathbb{R}^2$ is differentiable, $E(c) = \int_0^1 \|c'(x)\| \, dx$.
- The area of a surface patch, where $s : [0, 1]^2 \to \mathbb{R}^3$ is differentiable, $E(s) = \int_0^1 \int_0^1 \sqrt{\det(J_s(x, y)^T J_s(x, y))} \, dx \, dy$. Minimizing this energy is a case of what is called Plateau's problem.
- $E(u) = \int_\Omega \|\nabla u\| + \frac{\lambda}{2} \|f(x) - u(x)\|^2 \, dx$, an energy whose minimization balances the total variation and the deviation from some target function. This is a popular functional in image denoising.
- (todo: correct physics) The work done by a particle moving as $c : [0, 1] \to \mathbb{R}$ in force field v, $\int_0^1 c'(t) \cdot v(c(t)) \, dt$.

- A light ray travels in a curve $c : [0, 1] \to \mathbb{R}^2$ with endpoints predetermined, through an inhomogeneous medium with varying speed-of-light $\phi(x) > 0$. Then $E(c) = \int_0^1 \frac{c'(t)}{\phi(c(x))}\, dx$ is minimized by Fermat's principle of least time.

The discussion here is not rigorous.

2.2.1 The Gâteaux derivative

Define U to be a suitably general function space for which this works.

Suppose we have some functional energy E, a function of $u \in U$ and its first n derivatives, which is a density integral over the *Lagrangian* $\mathcal{L}$,

$$E = E(u, u', \cdots, u^{(n)}) = \int_\Omega \mathcal{L}(u(x), u'(x), \cdots, u^{(n)}(x))\, dx.$$

Here $u^{(k)}$ is understood to be the tensor of all k'th partial derivatives. To measure the response of the functional E to small perturbations of u, $u + \epsilon h$ for some h, we need $u + \epsilon h$ to have sufficiently many derivatives and bounded integrals (or is "admissable" in Dirichlet's terminology [2]). Suppose h is defined to be of some function space V such that this is true. The Gâteaux derivative is derived through taking limits in the exact same way as the regular derivative. If u is perturbed by h then u' is perturbed by h', and so on. If the limit exists, then the linear response of E to this perturbation is

$$\frac{dE}{du}\Big|_h := \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(E(u + \epsilon h, u' + \epsilon h', \cdots, u^{(n)} + \epsilon h^{(n)}) - E(u, u', \cdots, u^{(n)}) \right).$$

For energy functionals based on a Lagrangian $\mathcal{L}$, this gives by Taylor expansion of E with respect to the $u^{(i)}$,

$$\begin{aligned} \frac{dE}{du}\Big|_h &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_\Omega \mathcal{L}(u + \epsilon h, u' + \epsilon h', \cdots, u^{(n)} + \epsilon h^{(n)}) - \mathcal{L}(u, u', \cdots, u^{(n)})\, dx \\ &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_\Omega \mathcal{L} + \epsilon h \frac{\partial \mathcal{L}}{\partial u} + \epsilon h' \frac{\partial \mathcal{L}}{\partial u'} + \cdots + \epsilon h^{(n)} \frac{\partial \mathcal{L}}{\partial u^{(n)}} + O(\epsilon^2) - \mathcal{L}\, dx \\ &= \int_\Omega h \frac{\partial \mathcal{L}}{\partial u} + h' \frac{\partial \mathcal{L}}{\partial u'} + \cdots + h^{(n)} \frac{\partial \mathcal{L}}{\partial u^{(n)}}\, dx \end{aligned} \tag{3}$$

Explain why this result is expected, as it just adds up the responses of the density.

This is a suitable definition of a "functional derivative" of an energy defined by a Lagrangian, but it would be nice to form instead a "functional gradient", an object encoding functional derivative information in all directions of perturbation.

In analogy to the gradient encoding all directional derivatives,

$$\lim_{\epsilon \to 0} \frac{f(x + \epsilon d) - f(x)}{\epsilon} = \langle \nabla f, d \rangle ,$$

the functional gradient $\frac{dE}{du}$ is defined such that

$$\frac{dE}{du}\Big|_h = \int_\Omega h \frac{\partial \mathcal{L}}{\partial u} + h' \frac{\partial \mathcal{L}}{\partial u'} + \cdots + h^{(n)} \frac{\partial \mathcal{L}}{\partial u^{(n)}} \, dx = \left\langle \frac{dE}{du}, h \right\rangle = \int_\Omega \frac{dE}{du}(x) h(x) \, dx.$$

This is possible due to what at first may seem like an unintuitive trick, repeated integration by parts.

