

Introduction to Mathematical Modeling

Class Notes

Based on the lectures and class notes of Prof. Robert Moser

Jacob Badger

1. Equations

The following equation sheets were provided by TA Gopal Yalla, http://users.odan.utexas.edu/~gopal/teaching/2019_CSE389C.

Local Conservation Laws

	Eulerian	Lagrangian
Mass	$\frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho \mathbf{v}) = 0$	$\varrho_0(X) = \varrho(\varphi(X)) \det F(X)$
Lin. Momentum	$\varrho \left(\frac{\partial(\mathbf{v})}{\partial t} + \mathbf{v} \cdot \operatorname{grad} \mathbf{v} \right) = \operatorname{div} T + \mathbf{f}_b$	$\varrho_0 \frac{\partial^2 u}{\partial t^2} = \operatorname{Div} \underbrace{FS}_P + f_0$
Ang. Momentum	$T = T^\top$	$S = S^\top$
Energy	$\varrho \left(\frac{\partial e}{\partial t} + \mathbf{v} \cdot \operatorname{grad} e \right) = T : D + r - \operatorname{div} \mathbf{q}$	$\varrho_0 \dot{e}_0 = S : \dot{E} - \operatorname{Div} \mathbf{q}_0 + r_0$
Entropy	$\varrho \left(\frac{\partial \eta}{\partial t} + \mathbf{v} \cdot \operatorname{grad} \eta \right) + \operatorname{div} \frac{\mathbf{q}}{\theta} - \frac{r}{\theta} \geq 0$	$\varrho_0 \dot{\eta}_0 + \operatorname{Div} \frac{\mathbf{q}_0}{\theta} - \frac{r_0}{\theta} \geq 0$

Maxwell's Equations

	Integral Form	Differential Form
Gauss's Law	$\epsilon_0 \int_{\partial \Omega} \mathbf{E} \cdot \mathbf{n} \, dA = q_\Omega = \int_\Omega \rho \, dx$	$\epsilon_0 \nabla \cdot \mathbf{E} = \rho$
Faraday's Law	$\int \mathbf{E} \cdot d\mathbf{s} = -\frac{d}{dt} \int_A \mathbf{B} \cdot \mathbf{n} \, dA$	$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$
Amp.-Max. Law	$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 \int_A \mathbf{j} \cdot \mathbf{n} \, dA + \mu_0 \epsilon_0 \int_A \mathbf{n} \cdot \frac{\partial \mathbf{E}}{\partial t} \, dA$	$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$
No Mag. Mono.	$\int_{\partial \Omega} \mathbf{B} \cdot \mathbf{n} \, dA = 0$	$\nabla \cdot \mathbf{B} = 0$

Schrödinger's Equation

General	$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V(x) \Psi$	$V = 0 \text{ for free particle}$
Time Invariant	$\tilde{H} \psi = E \psi$	$\tilde{H}(x, \frac{\hbar}{i} \frac{\partial}{\partial x}) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x)$
Ladder Op Form	$\tilde{H} \psi = \hbar \omega \left(\tilde{a}_\pm \tilde{a}_\mp \pm \frac{1}{2} \right) \psi = E \psi$	$\tilde{a}_\pm = \frac{1}{\sqrt{2\hbar m \omega}} \left(\mp i \tilde{p} + m \omega \tilde{q} \right)$

Kinematics of Deformable Bodies

(pg)

$$1. u = \varphi(X) - X \quad (15)$$

$$2. F(X) = \nabla \varphi(X) = I + \nabla u(X) \quad (15)$$

$$3. C = F^T F \quad (16)$$

$$4. E = \frac{1}{2}(C - I) \quad (16)$$

$$5. v = \dot{x}(\varphi^{-1}(x, t), t) \quad (19)$$

$$6. \frac{D\psi}{Dt} = \frac{\partial \psi}{\partial t} + v \cdot \text{grad } \psi \quad (20)$$

$$7. L = \text{grad } v \quad (20)$$

$$8. L = D + W \quad (20)$$

$$9. D = \frac{1}{2}(L + L^T) \quad (20)$$

$$10. W = \frac{1}{2}(L - L^T) \quad (20)$$

$$11. L_m = \dot{F}F^{-1} \quad (20)$$

$$12. \dot{F} = \text{grad } vF = L_m F \quad (20)$$

$$13. \frac{d}{dt} \det F = \det F \text{div } v \quad (21)$$

$$14. \text{Piola Transform} \quad (22)$$

$$T_0(X) = [\det F(X)] T(X) F(X)^{-T}$$

$$15. \text{Polar Decomposition Theorem} \quad (18)$$

$$F = RU = VR$$

– R orthogonal; U, V sym. P.D.

$$C = F^T F = U^2$$

Divergence Theorem

(21)

$$\int_{\Omega} \text{div } \Psi \, dx = \int_{\partial \Omega} \Psi \cdot \hat{n} \, dA.$$

Reynold's Transport Theorem

(22)

$$\begin{aligned} \frac{d}{dt} \int_{\omega_t} \Psi \, dx &= \int_{\omega_t} \frac{\partial \Psi}{\partial t} \, dx + \int_{\partial \omega_t} \Psi v \cdot \hat{n} \, dx \\ &= \int_{\omega_t} \left(\frac{\partial \Psi}{\partial t} \, dx + \text{div}(\Psi v) \right) \, dx \end{aligned}$$

Mass

$$1. M(\mathcal{B}) = \int_{\Omega_t} \varrho \, dx \quad (30)$$

$$2. \int_{\Omega_0} \varrho_0(X) \, dX = \int_{\Omega_t} \varrho(x) \, dx \quad (30)$$

$$3. \text{Material Conservation of Mass} \quad (37)$$

$$\varrho_0(X) = \varrho(x) \det F(X)$$

$$4. \text{Spatial Conservation of Mass} \quad (30)$$

$$\frac{\partial \varrho(x)}{\partial t} + \text{div}(\varrho(x)v) = 0$$

Linear & Angular Momentum

$$1. \frac{dI(\mathcal{B}, t)}{dt} = \int_{\Omega_t} \varrho \frac{dv}{dt} \, dx = \mathbf{F}_b + \mathbf{F}_s \quad (26)$$

$$2. \int_{\Omega_t} \varrho \frac{dv}{dt} \, dx = \int_{\Omega_t} f \, dx + \int_{\partial \Omega_t} \sigma(n) \, dA.$$

$$3. \text{Cauchy's Theorem}$$

$$\sigma(n, x, t) = T(x, t)n \quad T = T^T$$

$$4. \text{Cauchy Stress}$$

$$T = (\det F)^{-1} P F^T = (\det F)^{-1} F S F^T$$

$$5. \text{First Piola-Kirchoff Stress}$$

$$P = (\det F) T F^{-T} = F S$$

$$6. \text{Second Piola-Kirchoff Stress}$$

$$S = (\det F) F^{-1} T F^{-T} = F^{-1} P$$

Energy

$$1. \text{Total Energy} = k + U$$

$\rightarrow k = \text{kinetic energy}, U = \text{internal energy}.$

$$2. \text{Principle Consv. Energy}$$

$$\frac{d}{dt}(k + U) = P + \dot{Q}.$$

$\rightarrow Q = \text{internal heating}.$

$$3. \varrho \frac{de}{dt} = T : D - \text{div} q + r$$

$$4. \varrho_0 \dot{e}_0 = S : \dot{E} - \text{Div} q_0 + r_0$$

$\rightarrow r = \text{heat per unit volume}.$

2nd Law of Thermodynamics

1. Clausius-Duhem

$$\varrho \frac{d\eta}{dt} + \operatorname{div} \frac{q}{\theta} - \frac{r_0}{\theta}$$

$$\varrho_0 \dot{\eta}_0 + \operatorname{Div} \frac{q_0}{\theta} - \frac{r_0}{\theta} > 0$$

$\rightarrow \theta = \text{temp}, \eta = \text{entropy density}$

Constitutive Equations

1. Material Frame Indifference

$$T^* = QTQ^T$$

2. MFI Solids

$$\mathcal{T}(\underbrace{QF}_{F^*}) = Q \mathcal{T}(F) Q^T$$

3. MFI Fluids

$$\mathcal{F}(\underbrace{QDQ^T}_{D^*}) = Q \mathcal{F}(D) Q^T$$

4. Coleman-Noll (Dissipative)

$$S = \varrho_0 \frac{\partial \Psi}{\partial E}, \quad \eta_0 = -\frac{\partial \Psi}{\partial \theta}$$

$$\frac{\partial \Psi}{\partial \nabla \theta} = 0, \quad -\frac{1}{\theta} q_0 \cdot \nabla \theta \geq 0$$

5. Coleman-Noll

$$S = F(E) + I(\dot{E})$$

$$\cdot F(E) = \varrho_0 \frac{\partial \Psi}{\partial E}$$

$$\cdot I(\dot{E}) : \dot{E} - \frac{1}{\theta} q_0 \cdot \nabla \theta \geq 0$$

Electromagnetic Waves

1. Coulomb's Law

$$F = k \frac{|q_1||q_2|}{r^2}, \quad k = \frac{1}{4\pi\epsilon_0}$$

2. Gauss's Law

$$q_\Omega = \epsilon_0 \oint_{\partial\Omega} E \cdot n \, dA = \int_\Omega \rho \, dx$$

$$\epsilon_0 \nabla \cdot E = \rho \quad (\text{charge density})$$

3. Ampere's Law

$$\oint B \cdot ds = \mu_0 i_{\text{enclosed}}, \quad i = \text{current.}$$

4. Ampere + Maxwell Law

$$\oint B \cdot ds = \mu_0 i + \mu_0 \epsilon_0 \frac{d}{dt} \Phi_E$$

$$\oint B \cdot ds = \mu_0 \int_A j \cdot n \, dA + \mu_0 \epsilon_0 \frac{d}{dt} \int_A E \cdot n \, dA$$

$$\nabla \times B = \mu_0 j + \mu_0 \epsilon_0 \frac{\partial E}{\partial t}$$

5. Faraday's Law

$$\oint E \cdot ds = -\frac{d}{dt} \int_A B \cdot n \, dA$$

$$\nabla \times E = -\frac{\partial B}{\partial t}$$

6. No Magnetic Monopoles

$$\int_{\partial\Omega} B \cdot n \, dA = 0$$

$$\nabla \cdot B = 0$$

Waves

1. $u(x, y) = \mu_0 e^{i(k \cdot x - \omega t)}$

· μ_0 = amplitude of wave

· k = wave number

· ω = angular frequency

· $\lambda = 2\pi/k$ = wave length

· $T = 2\pi/\omega$ = period of oscillation

· $v = \omega/k$ = wave speed

2. General Wave Equation

$$\frac{\partial^2 u}{\partial t^2} = \left(\frac{\omega^2}{k^2} \right) \frac{\partial^2 u}{\partial x^2}$$

3. Electromagnetic Waves

$$E = E_0 e^{i(k \cdot x - \omega t)}, \quad B = B_0 e^{i(k \cdot x - \omega t)}$$

· $\omega/|k|$ = propagation speed

· $\hat{k} = k/|k|$ = propagation direction

· $c = 1/\sqrt{\epsilon_0 \mu_0}$ = speed of light

4. Electromagnetic Wave Equation

$$\frac{\partial^2 E}{\partial t^2} = \frac{1}{\mu_0 \epsilon_0} \Delta E$$

5. E-M Waves & Maxwell's Equations

$$k \cdot E = 0 \implies k \cdot E_0 = 0$$

$$k \cdot B = 0 \implies k \cdot B_0 = 0$$

$$k \times E = \omega B \implies \hat{k} \times E = cB$$

$$k \times B = -\frac{1}{c^2} \omega E \implies \hat{k} \times B = -\frac{1}{c} E$$

$$E \cdot B = 0$$

$$|E_0| = c|B_0|$$

$$E \times B = \frac{k}{\omega}|E|^2 = \frac{1}{c}|E|^2 \hat{k}$$

Quantum Mechanics

$$1. E = \hbar\omega = h\nu$$

$$2. \lambda = \frac{h}{p}$$

$$3. \text{Wave Equation}$$

$$\Psi(x, t) = \psi_0 e^{i(px - Et)/\hbar}$$

$$\rightarrow \Psi(x, t) = \psi_0 e^{i(kx - \omega t)}$$

$$\rightarrow k = 2\pi/\lambda = p/\hbar$$

$$\rightarrow \omega = 2\pi\nu = E/\hbar$$

$$4. E\Psi = \left(-\frac{\hbar}{i}\frac{\partial}{\partial t}\right)\Psi$$

$$5. p\Psi = \left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\Psi$$

$$6. \text{Hamiltonian Operator } H(q, p) = E$$

$$H(q, p) = \frac{p^2}{2m} + V(q) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$$

$$7. \text{Schrodinger's equation (general)}$$

$$i\hbar\frac{\partial\Psi}{\partial t} + \frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} - V\Psi = 0$$

$$8. \Psi^*\Psi = |\Psi(x, t)|^2 = \rho(x, t)$$

$$\rightarrow \int_{-\infty}^{\infty} |\Psi(x, t)|^2 dx = 1$$

$$\rightarrow \frac{d}{dt} \int_{-\infty}^{\infty} |\Psi(x, t)|^2 dx = 0$$

Dynamic Variables & Observables

$$1. \text{Dynamic Variable}$$

$$Q = Q(q_1, q_2, \dots, q_N; p_1, p_2, \dots, p_N)$$

$$\tilde{Q} = \tilde{Q}(q_1, \dots, q_N; -i\hbar\frac{\partial}{\partial q_1}, \dots, -i\hbar\frac{\partial}{\partial q_N})$$

$$2. \langle Q \rangle = (\Psi, \tilde{Q}\Psi)_{L^2} = \int \Psi^* \tilde{Q}\Psi dq$$

$$3. \text{Hermitian}$$

$$\langle \psi, A\phi \rangle = \langle A\psi, \phi \rangle \quad \forall \phi, \psi \in L^2$$

$$4. \sigma_Q^2 = \langle Q^2 \rangle - \langle Q \rangle^2$$

$$5. \text{Generalized Uncertainty Principle}$$

$$\sigma_Q^2 \sigma_M^2 \geq \left(\frac{1}{2i} \langle [\tilde{Q}, \tilde{M}] \rangle \right)^2$$

Hydrogen Atom

$$1. \text{The complete hydrogen wave functions}$$

$$\psi_{n\ell m} = R_{n\ell}(r)Y_{\ell m}(\theta, \phi)$$

$$n = 1, 2, \dots \quad (\text{energy level})$$

$$\ell = \underbrace{0}_s, \underbrace{1}_p, \underbrace{2}_d, \underbrace{3}_f, \dots, n-1 \quad \begin{matrix} \text{(shape)} \\ \text{orbital} \end{matrix}$$

$$m = 0, \pm 1, \pm 2, \dots, \pm \ell \quad (\text{orientation})$$

2. Dimensional Analysis

When modeling complex systems, we commonly have a set of applicable, reliable theories, which are not "closed." We then need "Closure models," often representing effects and phenomena that are not accessible by the reliable theory.

EXAMPLE: Continuum Mechanics:

Reliable Theory: Conservation laws

Closure Models: Constitutive models.

Reliable theory represents centuries of accumulated empirical knowledge, their validity is generally not questioned for the problem at hand. Constitutive models are basically "made up" to be consistent with what is known and available data. This is part of the challenge of mathematical modeling. These closure models can be formulated with the i) Empirical observations, ii) Invariance principles and dependency assumptions, iii) Established theories and principles, iv) insightful guessing.

The invariance principles can significantly constrain the closure models we formulate. Let us consider two broadly useful invariance constraints: Dimensional invariance and Coordinate invariance.

2.1 Dimensional Invariance (Homogeneity)

A mathematical model of a physical system is a set of relationships among mathematical descriptors of that system.

Question 1: *Consider an object as a physical system. What descriptors might be useful?*

Dimensions, measures on properties.

These descriptors (measures of properties) are defined in terms of certain arbitrary references. The principle of dimensional homogeneity is this: The validity of a model relationship cannot depend on the arbitrary references used to define the descriptors—because these descriptors are arbitrary!

Consider a relationship between two descriptors: $A = B$ in some unit system U . In the usual way, we can express the same descriptors in terms of a second unit system U' , i.e. $A' = C'_A A$ and $B' = C'_B B$ where C'_A, C'_B are conversion factors. To be independent of our arbitrary unit system, we must have $A' = B' \implies C'_A = C'_B$. But the unit system U' is arbitrary, so it must be that $C'_A = C'_B$ for conversion from U to any other unit system U' .

This requirement of invariance to the unit system leads to the algebra of dimensions. If two descriptors A and B have dimension (e.g. length, mass, time) denoted $[A]$ and $[B]$ respectively, then AB has dimensions $[A][B]$ and A/B has dimensions $[A]/[B]$ where having the dimension $[A][B]$ means that the conversion factor for transforming units from U into U' is $C'_A C'_B$

EXAMPLE: Area has dimensions of $[length]^2$ or $[L]^2$. Velocity has dimensions $\frac{[L]}{[T]}$

EXAMPLE: Newton's law implies $F = ma$, then dimensions of F are $[M] \cdot \frac{[L]}{[T]^2} = \frac{[M][L]}{[T]^2}$

This is of course trivial, but the requirement of dimensional homogeneity places important constraints on any model of a physical system.

Consider a model of some physical system that asserts a relationship among n descriptors of the system, x_1, \dots, x_n :

$$F(x_1, x_2, \dots, x_n) = 0$$

which satisfies the constraint of invariance to the unit system. Generally, the variables will have dimensions that can be expressed in terms of a set of $m < n$ independent dimensions or fundamental units (e.g. mass, length, time).

We can choose m of the n variables (WLOG, x_1, \dots, x_m) that have independent dimensions. Note that independent dimensions are analogous to linearly independent vectors in linear algebra and can indeed be represented this way by expressing each variable as a vector with coefficients equal to the power of a fundamental unit present in its units. Now we are free to choose a unit system—so let us choose the units in which x_1, \dots, x_m are all one. In these units the variables \tilde{x}_i are given by,

$$\tilde{x}_i = \begin{cases} 1 & i \leq m \\ x_i \prod_{j=1}^m x_j^{P_{ij}} & m < i \leq n \end{cases}$$

where the powers P_{ij} are chosen such that \tilde{x}_i are dimensionless. The assumptions of m independent dimensions and m variables with independent dimensions guarantees that we can find appropriate P_{ij} (Linear Algebra). Expressing the model in these units gives,

$$F(\underbrace{1, \dots, 1}_m, \tilde{x}_{m+1}, \dots, \tilde{x}_n) = 0.$$

Since the first arguments are constants we can rewrite as,

$$\tilde{F}(\tilde{x}_{m+1}, \dots, \tilde{x}_n) = 0.$$

This is a big deal as we just reduced the dimensionality of our model by m .

This result is called the Buckingham II Theorem because Buckingham used Π_i instead of \tilde{x}_i but the result was obtained much earlier.

The variables x_1, \dots, x_m which we used to non-dimensionalize the remaining variables are referred to as *scaling* variables.

Question 2: *How do you choose scaling variables?*

There is no exact set of rules, but typically you want to choose variables that you believe will have the strongest or most direct impact on the phenomenon you are trying to model. This is especially important when later trying to simplify your model, if you chose an insignificant parameter it is then difficult if not impossible to remove it without going back to square one and re-picking scaling variables.

The bottom line. If the dependencies of a model are known or can be postulated, dimensional analysis can be used to infer simpler models and consolidate data.

2.2 Coordinate Invariance (Tensor Analysis)

Cartesian tensor representation (in 3-D). Let us define three mutually orthogonal unit vectors \mathbf{e}_i ,

$$(\mathbf{e}_i, \mathbf{e}_j) = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

where (\cdot, \cdot) denotes the inner product. Any vector $\mathbf{x} \in \mathbb{R}^3$ can be represented as a linear combination of these basis vectors,

$$\mathbf{x} = \sum_{i=1}^3 x_i \mathbf{e}_i,$$

where $x_i \in \mathbb{R}$. Any x that can be represented in this way is called a rank-1 tensor (scalars are rank-0 tensors).

A rank-2 tensor \underline{Y} can be represented in terms of basis vectors by defining what we call a tensor product. The tensor product of two vectors \mathbf{a} and \mathbf{b} —denoted $\mathbf{a} \otimes \mathbf{b}$ is the operator that assigns for each vector \mathbf{c} the vector $(\mathbf{b} \cdot \mathbf{c})\mathbf{a}$; that is,

$$\mathbf{a} \otimes \mathbf{b} := (\mathbf{b} \cdot \mathbf{c})\mathbf{a}.$$

Note that a tensor product is a homogeneous linear operator that maps a vector (\mathbf{c} in the definition) to another vector (\mathbf{a}) scaled by the inner product of \mathbf{b} and \mathbf{c} .

In particular, we know that homogeneous linear operators on \mathbb{R}^3 are just 3×3 matrices, so rank-2 tensors can be represented using 3×3 matrices. In particular,

$$\underline{Y} = \sum_{i=1}^3 \sum_{j=1}^3 Y_{ij} \mathbf{e}_i \otimes \mathbf{e}_j.$$

Taking the product of \underline{Y} with some vector \mathbf{x} gives,

$$\begin{aligned} \underline{Y}\mathbf{x} &= \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 Y_{ij} (\mathbf{e}_i \otimes \mathbf{e}_j) \mathbf{e}_k x_k \\ &= \sum_{i=1}^3 \left(\sum_{j=1}^3 \sum_{k=1}^3 Y_{ij} \delta_{jk} \mathbf{x}_k \right) \mathbf{e}_i = \sum_{i=1}^3 \left(\sum_{j=1}^3 Y_{ij} \mathbf{x}_j \right) \mathbf{e}_i = \mathbf{z}, \text{ a vector.} \end{aligned}$$

But note this is a pain to carry around all the summations and unit vectors so we adopt the Einstein summation convention.

Einstein summation convention:

Tensors are represented as indexed objects, e.g. x_i, y_{ij}

In any term, an index can appear at most twice

A repeated index implies summation over that index

A non-repeated index implies multiplication with unit vector with that index

EXAMPLE:

$$z_i = y_{ij} x_j = \sum_{i=1}^3 \left(\sum_{j=1}^3 y_{ij} x_j \right) \mathbf{e}_i.$$

Scalars and Invariants Invariants of vectors/tensors are just scalar functions of those vectors/tensors. They are called invariants because they are invariant to the coordinate system in which the tensors are expressed.

EXAMPLE: Which of the following are scalar invariants?

$$\begin{aligned} a_{ij} \\ a_{ii} \leftarrow \\ b_i \\ b_i b_i \leftarrow \\ a_{ij} a_{ji} \leftarrow \end{aligned}$$

For rank-2 tensors in 3-D, two important results:

The Cayley-Hamilton Theorem. A matrix is a solution to its own characteristic equation. This means that A^n for $n > 2$ is a linear combination of I, A, A^2 (A is a rank-2 tensor).

The key point of this theorem is that any analytic tensor function of a tensor $F(A)$ is a linear combination of I, A, A^2 , with coefficients that are scalar functions of the invariants of A .

There are only 3 independent invariants of a 3-D rank-2 tensor. This implies that any scalar function of a tensor A can be expressed as a function of its eigenvalues.

A "standard set" of invariants are defined as the coefficients of the characteristic equation. That is, the eigenvalues λ of A are the solutions of:

$$\lambda^3 - I_1 \lambda^2 + I_2 \lambda - I_3 = 0$$

where,

$$\begin{aligned} I_1 &= A_{ii} \\ I_2 &= \frac{1}{2} (A_{ii} A_{jj} - A_{ij} A_{ji}) \\ I_3 &= \det(A) \end{aligned}$$

Another commonly used set of invariants is:

$$\begin{aligned} I_A &= A_{ii} \\ II_A &= \frac{1}{2} (A_{ii} A_{jj} - A_{ij} A_{ji}) \\ III_A &= \det(A). \end{aligned}$$

The Cayley-Hamilton theorem, and the result on the number of independent scalar invariants allows us to write any analytic tensor function F of a tensor A as,

$$F(A) = \gamma_1(I_1, I_2, I_3)I + \gamma_2(I_1, I_2, I_3)A + \gamma_3(I_1, I_2, I_3)A^2.$$

Suppose instead that a tensor A is a function of some vector r . What can we say about the general form of this function?