2.2.2 Integration by parts

Make good derivation and reasoning for integration by parts. Is exterior calculus necessary? It seems to be for making sense of multi-dimensional integration by parts. Introduce any formalisms used in the derivation.

Integration by parts is a generalization of the generalized Stokes theorem. The *generalized Stokes theorem* is written

$$\int_\Omega d\omega = \int_{\partial \Omega} \omega. \tag{4}$$

Integration by parts formulae are a consequence of the (alternating) Leibniz product rule satisfied by the exterior derivative,

$$\int_\Omega d(\phi \wedge \omega) = \int_{\partial \Omega} \phi \wedge \omega = \int_\Omega d\phi \wedge \omega + (-1)^{\deg(\phi)} \int_\Omega \phi \wedge d\omega.$$

Work out boundary terms, here the boundary integral is assumed zero (not always valid).

Repeated integration by parts is then induced,

$$\begin{aligned}
\frac{dE}{du}\Big|_h &= \int_\Omega h \frac{\partial \mathcal{L}}{\partial u} + h' \frac{\partial \mathcal{L}}{\partial u'} + \cdots + h^{(n)} \frac{\partial \mathcal{L}}{\partial u^{(n)}} \, dx \\
&= \sum_{k=0}^{n} \int_\Omega \frac{\partial \mathcal{L}}{\partial u^{(k)}} d^k h \, dx \\
&= \sum_{k=0}^{n} \int_\Omega (-1)^k h d^k \frac{\partial \mathcal{L}}{\partial u^{(k)}} \, dx \\
&= \int_\Omega \left(\sum_{k=0}^{n} (-1)^k d^k \frac{\partial \mathcal{L}}{\partial u^{(k)}} \right) h \, dx.
\end{aligned}$$

It follows that

$$\frac{dE}{du} \coloneqq \sum_{k=0}^{n} (-1)^k d^k \frac{\partial \mathcal{L}}{\partial u^{(k)}} \tag{5}$$

is a suitable definition. Setting this to zero gives the Euler-Lagrange equation of the functional E, whose solutions are stationary points of E with respect to variations.

2.2.3 Some examples of Euler-Lagrange equations

Here the Laplace equation and Plateau's problem for minimal surfaces are considered. It is shown that the Laplace equation solves for a stationary function of the *Dirichlet energy*, whose gradient descent algorithm leads to the heat equation.

---

example

Applying the Euler-Lagrange equations to find stationary points of the arc-length/area functional of a curve/surface leads to a non-linear diffusion process called mean-curvature flow, which can be implemented on a computer to find locally *minimal surfaces*, surfaces which interpolate a boundary and whose area can only increase under small perturbations.

---

example

---

3 The finite element method

Discuss numerical solution of partial differential equations, central ideas in finite element methods. I don't think a comparison to other methods would be very useful, should give a self-contained motivation for this particular set of discretization ideas. Describe connection to variational methods. Galerkin methods.

3.1 Weak forms of partial differential equations

The partial differential equation (without initial or boundary conditions) is formulated as

$$\mathcal{A}u = f,$$

where $\mathcal{A}$ is a (not necessarily linear) differential operator. The weak form of this PDE is

$$\langle \mathcal{A}u, \phi \rangle = \langle f, \phi \rangle \tag{6}$$

for $\phi \in S \subset V$. If S is finite dimensional with a basis of $\{\phi_1, \cdots, \phi_n\}$, then a solution to (6) is a solution to the system of equations

$$\langle \mathcal{A}u, \phi_i \rangle = \langle f, \phi_i \rangle$$

for $1 \leq i \leq n$. Supposing $u_s = u^i \phi_i$ is an approximate solution, and f_I is f projected onto S, we can instead solve a finite system of equations

$$\begin{aligned} & \langle \mathcal{A} u_s, \phi_i \rangle = \langle f, \phi_i \rangle \\ \equiv \quad & \left\langle \mathcal{A}(u^j \phi_j), \phi_i \right\rangle = \langle f_I, \phi_i \rangle \,. \end{aligned}$$

This system can be expanded, even for non-linear $\mathcal{A}$, into a system of equations in mass matrices. For example, consider the non-linear equation with its weak form for test functions in S:

$$\begin{aligned} \text{Strong form} \quad & -u'' + u^2 = f \\ \text{Weak form} \quad & \int_\Omega u' \phi_i' \, dx + \int_\Omega u^2 \phi_i \, dx = \langle f, \phi_i \rangle \end{aligned}$$