Consider two arbitrary vectors a, r . Then,

$$a_i A_{ij} b_j = f(a, b, r),$$

but we know that the scalar function f must be expressed in terms of the invariants that can be formed from a, b, r . These are,

$$|a|, |b|, |r|, a_i b_i, a_i r_i, b_i r_i$$

Furthermore, the left hand side is clearly a bilinear function of a, b , thus we must choose only scalar invariants which are bilinear in a, b ; we are therefore limited to

$$a_i r_i, b_i r_i$$

Which can be multiplied by some function of $|r|$. This implies,

$$a_i A_{ij} b_j = g(|r|) a \cdot b + h(|r|) a \cdot r b \cdot r$$

which in turn implies that,

$$A_{ij} = g(|r|) + h(|r|) r_i r_j.$$

Cross products and tensor consistency. Finally, in dealing with cross products in tensor notation we introduce the "alternating tensor" or "Levi-Cevita Symbol"

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } i, j, k \text{ is an even permutation of } 1, 2, 3 \\ -1 & \text{if } i, j, k \text{ is an odd permutation of } 1, 2, 3 \\ 0 & \text{otherwise.} \end{cases} \quad (2.2.1)$$

The challenge in introducing this tensor is that if we swap any two of our basis vectors we change the sign of this tensor—meaning this tensor is not coordinate system invariant. It is however almost coordinate system invariant, up to a sign, thus we call it a "pseudo-vector." Note however that if we were to apply the tensor again, (introduce another cross product in our term) that the result would again be a vector ($-1^2 = 1$). Coordinate system invariance then imposes the restriction that models for pseudo-vectors must be expressed in terms of pseudo-vectors.

EXAMPLE: Which of the following would be a permissible/valid model?

$$\begin{aligned} a \times b &= c \times d && \leftarrow \\ a &= b \times d && \\ a &= b \times c \times d && \leftarrow \end{aligned}$$

We conclude this section by giving the tensor representation of vector calculus differential operators:

$$\begin{aligned} \text{Gradient:} \quad (\nabla \varphi)_i &= \frac{\partial \varphi}{\partial x_i} \\ \text{Divergence:} \quad (\nabla \cdot \Psi) &= \frac{\partial \Psi_i}{\partial x_i} \\ \text{Curl:} \quad (\nabla \times \Psi)_i &= \epsilon_{ijk} \frac{\partial \Psi_j}{\partial x_k} \end{aligned} \quad (2.2.2)$$

Bottomline. Tensor consistency and dimensional consistency impose significant constraints on model forms.

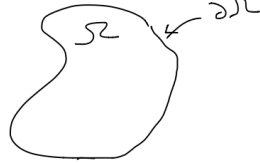
3. Kinematics of Deformable Bodies

3.1 Motivation

Continuum mechanics applies a form of Newton's law to a continuum, rather than discrete masses. Instead of $\mathbf{F} = m\mathbf{a}$ we use:

$$\frac{d\mathbf{M}}{dt} = \mathbf{F}, \quad \text{where } \mathbf{M} = m\mathbf{v}$$

Consider some domain Ω within the material (continuum):



Question 1: *What is the momentum of the material in Ω ? What forces are acting on it?*

The momentum is the sum (integral) of momentum of constituent particles. Two types of forces are considered, body and surface forces.

The analog of $\mathbf{F} = m\mathbf{a}$ (conservation of momentum) is:

$$\frac{\partial}{\partial t} \int_{\Omega} \rho \mathbf{v} \, dx = \int_{\Omega} \mathbf{f}_b \, dx + \int_{\partial\Omega} \mathbf{f}_s \, dA$$

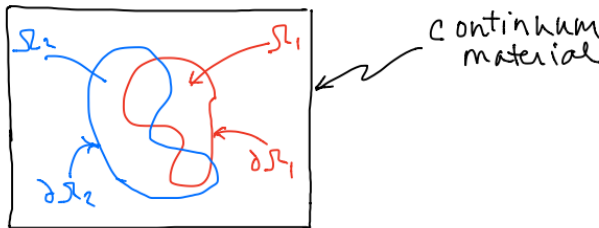
where ρ is the mass density, f_b is the body force per unit volume acting on Ω , and f_s is the surface force per unit area acting on $\partial\Omega$.

Question 2: *What physical phenomena might f_b and f_s represent?*

f_b : gravity,

f_s : stress, pressure,

Consider two subdomains of a continuum material, Ω_1, Ω_2 . Neglecting body forces (their model will depend on corresponding "force fields") we can ask what the momentum of the material is on either subdomain and write the conservation of momentum for each:



$$\begin{aligned} \frac{\partial}{\partial t} \int_{\Omega_1} \rho \mathbf{v} \, dx &= \int_{\partial\Omega_1} \mathbf{f}_{s1} \, dx \\ \frac{\partial}{\partial t} \int_{\Omega_2} \rho \mathbf{v} \, dx &= \int_{\partial\Omega_2} \mathbf{f}_{s2} \, dx \end{aligned}$$

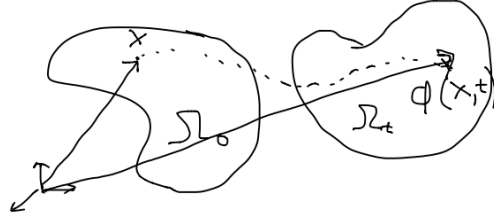
Question 3: *What should f_{s1}, f_{s2} depend on? How would this be different for a fluid or solid? Should f_{s1}, f_{s2} be the same at points where $\partial\Omega_1, \partial\Omega_2$ intersect?*

The key point here is that f_s will depend on the orientation of the curve (surface) $\partial\Omega$ (i.e. the surface normal of the boundary). For solids we care about deformation, for fluids we care about rate of deformation. The surface forces will not be the same where domains intersect since they will depend on the orientation of the boundary.

Models for f_s, f_b are needed to complete the description of the motion of the continuum. We again neglect f_b , focusing only on f_s for now. In particular, we hypothesize that for solids, internal surface forces f_s will depend on deformation—or strain. For fluids, we hypothesize instead that internal surface forces f_s will depend on the *rate* of deformation—or strain-rate. The insight behind such choices will become clear eventually, but for now we build up the analytic tools necessary to define continuum motion, paying particular attention to developing representations for strain and strain-rate that we will use to model surface forces.

3.2 Solids

Material configurations and motion. Consider a deformable body \mathcal{B} . The material points within \mathcal{B} can be labeled by a vector X representing the position with respect to the origin in some "reference configuration" occupying some region $\Omega_0 \subset \mathbb{R}^3$.



The motion and deformation of the body can then be expressed at some time t by a function $x = \varphi(X, t)$. $\varphi(X, t)$ is bijective and it cannot turn the material inside-out, thus

$$\det(\nabla\varphi(X, t)) > 0.$$

The deformation of the body is denoted

Deformation

$$u = \varphi(X) - X \quad (3.2.1)$$

where the dependence of u, φ on t has been neglected. Note that φ is a vector field, so its gradient $\nabla\varphi$ is a rank-2 tensor. This rank-2 tensor will be called the *deformation gradient* and denoted:

Deformation Gradient

$$F(X) = \nabla\varphi(X) = I + \nabla u(X) \quad (3.2.2)$$

or equivalently,

$$F_{ij}(X) = \frac{\partial\varphi_i}{\partial X_j} = \delta_{ij} + \frac{\partial u_i}{\partial X_j}$$

The motion is *rigid* if the body does not deform. This requires

$$\varphi(X) = \mathbf{a} + Qx,$$

where $\mathbf{a} \in \mathbb{R}^3$ is a translation and Q is a unitary tensor—a rigid body rotation.

Question 4: *Should the internal surface forces be affected by rigid body motion? Why or why not?*

No they shouldn't since the continuum is not compressed or altered, just shifted and rotated.

Deformation and Strain. We now consider the deformation of differential line segments. Let dX denote a differential segment in the reference configuration, it will be mapped to

$$dx = FdX.$$

The deformations we are interested in (non-rigid body) will result in changes of the lengths of line segments. Considering the length of these segments,

$$\begin{aligned} dS_0^2 &= |dX|^2 = dX \cdot dX = dX_i dX_i \\ dS^2 &= |dx|^2 = (FdX) \cdot FdX = dX^\top \underbrace{F^\top F}_C dX = dX_i \underbrace{F_{ij}^\top F_{jk}}_{C_{ik}} dX_k \end{aligned}$$

Thus we define the *Cauchy-Green deformation tensor*, a rank-2 tensor C that quantifies "stretching" of continuum in different directions as

Cauchy-Green Deformation Tensor

$$C = F^\top F. \quad (3.2.3)$$

Tensor C is symmetric positive definite. Additionally, we can define the *Green-St. Venant strain tensor* E that denotes the change in length of line segments as

Green-St. Venant Strain Tensor

$$E = \frac{1}{2}(C - I). \quad (3.2.4)$$

Note that tensor C can be viewed as the dilation of a line segment, while E denotes (half) the actual change in length of the segment (e.g. a segment stretched to a ratio of 1.1 times its original size only changed .1 in length—1.1 would be value of C in this direction, .1 would be (twice) the value of E in this direction). Thus if φ is a rigid-body motion, then $E = 0$.

Principle values and directions of deformation. Since E is symmetric it has real eigenvalues and orthogonal eigenvectors. These eigenvectors are the *principle coordinates* or *principle directions* of strain E . In this coordinate system, $E = E_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$ and

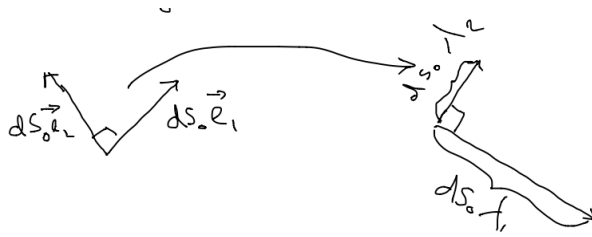
$$E_{ij} = \begin{cases} \frac{1}{2}(\lambda_i^2 - 1) & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (3.2.5)$$

where λ_i^2 are the eigenvalues of C and it is customary to write λ_i^2 since $C = F^\top F$. Since $E = \frac{1}{2}(C - I)$, eigenvectors (principle directions) \mathbf{e}_i are also eigenvectors of C corresponding to λ_i^2 . Consider how the differential line segments in the principle directions transform into the current configuration:

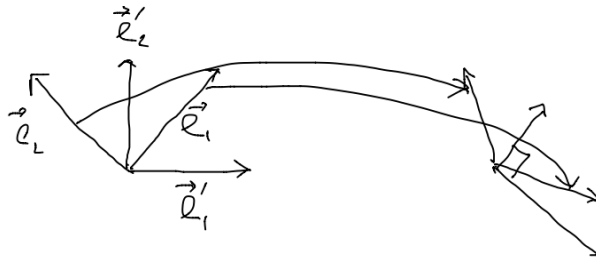
Note in particular that the images of principle directions are also mutually orthogonal.

Question 5: *Why are the images of vectors \mathbf{e}_i also mutually orthogonal?*

By the properties of SVD of C , see conditions for orthogonality of right singular vectors V .



Shear strains (non-principle deformations). Now consider what happens to differential segments aligned with a more general set of orthogonal bases:



The images of $dS_0 \mathbf{e}_i'$, that are not the eigenvectors to C and E are not mutually orthogonal since the projection along direction \mathbf{e}_i is scaled according to λ_i which generally are not equal.

The apparent rotation of the non-principle basis vectors is encoded in the off-diagonal elements of E when expressed in these coordinates. They are referred to as *shear strains* γ_{ij} and can be viewed as a consequence of anisotropic stretching. The following equations for shear strains are not given in the notes but are included here in case they come in handy:

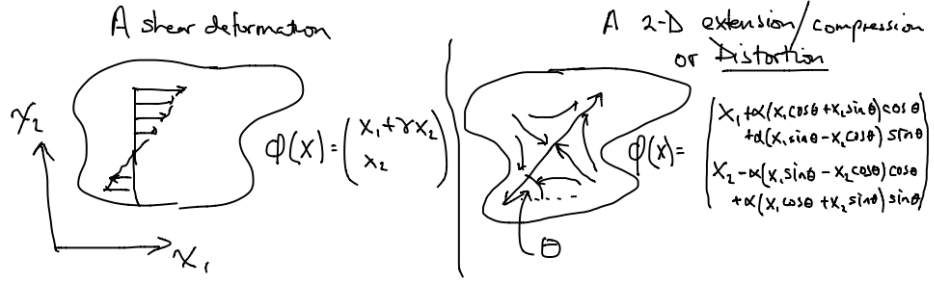
Shear Strain

$$\sin \gamma_{ij} = \frac{2E_{ij}}{\sqrt{1 + 2E_{ii}} \sqrt{1 + 2E_{jj}}}.$$

Angle γ_{ij} measures the deviation of the angle between $\mathbf{e}_i, \mathbf{e}_j$ from perpendicular.

Relating shear and principle strain—the Polar Decomposition Theorem. Consider two deformations:

For a given γ the two can be related by a scalar α and rotation angle θ . In particular, the two deformations will have the same C and E .



Theorem 3.2.1 (Polar Decomposition Theorem). *For any invertible F , \exists unique R, U, V s.t.*

1. R is unitary (i.e. a rotation)
2. U, V are symmetric positive definite
3. $F = RU = VR$.

The proof is outlined on pg. 19 of the book.

When F is the deformation gradient:

Polar Decomposition

$$\begin{aligned} F &= RU = VR \\ C &= F^T F = U^T R^T RU = U^2 \\ B &= FF^T = RVV^T R^T = V^2 \end{aligned}$$

So the left and right deformation gradients do not depend on R . This is clearly the explanation for our example: in the shear case, $R \neq I$; in the distortion case, $R = I$. Thus a pure shear in 2-D is the composition of a 2-D distortion and a rotation.

In general, the Polar Decomposition Theorem implies that any local deformation as characterized by the deformation gradient is composed of a distortion and a rotation—thus DEFORMATION TENSORS AND STRAIN TENSORS ARE INDEPENDENT OF ROTATION.

Question 6: Recall—we started this discussion expecting internal forces to depend on the deformation, is the deformation tensor/strain tensor a good candidate for this dependence?

Yes.

NOTE. This theorem has significant real-world consequences. For example, consider a rod in torsion (think a drive shaft or a bolt with a wrench applied to it). Picture a differential material element on the surface of the rod, it is in nearly pure shear. The principle stresses and directions are then a 45° rotation from the axis of the rod. This is often how failure occurs in shafts under torsion, note the nearly 45° angle of the fault in the following image.



3.3 Fluids

We now turn our attention to modeling internal surface forces in fluids. Note that fluids may undergo deformation without carrying residual stresses into a new reference configuration (e.g. think water bottle being shaken). Thus for fluids we consider the dependence of internal surface forces on the *rate of deformation* instead.

Consider the motion as a function of time:

$$\begin{aligned}
 & \left. \begin{aligned} \dot{x} &= v(x) = \frac{\partial \varphi}{\partial t} \\ \ddot{x} &= v(x) = \frac{\partial^2 \varphi}{\partial t^2} \end{aligned} \right\} \text{Lagrangian (material frame)} \\
 & \left. \begin{aligned} v(x) &= \frac{\partial x}{\partial t} \Big|_x = \dot{x}(\varphi^{-1}(x, t), t) \\ a(x) &= \frac{\partial^2 x}{\partial t^2} \Big|_x = \ddot{x}(\varphi^{-1}(x, t), t) \end{aligned} \right\} \text{Eulerian (spatial frame)}
 \end{aligned} \tag{3.3.1}$$

The Lagrangian frame describes the motion of a point that originated at X throughout time. The Eulerian frame describes the motion of particles passing through point x throughout time.

Now consider the material time derivative of some field quantity Ψ in each frame of reference. Elementary multivariate chain rule gives:

$$\begin{aligned}
 \text{Lagrangian: } \frac{d\Psi_m(X, t)}{dt} &= \frac{\partial \Psi_m}{\partial t} + \frac{\partial \Psi_m}{\partial x_i} \frac{\partial X_i}{\partial t} = \frac{\partial \Psi_m}{\partial t} = \frac{\partial \Psi}{\partial t} \Big|_x \\
 \text{Eulerian: } \frac{d\Psi(x, t)}{dt} \Big|_x &= \frac{\partial \Psi}{\partial t} \Big|_x + \frac{\partial \Psi_m}{\partial x_i} \frac{\partial x_i}{\partial t} \Big|_x = \frac{\partial \Psi}{\partial t} + v_i \frac{\partial \Psi}{\partial x_i} = \frac{\partial \Psi}{\partial t} + v \cdot \text{grad } \Psi
 \end{aligned}$$

We write these relations again for clarity,

$$\begin{aligned}
 \text{Lagrangian: } \frac{d\Psi_m(X, t)}{dt} &= \frac{\partial \Psi_m}{\partial t} \Big|_x \\
 \text{Eulerian: } \frac{d\Psi(x, t)}{dt} \Big|_x &= \frac{\partial \Psi}{\partial t} + v \cdot \text{grad } \Psi
 \end{aligned}$$

In the Eulerian description, it is convenient to have a special nomenclature for $\frac{d}{dt}\Big|_X$, we define:

Total Derivative

$$\frac{D\Psi}{Dt} := \frac{d\Psi}{dt}\Big|_X = \frac{\partial\Psi}{\partial t} + v \cdot \text{grad } \Psi, \quad (3.3.2)$$

as the so-called material time derivative, or substantial derivative.

Quantifying rates of deformation. In the Eulerian description we define the *velocity gradient tensor* L as

Velocity Gradient Tensor

$$L := \text{grad } v \quad (3.3.3)$$

We can consider the rate of change of the (Lagrangian) deformation gradient F now as:

$$\begin{aligned} \dot{F} &= \frac{\partial}{\partial t} \nabla \varphi = \nabla \frac{\partial \varphi}{\partial t} = \nabla v \\ \implies \dot{F}_{ij} &= \frac{\partial v_i}{\partial X_j} = \frac{\partial v_i}{\partial x_j} \frac{\partial x_j}{\partial X_j}, \end{aligned}$$

thus

$$\dot{F} = \text{grad } v F = L_m F \quad (3.3.4)$$

where L_m is L expressed in Lagrangian frame. We can also observe the identity

$$L_m = \dot{F} F^{-1} \quad (3.3.5)$$

We note that L can be decomposed into symmetric part D and anti-symmetric part W , i.e.

$$L = D + W. \quad (3.3.6)$$

where,

$$D = \frac{1}{2} (L + L^\top) \quad (\text{strain-rate tensor}) \quad (3.3.7)$$

$$W = \frac{1}{2} (L - L^\top) \quad (\text{rotation-rate tensor}) \quad (3.3.8)$$

Note that $Wv = \frac{1}{2}\omega \times v$ where $\omega = \text{curl } v$ is the vorticity.

We will now illustrate the reason behind the naming of the symmetric part D (*strain-rate tensor*) and anti-symmetric part W (*rotation-rate tensor*).

Consider a the rate of change of a differential line segment dS , but first note that

$$\begin{aligned} \frac{\partial}{\partial t}(dX \cdot C dX) &= dX \cdot \frac{\partial C}{\partial t} dX = dX \cdot \frac{\partial}{\partial t}(F^\top F) dX \\ &= dX \cdot (\dot{F}^\top F + F^\top \dot{F}) dX \\ &= dX \cdot ((LF)^\top F + F^\top (LF)) dX \\ &= dX \cdot (F^\top (L + L^\top) F) dX. \end{aligned}$$

Then,

$$\begin{aligned}\frac{\partial dS^2}{\partial t} &= dX \cdot [(F^\top(L + L^\top)F)dX] \\ &= (FdX) \cdot [(L + L^\top)(FdX)] \\ &= (dx) \cdot [(L + L^\top)(dx)] = 2(dx) \cdot [Ddx]\end{aligned}$$

Thus the strain-rate tensor is exactly what it claims to be, the rate that segments are stretched (strained). At the beginning of this section we hypothesized the dependence of surface forces on the rate of deformation. It can be shown (see homework) that under the assumption that the current configuration is the reference configuration (usually a good assumption for fluids as the choice of reference configuration is somewhat arbitrary), $D = \dot{E}$. Under these assumptions the tensor D is also rotation independent. The strain-rate tensor D could therefore be a good candidate for modeling such forces (foreshadowing).

In addition to quantifying the strain and rotation rates, we can quantify the rate of volume change as

$$\frac{\dot{\det F}}{\det F} = \det F \operatorname{div} v. \quad (3.3.9)$$

This identity will come in handy shortly.

Divergence theorem. Before continuing we briefly note the Divergence Theorem, which will allow us to transition between volume and flux integrals. Let Ψ represent some vector field, then

Divergence Theorem

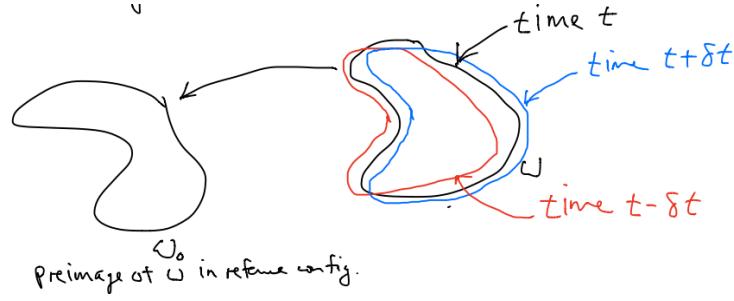
$$\int_{\Omega} \operatorname{div} \Psi \, dx = \int_{\partial\Omega} \Psi \cdot \hat{n} \, dA. \quad (3.3.10)$$

Reynold's Transport Theorem. Conserved quantities like mass, momentum, and energy are carried by the material. In writing conservation laws for continua in an Eulerian description it will be convenient to determine the material derivative of an intensive quantity integrated over some volume. Let Ψ be the quantity of interest. We want to know $\frac{\partial}{\partial t} \int_{\omega_t} \Psi dx$ in some subdomain ω_t corresponding to a subdomain ω_0 in the reference configuration. In particular we note that the preceding integral is difficult to evaluate because the domain of integration is changing with time. The following identity provides the multivariate analogue of the Leibniz rule (for 1D integrals with varying bounds of integration).

The material occupying ω_t at time t was in some region $\omega_0 \in \Omega_0$ in the reference configuration—and the region occupied by this material evolves continually in time as shown.

The time derivative we want is of the material occupying ω_t at time t , but following the material. So we can write,

$$\begin{aligned}\frac{d}{dt} \int_{\omega_t} \Psi dx &= \frac{d}{dt} \int_{\omega_0} \Psi_m \det F dX = \int_{\omega_0} \frac{d}{dt} (\Psi_m \det F) dX \\ &= \int_{\omega_0} \left(\frac{\partial \Psi_m}{\partial t} + v \cdot \operatorname{grad} \Psi_m \right) \det F dX + \int_{\omega_0} \Psi_m \frac{\dot{\det F}}{\det F} dX\end{aligned}$$



The term inside the first integral is obtained for the time derivative of Ψ_m using the multivariate chain rule, as Ψ_m is dependant on variables $(x(t), t)$. Switching back to the Eulerian integral,

$$\begin{aligned} &= \int_{\omega_t} \left(\frac{\partial \Psi}{\partial t} + v \cdot \text{grad } \Psi \right) dx + \int_{\omega_t} \Psi \text{div } v dx \\ &= \int_{\omega_t} \left(\frac{\partial \Psi}{\partial t} + \text{div}(\Psi v) \right) dx \end{aligned}$$

Finally, applying the divergence theorem we get,

Reynold's Transport Equation

$$\begin{aligned} \frac{d}{dt} \int_{\omega_t} \Psi dx &= \int_{\omega_t} \frac{\partial \Psi}{\partial t} dx + \int_{\partial \omega_t} \Psi v \cdot n dx \\ &= \int_{\omega_t} \left(\frac{\partial \Psi}{\partial t} + \text{div}(\Psi v) \right) dx \end{aligned} \quad (3.3.11)$$

This result is known as Reynold's Transport Theorem. The surface integral can be interpreted as the net flux of the Ψ quantity carried by the material across the boundary of ω .