Considering again $u_s = u^i \phi_i$ and f_I the projected f, the discrete weak form is

todo

——— INCORRECT ——— This is a non-linear system of equations that is feasibly solvable with a computer. Letting $A = \left(\left\langle \phi_i', \phi_j' \right\rangle\right)$, $B = \left(\left\langle \phi_i^2, \phi_j \right\rangle\right)$, $C = \left(\langle \phi_i, \phi_j \rangle\right)$, the equations are

$$A\bar{u} + B\bar{u}^2 = C\bar{f},$$

where $\bar{u}^2$ is squared component-wise, and $\bar{u}$, $\bar{f}$ are respective vectors of coefficients.

3.2 Computing integrals over trial functions

Let $G_T : T \to N \subset \mathbb{R}^n$ be defined as

$$G_T(x) = p^i \Phi_i(x)$$

where $p^i \in \mathbb{R}^n$ are *control points* and Φ_i are basis functions on the reference domain which form a partition of unity. This is a common paradigm for parametric surfaces, such as Bézier patches, in geometric design.

$$J = \frac{\partial G_T(x)_\gamma}{\partial x_\tau} = p^i_\gamma \frac{\partial \Phi_i(x)_\gamma}{\partial x_\tau}$$

Consider a point-quadrature rule over points $g_i \in T$ with corresponding weights w^i. The change of variables formula and quadrature is

$$\int_N f(x) \, dx = \int_T f(x) |J(x)| \, dx \approx w^i f(g_i) |J(g_i)|.$$

We can then reduce the amount of computation needed for these quadratures, if we know the basis functions over each element *a priori*, by pre-computing a tensor of values

$$\mathrm{TABLE}(i, \gamma, \tau) = \frac{\partial \Phi_i(g_i)_\gamma}{\partial x_\tau}.$$

This stores all information about Jacobians of the element-mapping basis functions at the relevant quadrature points, and thus forms a basis for the discrete Jacobian-field on the reference element. The unknowns are the control points p^i. The quadrature is then computed as

$$\int_N f(x)\,dx \approx w^i f(g_i) \left| \left(p^k_\gamma \,\mathrm{TABLE}(k, \gamma, \tau)\right)_{\gamma,\tau} \right|.$$

$dPhi(i, ..., \tau), dPsi, Phi(i, ...)$ "..." are for vector-valued basis functions. τ is the axis index of the partial derivative.

4 Solving Poisson's equation

5 Solving the inviscid Burgers equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0. \tag{7}$$

Using implicit Euler,

$$\begin{aligned} u^0 &= g \\ \frac{u^{n+1} - u^n}{\Delta t} &+ u^{n+1}\frac{\partial u^{n+1}}{\partial x} = 0. \end{aligned}$$

In the update equation, letting $u = u^{n+1}$ be the unknown and $u_p = u^n$ be constant, integrating against a test function gives

$$\int_\Omega \frac{u}{\Delta t} v + u\frac{\partial u}{\partial x} v\,dx = \int_\Omega \frac{u_p}{\Delta t} v\,dx.$$

Let $u = u^i\phi_i$ be a combination of basis trial functions. Since the left-hand-side is linear in the test function v, it is necessary and sufficient that the equation is satisfied separately for each $v = \psi_j$. This gives a finite system of equations

$$\frac{u^i}{\Delta t}\int_\Omega \phi_i\psi_j\,dx + \int_\Omega (u^i\phi_i)(u^i\frac{\partial \phi_i}{\partial x})\psi_j\,dx = \frac{1}{\Delta t}\int_\Omega u_p\psi_j\,dx.$$

6 Deriving the Navier-Stokes equations

Motivation

6.1 Conservation laws and continuity equations

Let $\Omega \subset U$ be a domain in U. A conserved quantity, named here energy (measured in Joules), can advect and disperse, and this is modelled in general by a flux function $j(x,t)$. For $U = \mathbb{R}^3$ this flux function has units $Jm^{-2}s^{-1}$. It is assumed that energy is introduced into the system purely through source term $p(x,t)$ with units $Jm^{-3}s^{-1}$.