The Piola Transformation. The Piola transform is detailed on pg. 17 of the book and will be considered in more detail later. While we do not go into detail here, we note the main equality and the resulting transformation for surface normals that will allow us to internal surface forces between reference and current configurations.

The following identity holds for tensor fields defined at two times, T_0 at t_0 and T at t :

$$\int_{\partial \Omega_0} T_0(X) \hat{n}_0(X) dA_0 = \int_{\partial \Omega} T(x) \hat{n}(x) dA \quad (3.3.12)$$

if the following equality is satisfied:

$$T_0(X) = [\det F(X)] T(X) F(X)^{-T}.$$

This result can be used to establish the following correspondence between surface normals at different times:

$$\hat{n} = \frac{\text{Cof} F \hat{n}_0}{\|\text{Cof} F \hat{n}_0\|}. \quad (3.3.13)$$

3.4 Summary

Deformation. Quantifies how much points in material have moved between two times.

$$u = \varphi(X) - X$$

Deformation Gradient. The gradient of deformation (plus I), a rank-2 tensor that essentially maps vectors in RC to CC, $dx = FdX$.

$$F(X) = \nabla\varphi(X) = I + \nabla u(X)$$

Cauchy-Green Deformation Tensor. A rank-2 tensor that quantifies how much distances are stretched from RC to CC

$$C = F^T F$$

Green-St. Venant Strain Tensor. A rank-2 tensor that quantifies the change in segment lengths (strain) from RC to CC.

$$E = \frac{1}{2}(C - I)$$

Shear Strain. The angle that basis vectors in RC deviate from perpendicular in CC.

$$\sin \gamma_{ij} = \frac{2E_{ij}}{\sqrt{1 + 2E_{ii}}\sqrt{1 + 2E_{jj}}}$$

Polar Decomposition. Allows deformation gradient tensor to be factored into a unitary tensor (rotation) and symmetric positive definite tensor. The resulting symmetric positive definite tensor U (when this form is used) are the principle directions of deformation, multiplied by diagonal the diagonal matrix of principle values of deformation. Tensor U can be computed by finding the eigen-decomposition of C , then taking the square root of the eigenvalue matrix and multiplying matrices back out.

$$\begin{aligned} F &= RU = VR \\ C &= F^T F = U^T R^T RU = U^2 \\ B &= FF^T = RVV^T R^T = V^2 \end{aligned}$$

Rates of deformation in different frames.

$$\left. \begin{aligned} \dot{x} &= v(x) = \frac{\partial \varphi}{\partial t} \\ \ddot{x} &= \dot{v}(x) = \frac{\partial^2 \varphi}{\partial t^2} \end{aligned} \right\} \text{Lagrangian (material description)}$$

$$\left. \begin{aligned} v(x) &= \dot{x}(\varphi^{-1}(x, t), t) \\ a(x) &= \ddot{x}(\varphi^{-1}(x, t), t) \end{aligned} \right\} \text{Eulerian (spatial description)}$$

Field rates in different frames.

$$\begin{aligned} \text{Lagrangian: } \frac{d\Psi_m(X, t)}{dt} &= \left. \frac{\partial \Psi}{\partial t} \right|_X \\ \text{Eulerian: } \left. \frac{d\Psi(x, t)}{dt} \right|_X &= \frac{\partial \Psi}{\partial t} + v \cdot \text{grad } \Psi \end{aligned}$$

Total Derivative. Short notation of field rate of change in Eulerian frame.

$$\frac{D\Psi}{Dt} := \left. \frac{d\Psi}{dt} \right|_X = \frac{\partial \Psi}{\partial t} + v \cdot \text{grad } \Psi$$

Velocity Gradient Tensor. Rank-2 tensor quantifying deformation (strain and rotation) rate. Its symmetric and anti-symmetric parts are called the strain-rate and rotation-rate tensors respectively.

$$L := \text{grad } v \quad (\text{velocity gradient tensor})$$

$$\dot{F} = \text{grad } v F = L_m F$$

$$L_m = \dot{F} F^{-1}$$

$$D = \frac{1}{2} (L + L^\top) \quad (\text{strain-rate tensor})$$

$$W = \frac{1}{2} (L - L^\top) \quad (\text{rotation-rate tensor})$$

Rate of volume change. Quantifies the rate at which volume of continuum is changing at a given location.

$$\frac{\dot{\det F}}{\det F} = \det F \operatorname{div} v$$

Divergence Theorem. Allows us to switch between volume and flux integrals of vector fields.

$$\int_{\Omega} \operatorname{div} \Psi \, dx = \int_{\partial \Omega} \Psi \cdot \hat{n} \, dA.$$

Reynold's Transport Equation. The equivalent of Liebniz rule in multiple dimensions, allows us to incorporate the rate of change of the domain of integration into integrals.

$$\begin{aligned} \frac{d}{dt} \int_{\omega_t} \Psi \, dx &= \int_{\omega_t} \frac{\partial \Psi}{\partial t} \, dx + \int_{\partial \omega_t} \Psi v \cdot n \, dx \\ &= \int_{\omega_t} \left(\frac{\partial \Psi}{\partial t} \, dx + \operatorname{div}(\Psi v) \right) \, dx \end{aligned}$$

Piola Transform (Corollary). Transformation of surface normals in RC to surface normals in CC.

$$\hat{n} = \frac{\operatorname{Cof} F \, \hat{n}_o}{\|\operatorname{Cof} F \, \hat{n}_o\|}$$

4. Eulerian Conservation (Fluids)

As we return to conservation laws for continua we could proceed in various ways. In particular, each of the conservation laws we derive will have both Eulerian and Lagrangian forms. While it may in some sense be natural to develop these two representations simultaneously, our focus here is on constructing a model for continuum mechanics. For fluid mechanics the choice of reference domain is somewhat arbitrary so we often prefer an Eulerian or spatial representation. For solid mechanics however, material largely maintains the same spatial relation (locally) as some reference configuration and a Lagrangian approach is often desired. For this reason we have decided to present Eulerian conservation laws in this chapter, and defer Lagrangian conservation for the next. Note however that the derivations in this chapter are not necessarily limited to fluids until we begin constructing models and assume dependencies.

This chapter proceeds as follows: We begin by developing Eulerian conservation of momentum to derive overarching fluid mechanical equations. In defining momentum conservation we introduce the concept of surface stresses σ , and we deviate slightly from rigorous derivation of conservation to define a model for surface stresses that will lead to the famous Navier-Stokes equations for Newtonian fluids. We then return to conservation to develop Eulerian representations of mass conservation and energy conservation. These later conservation laws lead to equations commonly studied under the guise of thermodynamics in engineering fields. In particular we hope to emphasize the inextricable connection between these fields through conservation laws.

4.1 Momentum

Question 1: *What is momentum?*

Recalling our formulation of Newton's law in a continuum however we can instead write $\frac{dm\mathbf{v}}{dt} = \mathbf{F}$, thus it can also be seen as the integral of force. In particular when no force is applied momentum remains constant—Newton's first law. It is this relation that we seek to conserve when we speak of momentum conservation.

Conservation of momentum is the fundamental relationship that describes motion of a body under the action of force. The momentum of a body, denoted $I(\mathcal{B}, t)$, is given by

$$I(\mathcal{B}, t) = \int_{\Omega_t} \varrho \mathbf{v} dx.$$

From Newton's law,

$$\frac{dI(\mathcal{B}, t)}{dt} = \frac{d}{dt} \int_{\Omega_t} \varrho \mathbf{v} dx = \int_{\Omega_t} \varrho \frac{d\mathbf{v}}{dt} dx = F_{net}, \quad (4.1.1)$$

where F_{net} are the net forces on body \mathcal{B} . As described in the beginning of the previous chapter, these forces are *body forces* (F_b)—exerted volumetrically, and *surface forces* (F_s)—exerted on the boundary of a region. Writing this relationship as before (in the previous chapter) we have,

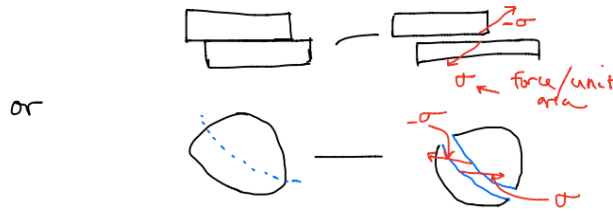
$$\frac{\partial}{\partial t} \int_{\Omega_t} \varrho \mathbf{v} dx = \mathbf{F}_b + \mathbf{F}_s = \int_{\Omega_t} \mathbf{f}_b dx + \int_{\partial\Omega_t} \mathbf{f}_s dA \quad (4.1.2)$$

Body Forces. Let $\mathbf{f}(\mathbf{x}, t)$ be the force per unit volume, then the body force F_b is

$$\mathbf{F}_b = \int_{\Omega_t} \mathbf{f}_b(\mathbf{x}, t) dx.$$

EXAMPLE: for gravity, $\mathbf{f} = \varrho \mathbf{g}$.

Surface Forces. Forces exerted on external surfaces may have different origins but they enter the formulation in the same way. Let σ be the force per unit area acting on the body occupying Ω_t .



For external forces, we can consider this to be imposed at the surface of an actual body. Momentum must be conserved everywhere, so (4.1.2) must hold for any subdomain Ω of the

material, not just for those domains whose boundaries align with physical boundaries. Thus in addition to imposed forces, we will have internal forces—that can be imposed on the many surfaces passing through any material point. Note in particular that these internal forces depend not only on the point at which a surface passes through, but the orientation of the surface passing through that point (as can be reasoned by considering material anisotropy and the varying force responses obtained in different orientations).

The following hypothesis and theorem identify a model (dependencies) for surface stresses σ .

Cauchy Hypothesis: There is a vector field $\sigma(\mathbf{x}, t, \hat{n})$ that defines the force/area (stress) at point x in the current configuration, which depends on the normal (outward pointing) \hat{n} to the surface $\partial\Omega$. Newton's laws imply that

$$\sigma = \sigma(\mathbf{x}, t, \hat{n}) = -\sigma(\mathbf{x}, t, -\hat{n}).$$

Theorem 4.1.1 (Cauchy Stress Theorem). *Assume the following conditions:*

1. *body forces f_b continuous on Ω*
2. *$\sigma(\mathbf{x}, t, \hat{n})$ is continuously differentiable w.r.t. \hat{n} at constant x*
3. *$\sigma(\mathbf{x}, t, \hat{n})$ is continuously differentiable w.r.t. x at constant \hat{n}*

Then \exists a tensor field $T(x, t) \ni \sigma(\mathbf{x}, t, \hat{n}) = T(x, t)n$, and $T(x, t) = T(x, t)^T$ where $T(x, t)n$ is a rank-2 tensor that maps normal vector to force per unit area on surface of Ω .

The proof of the Cauchy Stress Theorem can be found on pages 37-38 of the book, but the basic idea is to consider a differential tetrahedron and apply conservation of momentum and conservation of angular momentum.

In short, this theorem allows us to represent the surface stress σ as the product of a (rank-2) tensor $T(x, t)$ and surface normal \hat{n} . Surface forces F_s then take the form

$$\mathbf{F}_s = \int_{\partial\Omega_t} T(x, t) \cdot \hat{n} dA_t.$$

While we yet lack a model for the tensor T , in terms of the displacements and displacement rates that were the subject of the previous chapter we begin to have some idea of the form such a model might take. We will develop this model in time—for fluids in this chapter and for solids in the next—but first we further develop momentum conservation (4.1.2).

Eulerian (*Spatial*) Conservation of Momentum. Substituting the Cauchy stress tensor T representation of surface forces and applying the Divergence Theorem equation (4.1.2) can now be written,

$$\frac{d}{dt} \int_{\Omega} \rho \mathbf{v} dx = \int_{\Omega} \mathbf{f}_b dx + \int_{\Omega} \operatorname{div} T dx.$$

We wish to derive a differential form of this relationship, thus aim to combine each of these integrands under a single integral. We can achieve this by applying Reynold's transport theorem to the term on the right hand side yielding:

$$\int_{\Omega} \frac{\partial(\varrho \mathbf{v})}{\partial t} + \operatorname{div}(\varrho \mathbf{v} \otimes \mathbf{v}) \, dx = \int_{\Omega} \mathbf{f}_b \, dx + \int_{\Omega} \operatorname{div} T \, dx$$

and combining integrands,

Momentum Conservation (Eulerian)

$$\int_{\Omega} \left[\frac{\partial(\varrho \mathbf{v})}{\partial t} + \operatorname{div}(\varrho \mathbf{v} \otimes \mathbf{v}) - \mathbf{f}_b - \operatorname{div} T \right] dx = 0. \quad (4.1.3)$$

Since Ω was arbitrary, this immediately implies the differential form,

Differential Momentum Conservation (Eulerian)

$$\frac{\partial(\varrho \mathbf{v})}{\partial t} + \operatorname{div}(\varrho \mathbf{v} \otimes \mathbf{v}) - \mathbf{f}_b - \operatorname{div} T = 0. \quad (4.1.4)$$

This is the Eulerian representation (since Ω was the current configuration domain). As tensor $\varrho \mathbf{v} \otimes \mathbf{v}$ is not symmetric however there is some ambiguity to how the divergence operator should be applied. Writing this relationship instead in Cartesian tensor notation removes the ambiguity:

Differential Momentum Conservation (Eulerian)

$$\frac{\partial \varrho v_i}{\partial t} + \frac{\partial \varrho v_i v_j}{\partial x_j} - f_{bi} - \frac{\partial T_{ij}}{\partial x_j} = 0. \quad (4.1.5)$$

We still lack a model for T . However, once we have developed such a model we will see that under certain assumptions (4.1.5) will take the form of the famous Navier-Stokes equations. We turn our attention to developing such a model.

Modeling the Cauchy Stress Tensor T . In Chapter 3 we hypothesized that the internal surface forces in fluids are dependent on the rate of deformation or strain. The reasoning for this hypothesis is that fluids can be deformed from a starting configuration and will again settle in a different configuration with no tendency to return to the starting configuration. Also in Chapter 3 we derived a relationship for the so-called *strain-rate* tensor D . Such a tensor is independent of rotations of the underlying coordinate system and is a good candidate to model surface stresses σ in the context of fluids. We assume therefore the dependence of T on D ,

$$T = T(D).$$

Thus we have a rank-2 tensor T as a function of another rank-2 tensor D . Without assuming anything about the structure of either quantity we can apply the Cayley-Hamilton Theorem (discussed in Chapter 2) to get a general form of the relationship as:

$$T = aI + bD + cD^2$$

where a, b, c are scalar functions of the three scalar invariants of D . It would be difficult to quantify all these functions experimentally or otherwise, so we make the following assumption:

ASSUME: T is a linear function in D .

This assumption leads to the definition of *Newtonian Fluids*, fluids where internal stresses depend linearly on strain-rate. This is a fair assumption for a wide class of fluids including water, oil, and others, but is a poor assumption for others including some important fluids including blood. However, models for these non-Newtonian fluids often rely on very similar assumptions, we will point out where one might modify this assumption to include a broader class of fluids.

From the assumption of linearity, we can immediately remove the dependence of T on c , as the D^2 term is not linear. Similarly, since scalar invariants of D depend on D , the scalar function b must necessarily be constant. Finally, only one scalar invariant of D is linear in D , i.e. $\text{tr}[D]$, so function a must be a scalar multiple of $\text{tr}[D]$, possibly plus some constant. This gives the form,

$$T = \alpha + \beta \text{tr}[D] + \gamma D$$

where α, β, γ are constants (in D). We can rewrite this expression in the form

$$T = P + \kappa \text{tr}[D]I + 2\mu \left(D - \frac{\text{tr}[D]}{3}I \right)$$

where constants are now given the names,

P — Hydrodynamic pressure

κ — Bulk viscosity

μ — Shear viscosity

The bulk viscosity κ is commonly assumed to be zero. This is the standard model for internal stresses of Newtonian fluids; substituting this model for T into (4.1.5) gives the Navier-Stokes equations as promised. To modify this model for non-Newtonian fluids, a common practice is to make μ a function of an invariant of D . The class of fluids that this extends to are known as shear-thinning or shear-thickening fluids, since the shear-viscosity is a function of shear-rate.

If we were to substitute T into (4.1.5) note that we would have three equations, one for each component of velocity (plus three equations imposing the symmetry of T)—but we have a number of unknowns including: $v_i, \varrho, P, \kappa, \mu$. It turns out that parameters μ will depend on additional variables like temperature T , but this will be dealt with later when we consider energy conservation. The point here is that our system is severely underdetermined. We now derive additional equations for conservation of mass and conservation of energy but note that this pattern of underdetermined-ness will persist, necessitating constitutive equations (the topic of Chapter 6).

4.2 Mass

We consider the conservation of mass first as it is fairly straightforward to derive.

Question 2: *What is mass? How is it measured?*

A measure of how much stuff is in an area. It can be measured as density times volume.

Consider a body \mathcal{B} occupying a domain Ω_0 in the reference configuration. Its mass is given by

$$M_0(\mathcal{B}) = \int_{\Omega_0} \varrho_0 dX,$$

where ϱ_0 is the mass density field in the reference configuration. At some time t the same body occupies Ω_t and has mass density $\varrho(x, t)$. At time t , the mass is

$$M(\mathcal{B}) = \int_{\Omega_t} \varrho(x, t) dx. \quad (4.2.1)$$

Question 3: *How are M_0 and M_t related? Why?*

They are the same since the mass of particles does not change and Ω_t tracks particles through space.

We then have that

$$\int_{\Omega_0} \varrho_0(X) dX = \int_{\Omega_t} \varrho(x) dx \quad (4.2.2)$$

Until now we have not limited ourselves to a Lagrangian or Eulerian frame and the preceding analysis holds for both. We again defer the Lagrangian formulation for now, as the Eulerian formulation is often desired for fluids.

Eulerian (*Spatial*) Conservation of Mass. The *Eulerian* (spatial) conservation of mass can be obtained by observing that (4.2.2) implies that

$$\frac{d}{dt} \int_{\Omega_t} \varrho(x) dx = 0$$

Recalling Reynold's transport theorem for a general field Ψ ,

$$\frac{\partial}{\partial t} \int_{\omega} \Psi dx = \int_{\omega} \frac{\partial \Psi}{\partial t} dx + \int_{\partial \omega} \Psi v \cdot n dx$$

and applying it to the density field $\varrho(x)$ we get the spatial description of conservation of mass,

$$\int_{\Omega_t} \left[\frac{\partial \varrho(x)}{\partial t} + \operatorname{div}(\varrho(x)v) \right] dx = 0,$$

and since Ω_t is arbitrary we get,

Mass Conservation (Eulerian)

$$\frac{\partial \varrho(x)}{\partial t} + \operatorname{div}(\varrho(x)v) = 0. \quad (4.2.3)$$

To better understand the meaning of this expression, we can re-express it using the divergence theorem on some spatial region ω .

$$\underbrace{\int_{\omega} \frac{\partial \varrho(x)}{\partial t} dx}_{\text{rate of change of mass in } \omega} = - \underbrace{\int_{\partial \omega} \varrho v \cdot \hat{n} dA}_{\text{rate at which mass leaves } \omega}$$

4.3 Energy

Another conserved quantity of interest is the energy. Let ϵ be the total energy per unit volume. Then energy conservation will follow the general outline of :

$$\frac{d}{dt} \int_{\Omega_t} \epsilon dx = \text{Sources}.$$

Question 4: *What constitutes the energy of a continuum? What are the possible sources?*

The energy consists of kinetic energy κ and internal energy U . Possible sources include power, or rate of work denoted \mathcal{P} and heating rate \dot{Q} .

Thus we have simply that

$$\frac{d}{dt} \int_{\Omega_t} (\kappa + U) = \mathcal{P} + \dot{Q}.$$

Writing out this relationship in a concrete form we get,

$$\frac{d}{dt} \int_{\Omega_t} \overbrace{\left(\frac{\rho \mathbf{v} \cdot \mathbf{v}}{2} + \rho e \right)}^{\epsilon} dx = \overbrace{\int_{\Omega_t} \mathbf{f} \cdot \mathbf{v} dx + \int_{\partial \Omega_t} \boldsymbol{\sigma} \cdot \mathbf{v} dA}^{\mathcal{P}} + \overbrace{\int_{\Omega_t} r dx - \int_{\partial \Omega_t} \mathbf{q} \cdot \hat{n}}^{\dot{Q}},$$

where, e is the internal energy per unit volume, r is the volumetric heating rate, and q is the heat flux (with \mathbf{q} pointing into the volume, or opposing \hat{n}). We begin to analyze this by expanding the stress term in \mathcal{P} :

$$\begin{aligned} \int_{\partial \Omega_t} \boldsymbol{\sigma} \cdot \mathbf{v} dA &= \int_{\partial \Omega_t} T \hat{n} \cdot \mathbf{v} dA = \int_{\partial \Omega_t} v_i T_{ij} n_j dA \\ &= \int_{\Omega_t} \frac{\partial v_i T_{ij}}{\partial x_j} dx = \int_{\Omega_t} \left(v_i \frac{\partial T_{ij}}{\partial x_j} + T_{ij} \frac{\partial v_i}{\partial x_j} \right) dx \\ &= \int_{\Omega_t} (v \cdot \operatorname{div} T + T : \operatorname{grad} v) dx \\ &= \int_{\Omega_t} (v \cdot \operatorname{div} T + T : D) dx, \end{aligned}$$

where $T : D$ denotes the contraction of T and D , or the trace of their product. so,

$$\mathcal{P} = \int_{\Omega_t} \mathbf{v} \cdot (\operatorname{div} T + \mathbf{f}) dx + \int_{\Omega_t} T : D dx.$$

But, from momentum conservation,

$$\operatorname{div} T + \mathbf{f} = \rho \frac{d\mathbf{v}}{dt} + \operatorname{div}(\rho \mathbf{v} \otimes \mathbf{v})$$

so,

$$\mathcal{P} = \int_{\Omega_t} \mathbf{v} \cdot \left(\rho \frac{d\mathbf{v}}{dt} + \operatorname{div}(\rho \mathbf{v} \otimes \mathbf{v}) \right) dx + \int_{\Omega_t} T : D dx$$

which, by expanding terms, adding the Eulerian conservation of mass equation, refactoring gives, and finally applying Reynold's transport theorem to the field $\Psi = \frac{\rho \mathbf{v} \cdot \mathbf{v}}{2}$

$$\mathcal{P} = \frac{d}{dt} \int_{\Omega_t} \frac{\rho \mathbf{v} \cdot \mathbf{v}}{2} dx + \int_{\Omega_t} T : D dx.$$

Note now that the left side in this term is equal to the kinetic energy term in ε . The equality between these terms is known as the conservation of kinetic energy, or the kinetic energy equation and is redundant with momentum conservation, so we can cancel these terms. To get \dot{Q} into a form amenable to our analysis we simply apply the divergence theorem to the right most term, giving

$$\dot{Q} = \int_{\Omega_t} r - \operatorname{div} \mathbf{q} \, dx.$$

Combining now all these terms into the conservation of energy equation, and cancelling the kinetic energy terms as noted we arrive at the internal energy equation,

Internal Energy Equation (Eulerian)

$$\frac{d}{dt} \int_{\Omega_t} \varrho e \, dx = \int_{\Omega_t} (T : D + r - \operatorname{div} \mathbf{q}) \, dx \quad (4.3.1)$$

which enforces the first law of thermodynamics. Applying Reynolds transport theorem on the left we arrive at the differential form of (Eulerian) internal energy conservation,

Differential form of Internal Energy Conservation (Eulerian)

$$\frac{\partial \varrho e}{\partial t} + \operatorname{div}(\varrho e \mathbf{v}) = T : D + r - \operatorname{div} \mathbf{q} \quad (4.3.2)$$

Thermodynamics and Entropy. As any student of thermodynamics will know however, there is also a second law of thermodynamics which must be satisfied. This law—that the entropy of a closed system cannot not decrease—is not a conservation equation, but rather a constraint on the internal energy equation. Entropy, a measure of the microscale randomness in a system, is a somewhat non-intuitive concept compared to other quantities we have thus far encountered. We neglect a rigorous definition here, however a simple explanation of this law is that things do not become more ordered absent some input of energy. Physical manifestations of this law are everywhere: gas fills a room on its own, liquids mix, and heat flows from hot to cold regions. Here we are particularly interested in entropy in constraining the direction of heat flow.