The rate of change of total energy in Ω is then

$$\frac{d}{dt}\int_\Omega E\,dx = \int_{\partial\Omega} -j\cdot\hat{n}\,dS + \int_\Omega p\,dx,$$

which by the divergence theorem is

$$\int_\Omega \frac{\partial E}{\partial t} + \nabla\cdot j - p\,dx = 0.$$

For this to be true of general domain Ω, the integrand must be identically zero:

$$\frac{\partial E}{\partial t} + \nabla\cdot j = p. \tag{8}$$

This is a general form for a conservation law. If the flux j is measured as the amount of energy being moved over a small boundary element in the direction of the normal by a vector field u, then the vector field on the boundary of region Ω is transfering the quantity there either into, along, or out of the region. Therefore in this case $j = Eu$ and the equation becomes

$$\frac{\partial E}{\partial t} + \nabla\cdot(Eu) = p. \tag{9}$$

6.1.1 Conservation of momentum

For example, the conserved quantity may be *momentum*, ρu, the product of scalar mass-density field ρ and flow-velocity field u. As this is a conserved vector quantity, the conservation law is written separately for each component of momentum, which is then compacted with $\otimes$ denoting the outer product:

$$\begin{aligned} & \frac{\partial(\rho u_i)}{\partial t} + \nabla\cdot(\rho u_i u) = p, \quad i = 1,2,3 \\ \equiv \quad & \frac{\partial(\rho u)}{\partial t} + \nabla\cdot(\rho u\otimes u) = p. \end{aligned} \tag{10}$$

6.1.2 Conservation of mass

Equation (10) is incomplete. Nothing has been said about restrictions on the mass density ρ. The completing equation is

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho u) = 0, \tag{11}$$

which expresses that the only way mass enters or exits a closed region is by advection by u. If u is conservative, mass is conserved. (10) and (11) combined give the correct formulation, from first principles, of Newton's second law in continuum mechanics. A more direct derivation, which makes the connection to Newton's second law much clearer, requires the material derivative.

6.2 Advection, transport, and the total and material derivatives

Consider a vector field u on a domain U. Any spatial quantity ϕ on U can be transported along u. The *total derivative* is defined as

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial t} + \frac{\partial x}{\partial t} \cdot \nabla \phi.$$

When evaluated at p, this measures the rate of change of ϕ with respect to t-parameterised motion along a curve through p. When this curve is given by the flow of a vector field u, this is called the *material derivative*, denoted

$$\frac{D\phi}{Dt} = \frac{\partial \phi}{\partial t} + u \cdot \nabla \phi.$$

Setting this to zero completes the derivation of the transport equation. $\frac{D\phi}{Dt} = 0$ (with u implicit) states that from the point of view of a particle moving along u, ϕ remains unchanged. u can, for example, model the deformation velocity field of a material.

6.3 The Cauchy momentum equation

In the 1820s Cauchy began the extension of Euler's laws of mechanics to deformable materials. A fundamental object introduced is the *Cauchy stress tensor*, denoted here by $T = T(x, t)$, which will now be motivated.

To distinguish continuum mechanics from point-particle mechanics, model-dependent internal forces are required. If these internal forces were not present then this could hardly be called a cohesive object, and the mechanics would effectively reduce to that of a continuum of non-interacting point-particles.

By *internal force* what is meant here is a force exerted on a small region of the material by its immediately adjacent particles. Consider $\Omega \subset U$ a domain. Integrating the exerted forces along the boundary $\partial\Omega$ gives a vector flux, for which the divergence theorem applies.

$$\int_{\partial\Omega} T \cdot \hat{n}\, dS = \int_{\Omega} \nabla \cdot T\, dV. \tag{12}$$

The divergence theorem here gives a differential force, a measurement of internal stress for each point in the material. Here $\nabla \cdot T$, when U is three-dimensional, is understood to be the vector

$$\nabla \cdot T = \begin{bmatrix} \nabla \cdot T_1 & \nabla \cdot T_2 & \nabla \cdot T_3 \end{bmatrix}^T,$$

where T_i denote the stress in the unit normal aligned to axis x_i (in matrix terms, the i'th column of T).

Geometric derivation of Cauchy stress tensor. (why is it a linear transformation?)

The *Cauchy momentum equation* is

$$\rho \frac{Du}{Dt} = \nabla \cdot T + \rho g. \tag{13}$$

$g = g(x,t)$ is a space-and-time-varying body force acting on the system, such as, for example, Earth's gravitational pull. This g is a force, which is per-unit-mass, which is why ρg appears in (13).