From thermodynamics we have the notion of "Thermodynamic state variables"

Temperature (absolute) – θ

Entropy – S

If we denote Q as the heat added to a system at a constant temperature θ , between times $t_2 > t_1$ then the second law of thermodynamics implies,

$$S(t_2) - S(t_1) - \frac{Q}{\theta} \geq 0.$$

When equality holds a process is *reversible*. The concept of reversibility is key in determining theoretical limits of performance and underlies measures of efficiency in engineering. For the time evolution we have,

2nd Law of Thermodynamics (Entropy Equation)

$$\frac{dS}{dt} - \frac{\dot{Q}}{\theta} \geq 0 \quad (4.3.3)$$

For a continuum, we define the variable η as the entropy per unit mass, thus

$$S = \int_{\Omega_t} \varrho \eta \, dx$$

and,

$$\frac{d}{dt} \int_{\Omega_t} \varrho \eta \, dx + \int_{\partial\Omega_t} \frac{\mathbf{q} \cdot \hat{\mathbf{n}}}{\theta} \, dA - \int_{\Omega_t} \frac{r}{\theta} \, dx \geq 0.$$

Finally, applying the divergence theorem to the surface integral, and applying Reynold's transport theorem we have,

Entropy Equation (Continuum, Eulerian)

$$\int_{\Omega_t} \left[\frac{\partial \varrho \eta}{\partial t} + \operatorname{div}(\varrho \eta \mathbf{v}) + \operatorname{div} \frac{\mathbf{q}}{\theta} - \frac{r}{\theta} \right] \, dx \geq 0. \quad (4.3.4)$$

and, since the domain Ω_t is arbitrary we get,

Clausius Duhem Inequality (Eulerian)

$$\frac{\partial \varrho \eta}{\partial t} + \operatorname{div}(\varrho \eta \mathbf{v}) + \operatorname{div} \frac{\mathbf{q}}{\theta} - \frac{r}{\theta} \geq 0. \quad (4.3.5)$$

Note that in defining our system of conservation equations we have introduced many more unknowns than equations. We will eventually remedy this in Chapter 6 where we develop constitutive or closure equations for our system. Note that up until this point we have made very few assumptions, only in defining our model for surface stresses σ and in defining our model for the resulting Cauchy stress tensor T . Thus, while we noted this representation is especially useful for fluids, it is valid for any continuum material.

4.4 Summary

Newton's law for Continua. Expresses conservation of momentum

$$\underbrace{\frac{\partial}{\partial t} \int_{\Omega_t} \rho \mathbf{v} dx}_{ma, \text{ momentum rate}} = \underbrace{\int_{\Omega_t} \mathbf{f}_b dx}_{\text{body forces}} + \underbrace{\int_{\partial\Omega_t} \mathbf{f}_s dA}_{\text{surface forces}}$$

Cauchy Stress Tensor. Surface stresses (force / unit area) represented by rank-2 symmetric Cauchy Stress tensor T . See associated theorem for assumptions on when this holds.

$$\mathbf{F}_s = \int_{\partial\Omega_t} T(x, t) \cdot \hat{n} dA_t$$

$$T = T^T \quad (\text{by conservation of angular momentum})$$

Conservation Laws.

- **Momentum Conservation (Eulerian).** Derived from Newton's law, $F = ma$ expressed in terms of continuum.

$$\underbrace{\frac{\partial(\rho \mathbf{v})}{\partial t}}_{\text{Momentum rate}} + \underbrace{\text{div}(\rho \mathbf{v} \otimes \mathbf{v})}_{\text{Momentum flux (Transport)}} - \underbrace{\mathbf{f}_b}_{\text{Body forces}} - \underbrace{\text{div} T}_{\text{Surface forces}} = 0$$

$$\frac{\partial \rho \mathbf{v}_i}{\partial t} + \frac{\partial \rho \mathbf{v}_i \mathbf{v}_j}{\partial x_j} - \mathbf{f}_{bi} - \frac{\partial T_{ij}}{\partial x_j} = 0 \quad (\text{Tensor Form})$$

- **Mass Conservation (Eulerian).**

$$\underbrace{\frac{\partial \rho(x)}{\partial t}}_{\text{Mass rate}} + \underbrace{\text{div}(\rho(x) \mathbf{v})}_{\text{Mass flux (Transport)}} = 0$$

- **Internal Energy Conservation (Eulerian).** Enforces first law of thermodynamics. We start with general expression,

$$\frac{d}{dt} \int_{\Omega_t} \left(\underbrace{\kappa}_{\text{Kinetic energy}} + \underbrace{U}_{\text{Internal energy}} \right) = \underbrace{\mathcal{P}}_{\text{Rate of work (Power)}} + \underbrace{\dot{Q}}_{\text{Heating rate}}$$

Expanding we get,

$$\frac{d}{dt} \int_{\Omega_t} \overbrace{\left(\frac{\rho \mathbf{v} \cdot \mathbf{v}}{2} + \rho e \right)}^{\epsilon \text{ (Total Energy / Unit Vol)}} dx = \overbrace{\int_{\Omega_t} \mathbf{f} \cdot \mathbf{v} dx + \int_{\partial\Omega_t} \boldsymbol{\sigma} \cdot \mathbf{v} dA}^{\mathcal{P} \text{ (Power)}} + \overbrace{\int_{\Omega_t} r dx - \int_{\partial\Omega_t} \mathbf{q} \cdot \hat{n}}^{\dot{Q} \text{ (Heating rate)}}$$

where,

e = Internal energy per unit volume

r = volumetric heating rate

q = heat flux

Kinetic energy and boundary work (second power term) redundant with momentum conservation so we take them out to get conservation of internal energy.

$$\underbrace{\frac{\partial \rho e}{\partial t}}_{\text{Int. energy rate}} + \underbrace{\text{div}(\rho e \mathbf{v})}_{\substack{\text{Int. energy flux} \\ \text{(Transport)}}} = \underbrace{T : D}_{\substack{\text{Rate of work} \\ \text{done by stress}}} + \underbrace{r}_{\text{Vol. heating rate}} - \underbrace{\text{div} \mathbf{q}}_{\substack{\text{Heat flux} \\ \text{(Transport)}}}$$

- **Clausius Duhem Inequality (Eulerian).** Enforces second law of thermodynamics.

$$\underbrace{\frac{\partial \rho \eta}{\partial t}}_{\text{Entropy rate}} + \underbrace{\text{div}(\rho \eta \mathbf{v})}_{\substack{\text{Entropy flux} \\ \text{Transport}}} + \underbrace{\text{div} \frac{\mathbf{q}}{\theta}}_{\substack{\text{Heat flux} \\ \text{Generation}}} - \underbrace{\frac{r}{\theta}}_{\text{Generation}} \geq 0$$

5. Lagrangian Conservation (Solids)

We have already defined conservation equations for momentum, mass, and energy in an Eulerian frame. We now turn to defining similar conservation laws for a Lagrangian frame that are more convenient in dealing with solids.

5.1 Mass

We begin with conservation of mass due to its simplicity. We start with relationship (4.2.2), given here for reference

$$\int_{\Omega_0} \varrho_0(X) dX = \int_{\Omega_t} \varrho(x) dx.$$

The integral in x can be transformed into an integral in X since $x = \varphi(X)$ and $dx = \det F dX$. The factor $\det F$ can be understood as the ratio of differential volume dx to dX . We thus have that

$$\int_{\Omega_0} (\varrho_0(X) - \varrho(\varphi(X)) \det F) dX = 0$$

From which we can conclude (since Ω_0 is arbitrary) that

Mass Conservation (Lagrangian)

$$\varrho_0(X) = \varrho(\varphi(X)) \det F(X). \tag{5.1.1}$$

This is the Lagrangian or *material* description of the conservation of mass.

5.2 Momentum

We give the Eulerian conservation equation again here for reference,

$$\int_{\Omega} \frac{\partial(\varrho \mathbf{v})}{\partial t} + \operatorname{div}(\varrho \mathbf{v} \otimes \mathbf{v}) - \mathbf{f}_b - \operatorname{div} T \, dx = 0.$$

To obtain a Lagrangian representation of momentum conservation equivalent to (4.1.3), we can perform a change of variable, i.e. "transform" from x to X . For the first three terms of (4.1.3) this is a straight forward process, again multiplying by a factor $\det F$ to transform differential volumes dx to differential volumes dX . However, for the last term in (4.1.3) we need a way to relate the changing orientation and size (stretching) of differential areas. The goal is to find a tensor field $T_0(X) \ni \operatorname{Div} T_0 = \det F \operatorname{div} T$. The Piola transform presented briefly at the end of Chapter 3 provides the appropriate tool to accomplish this objective.

Piola Transform. The derivation of the Piola transform is given on pages 16-18 of the book. The result is, if

$$T_0(X) = [\det F(X)] T(X) F(X)^{-\top},$$

then

Piola Transform

$$\operatorname{Div} T_0 = \operatorname{div} T \det F.$$

Observe,

$$\int_{\Omega_0} \operatorname{Div} T_0 \, dX = \int_{\Omega_0} \operatorname{div} T \det F \, dX = \int_{\Omega} \operatorname{div} T \, dx.$$

and by the divergence theorem,

$$\int_{\partial\Omega_0} T_0 \hat{n}_0 \, dA_0 = \int_{\partial\Omega} T \hat{n} \, dA$$

Tensor T is a linear transformation from \hat{n} (CC) to a CC force/CC area. Tensor T_0 is a linear transformation of a normal \hat{n}_0 (RC) to a CC force/RC area. Since $T_0 = \operatorname{Cof} F$, $\|\operatorname{Cof} F\|$ is the area scaling from the reference configuration to the current configuration. The direction of $\operatorname{Cof} F \hat{n}_o$ is then in the direction of \hat{n} , and

$$\hat{n} = \frac{\operatorname{Cof} F \hat{n}_o}{\|\operatorname{Cof} F \hat{n}_o\|}.$$

Tensor T_0 is given the symbol $P(X)$, the first Piola-Kirchhoff Stress Tensor,

1st Piola-Kirchhoff Stress Tensor

$$P = \det F T F^{-\top} = T \operatorname{Cof} F \tag{5.2.1}$$

We now have the tools to rewrite the conservation of momentum equation (4.1.3) in Lagrangian form,

Momentum Conservation (Lagrangian)

$$\varrho_0 \frac{\partial^2 u}{\partial t^2} = \text{Div} P + f_0 \quad (5.2.2)$$

where,

$$\begin{aligned} \varrho_0 &= \varrho(\varphi(X)) \det F \\ f_0 &= f(\varphi(X)) \det F. \end{aligned}$$

But note, tensor P is not symmetric (since F and therefore its cofactor are not symmetric). But, since $P = \det F T F^{-\top}$, it follows that $P F^T = \det F T$ is symmetric (since $\det F$ is a scalar and T is symmetric). A symmetric tensor is also recovered if we map the CC force/RC area to an RC force/RC area. This is accomplished by pre-multiplication of F^{-1} , and we define the second Piola-Kirchhoff Stress Tensor S in this way.

2nd Piola-Kirchhoff Stress Tensor

$$S = F^{-1} P = \det F F^{-1} T F^{-\top} \quad (5.2.3)$$

In terms of S , momentum conservation becomes,

Momentum Conservation (in terms of S)

$$\varrho_0 \frac{\partial^2 u}{\partial t^2} = \text{Div} \underbrace{F S}_P + f_0 \quad (5.2.4)$$

The tensor S is a linear transformation of a normal \hat{n}_0 (RC) to an RC force/RC area (fully RC map).