—rederive this, maybe incorrect masses

As defined above, $\frac{D}{Dt}$ denotes the *material derivative*. In the mechanics of a point particle, the particle has a velocity at each time. The definition of velocity and acceleration *follow* this point. This leads to the relevant idea of acceleration in continuum mechanics: $\frac{Du}{Dt} = \frac{\partial u}{\partial t} + u \cdot \nabla u$. By inertia we want to follow a particle as it continues in the direction of its velocity, and see if this velocity is changing. This readily gives an interpretation of (13) as simply an instance of

$$F = ma. \tag{14}$$

The left-hand-side of (13) is mass (here a differential density) times acceleration, and the right-hand-side is force. Rather than a general force term, (13) splits the forces into internal stress forces determined by the Cauchy stress tensor, and other external body forces. This is a general momentum equation for continuum mechanics. A specific instance of a problem with a certain continuum material model requires

- A definition of the Cauchy stress tensor, which likely relies on differential operators.
- A definition of any other external body forces acting on the material.
- Any other restrictive equations (for example, $\nabla \cdot u = r$).
- Initial/boundary conditions for which these combined equations form a well-posed problem.

6.4 The Euler equations for inviscid flow

The Cauchy momentum equation is the basic skeleton for developing equations of fluid motion.

6.4.1 Incompressibility

We will assume that the fluid is *incompressible*. Compressibility is the basis of acoustic waves, so of course this assumption is non-physical. However, if we are interested in the bulk-movement flow, in many cases acoustic waves will have neglible effect.

6.4.2 Inviscid flow

We need to define a Cauchy stress tensor. Internal forces are measured across an oriented surface element δS with normal $\hat{n}$ by evaluating $T \cdot \hat{n}$. The component of this force perpendicular to $\hat{n}$ is called a *shear force*. If there is no shear force, then the matrix representation of the Cauchy stress tensor will be αI for some scalar function α. The tensor divergence is equal to a gradient: $\nabla \cdot (\alpha I) = \nabla \alpha$. Letting α here be $\alpha = -p$, the *Euler equations of motion* are

$$\begin{aligned} \rho \frac{Du}{Dt} &= -\nabla p + \rho g \\ \nabla \cdot u &= 0. \end{aligned} \tag{15}$$

The *pressure* p is an unknown. The Cauchy stress tensor, and thus the pressure, is *defined* to give those internal non-shear forces which give incompressibility.

6.5 Viscous flow and Newtonian fluids

So far the derivations have been geometric rather than physical.

7 Solving (certain cases of) the Navier-Stokes equations

8 Appendix A: Classical mechanics

8.1 Potential and kinetic energy

Consider the paths of particles in a closed system with position and momentum evolving in time due to the action of a conservative force, $f = -\nabla U$ where U is called the *potential energy*. Unit mass is assumed here.

$$x'' = f(x) = -\nabla U(x). \tag{16}$$

Given the conservative force field f, U can be defined by quadrature as

$$U(x(t)) = x_0 + \int_{t_0}^{t} x'(\tau) \cdot (-f(x(\tau)))\, d\tau. \tag{17}$$

We want to derive an energy conservation law $\frac{dE}{dt} = 0$, so as $U(x(t))$ varies, we need to define some energy term T where the potential energy "comes from"

and "goes", and let $E = U + T$. Potential energy U is purely a function of position x. Assume that $T = T(x'(t))$ is purely a function of momentum, with the assumption of unit mass. We then need $\frac{dT}{dt} = -\frac{dU}{dt}$. By writing T in terms of quadrature along the phase curve $(x(t), x'(t))$, we have

$$T(x'(t)) = x'_0 + \int_{t_0}^{t} f(x(\tau)) \cdot \frac{dT}{dx'}\, d\tau. \tag{18}$$

Therefore we want $\frac{dT}{dx'} = x'$, and a possible energy term is $T = |x'|^2/2$. This, by construction, gives

$$\frac{d}{dt}(U + T) = [x'(t) \cdot (-f(x(t)))] + [f(x(t)) \cdot x'(t)] = 0. \tag{19}$$

$T(x')$ is called the *kinetic energy*. $E = U + T$ is called the *total energy* of the configuration $(x(t), x'(t))$ in the closed system.

References

[1] Johann Bernoulli, *"Problema novum ad cujus solutionem Mathematici invitantur." (A new problem to whose solution mathematicians are invited.)*, 1696. (retrieved from wikipedia/brachistochrone_curve)

[2] A. F. Monna, *Dirichlet's principle: A mathematical comedy of errors and its influence on the development of analysis*, 1975

[3] Stig Larsson, *Partial differential equations with numerical methods*, 2003