5.3 Energy

We conclude with the derivation of Lagrangian energy conservation (and entropy equation). We proceed by treating each of the following terms in order.

$$\int_{\Omega_t} \underbrace{\frac{\partial \varrho e}{\partial t} + \operatorname{div}(\varrho e \mathbf{v})}_{(i)} dx = \int_{\Omega_t} \underbrace{T : D}_{(ii)} + \underbrace{r}_{(iii)} - \underbrace{\operatorname{div} \mathbf{q}}_{(iv)} dx$$

(i) Observe,

$$\begin{aligned} \frac{\partial \varrho e}{\partial t} + \operatorname{div}(\varrho e \mathbf{v}) &= \varrho \left(\frac{\partial e}{\partial t} + \operatorname{div}(e \mathbf{v}) \right) \\ &= \varrho \left(\frac{\partial e}{\partial t} + \mathbf{v} \cdot \operatorname{grad} e \right) \\ &= \varrho \frac{De}{dt} = \varrho \dot{e} \\ &= \varrho_0 \dot{e}_0 \end{aligned}$$

where the integrals have been neglected and,

$$\begin{aligned} \varrho_0 &= \varrho(\varphi(X)) \det F \\ e_0 &= e(\varphi(X)). \end{aligned}$$

Note that e_0 is not scaled by the volume ratio $\det F$ as it is measured per unit mass not per unit volume.

(ii)

$$\begin{aligned} \int_{\Omega_t} T : D dx &= \int_{\Omega_0} T : \operatorname{grad} v \det F dX \\ &= T_{ij} \frac{\partial^2 \varphi_i}{\partial X_k \partial t} \frac{\partial X_k}{\partial x_j} \det F \\ &= T_{ij} \dot{F}_{ik} F_{kj}^{-1} \det F \\ &= \dot{F}_{ik} (T_{ij} F_{jk}^{-\top} \det F) \\ &= \dot{F}_{ik} P_{ik} = \dot{F}_{ik} F_{ij} S_{jk} \\ &= (F^\top \dot{F}) : S \\ &= \dot{E} : S = S : \dot{E} \end{aligned}$$

where the integrals have again been neglected after the first line.

(iii)

$$r = r_0 \det F$$

(iv) See homework,

$$\operatorname{div} \mathbf{q} = \operatorname{Div} \mathbf{q}_0$$

where,

$$\mathbf{q}_0 = (\operatorname{Cof} F)^\top \mathbf{q}$$

Substituting the preceding relationships we arrive at the Lagrangian conservation of internal energy,

Energy Conservation (Lagrangian)

$$\varrho_0 \dot{e}_0 = S : \dot{E} - \text{Div } \mathbf{q}_0 + r_0 \quad (5.3.1)$$

Finally, the entropy equation can be transformed by combining trivial analogues of the preceding transformations,

Clausius Duhem Inequality (Lagrangian)

$$\varrho_0 \dot{\eta}_0 + \text{Div } \frac{\mathbf{q}_0}{\theta} - \frac{r_0}{\theta} \geq 0 \quad (5.3.2)$$

5.4 Summary

Stress Tensors. All rank-2 tensors mapping surface normals to stresses (forces per unit areas), but configurations in which these are defined differ.

- **Cauchy Stress Tensor.** (CC) surface normal to (CC) force / (CC) area (thus typically used in Eulerian frame).

$$T = T^\top$$

- **1st Piola-Kirchhoff Stress Tensor.** (RC) surface normal to (CC) force / (RC) area.

$$P = T_0 = \det F T F^{-\top} = T \text{Cof } F$$

$$P \neq P^\top$$

- **2nd Piola-Kirchhoff Stress Tensor.** (RC) surface normal to (RC) force / (RC) area.

$$S = F^{-1} P = \det F F^{-1} T F^{-\top}$$

$$S = S^\top$$

Piola Transform. The Piola tranform is essentially a transformation of surface normal direction, and corresponding area scaling. Its problem is essentially to find a tensor T_0 s.t.

$$\text{Div} \underbrace{T_0}_P = \text{div } T \det F.$$

This is achieved if

$$\underbrace{T_0}_P = \det F T F^{-\top} = T \text{Cof } F$$

We also get the following identity for transforming surface normals from (RC) to (CC)

$$\hat{n} = \frac{\text{Cof } F \hat{n}_o}{\|\text{Cof } F \hat{n}_o\|}$$

Conservation Laws. All these were derived by applying appropriate transformations to Eulerian representations.

- **Momentum Conservation (Lagrangian).**

$$\underbrace{\rho_0 \frac{\partial^2 u}{\partial t^2}}_{\text{Momenutum rate}} = \underbrace{\text{Div} \overbrace{FS}^P}_{\substack{\text{Surface forces} \\ (\text{Stress})}} + \underbrace{f_0}_{\text{Body forces}}$$

- **Mass Conservation (Lagrangian).** This one is particularly boring, just expresses that density changes with expansion or compression.

$$\rho_0(X) = \rho(\varphi(X)) \det F(X)$$

- **Energy Conservation (Lagrangian).** Enforces first law of thermodynamics.

$$\underbrace{\rho_0 \dot{e}_0}_{\text{Int. energy rate}} = \underbrace{S : \dot{E}}_{\substack{\text{Rate of work} \\ \text{done by stress}}} - \underbrace{\text{Div } \mathbf{q}_0}_{\text{Heat flux}} + \underbrace{r_0}_{\text{Vol heating rate}}$$

- **Clausius Duhem Inequality (Lagrangian).** Enforces second law of thermodynamics.

$$\rho_0 \dot{\eta}_0 + \text{Div } \frac{\mathbf{q}_0}{\theta} - \frac{r_0}{\theta} \geq 0$$

Exercise

Suppose the second Piola-Kirchoff Stress tensor is given by,

$$S = \lambda I \nabla \cdot \mathbf{u} + \mu (\nabla \mathbf{u} + \nabla \mathbf{u}^T).$$

- Express S in index notation.
- Assuming no body forces and starting with conservation of momentum, derive an equation for the evolution of \mathbf{u} .
- Show that $\mathbf{u} = f(\mathbf{x} \cdot \mathbf{e} - \alpha t)\mathbf{e}$ for any sufficiently smooth f and any unit vector \mathbf{e} is a solution. What is the value of α (pressure wave speed)?
- Show that $\mathbf{u} = f(\mathbf{x} \cdot \mathbf{e} - \beta t)\mathbf{g}$ for any \mathbf{g} such that $\mathbf{g} \cdot \mathbf{e} = 0$. What is the value of β (shear wave speed)?

Solution:

- Simple application of definitions

$$\begin{aligned} S_{ij} &= \lambda \delta_{ij} \frac{\partial u_k}{\partial x_k} + \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) && \text{(tensor notation)} \\ &= \lambda \delta_{ij} u_{k,k} + \mu (u_{i,j} + u_{j,i}). && \text{(engineering notation)} \end{aligned}$$

- Assuming $\mathbf{f}_b = 0$, Lagrangian conservation of momentum can be written as,

$$\begin{aligned} \rho_0 \frac{\partial^2 u}{\partial t^2} &= \nabla \cdot (F_{ij} S_{jk}) \\ &= \left[(\delta_{ij} + u_{i,j}) (\lambda \delta_{ij} u_{\ell,\ell} + \mu (u_{i,j} + u_{j,i})) \right]_{,k} \end{aligned}$$

since \mathbf{u} is assumed to be small we can neglect quadratic terms

$$\begin{aligned} &= \left[(\delta_{ij} + \cancel{u_{i,j}}) (\lambda \delta_{jk} u_{\ell,\ell} + \mu (u_{j,k} + \cancel{u_{k,j}})) \right]_{,k} \\ &= \left[\lambda \delta_{ik} u_{\ell,\ell} + \mu (u_{i,k} + u_{k,i}) \right]_{,k} \\ &= \underbrace{\lambda u_{\ell,\ell k}}_{u_{\ell,k\ell}} + \mu (u_{i,kk} + u_{k,ik}) \\ &= (\lambda + \mu) u_{k,ik} + \mu u_{i,kk} \\ &= (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla \cdot (\nabla \mathbf{u}). \end{aligned}$$

- First let \mathbf{e}, \mathbf{g} be arbitrary vectors and let $\mathbf{u} = f(\mathbf{x} \cdot \mathbf{e} - \alpha t)\mathbf{g}$. Then by application of the chain rule,

$$\begin{aligned} u_{i,j} &= f'(\mathbf{x}, t) g_i e_j \\ u_{i,ji} &= f''(\mathbf{x}, t) g_i e_j e_i = f''(\mathbf{x}, t) (\mathbf{e} \cdot \mathbf{g}) \mathbf{e} \end{aligned} \quad \leftarrow$$

similarly,

$$u_{i,jj} = f''(\mathbf{x}, t)g_i e_j e_j = f''(\mathbf{x}, t)g_i = f''(\mathbf{x}, t)\mathbf{g} \quad \leftarrow$$

for this part, $\mathbf{g} = \mathbf{e}$ which, by substituting into our relationship from part (b) and performing differentiation with respect to t gives,

$$\rho_0 \alpha^2 f''(\mathbf{x}, t)\mathbf{g} = (\lambda + 2\mu)f''(\mathbf{x}, t)\mathbf{g}$$

which finally gives,

$$\alpha = \sqrt{\frac{\lambda + 2\mu}{\rho_0}}.$$

(d) This follows from part (c) by letting $\mathbf{g} \cdot \mathbf{e} = 0$. Then, $u_{i,ji} = 0$ immediately implies,

$$\beta = \sqrt{\frac{\mu}{\rho_0}}.$$

Exercise

A cube of incompressible material of dimension a is subjected to a uniaxial load in the x_1 direction, resulting in a deformation that reduces the dimension in the x_1 direction to $\alpha^2 a$, with $0 < \alpha < 1$. The material has an incompressible neo-Hookean strain energy density function:

$$W(\mathbf{E}) = \mu I_{\mathbf{E}}$$

where \mathbf{E} is the strain tensor, $I_{\mathbf{E}}$ is its first invariant, and μ is the shear modulus.

- (a) Determine the deformation gradient tensor (\mathbf{F}), the Cauchy-Green deformation tensor (\mathbf{C}) and the Green-St. Venant strain tensor (\mathbf{E}) in the material.
- (b) How much work was required to deform the material?
- (c) Because the material is incompressible, there is a degeneracy in the stress relation, so that the second Piola-Kirchhoff stress is given by

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{E}} + p \mathbf{F}^{-1} \mathbf{F}^{-\top}$$

where p can be determined from boundary conditions, in this case on the unloaded sides. Determine \mathbf{S} .

- (d) What is the total load force?

Solution:

- (a)
- (b)
- (c)
- (d)

6. Constitutive Equations (Closure Models)

Consider the Eulerian conservation equations (note that this form of Eulerian conservation is obtained by applying product rule and applying conservation of mass, it does not simply assume ϱ is time independent as this is not generally true):

Eulerian Conservation

Mass	$\frac{\partial \varrho}{\partial t} + \text{div}(\varrho \mathbf{v}) = 0$
Lin. Momentum	$\varrho \left(\frac{\partial \langle \mathbf{v} \rangle}{\partial t} + \mathbf{v} \cdot \text{grad } \mathbf{v} \right) = \text{div } T + \mathbf{f}_b$
Ang. Momentum	$T = T^T$
Energy	$\varrho \left(\frac{\partial e}{\partial t} + \mathbf{v} \cdot \text{grad } e \right) = T : D + r - \text{div } \mathbf{q}$
Entropy	$\varrho \left(\frac{\partial \eta}{\partial t} + \mathbf{v} \cdot \text{grad } \eta \right) + \text{div } \frac{\mathbf{q}}{\theta} - \frac{r}{\theta} \geq 0$

Given initial conditions, we want to solve these equations for evolution of variables describing state of the continuum (body). Note however that we have far more unknowns than equations.

Counting unknowns we have,

ϱ	\mathbf{v} (or u)	T	\mathbf{f}	e	\mathbf{q}	\mathbf{r}	η	θ	Total
1	3	6	3	1	3	3	1	1	20 16

where the crossed out variables are not counted because they are determined by external data and other physical phenomenon (e.g. gravity, E&M, chemistry), and the conservation of angular momentum equation implying the symmetry of T is used to reduce the number of unknowns in T to 6.

Now counting equations we note,

Mass	Momentum	Energy	Total
1	3	1	5

Thus we clearly require additional relationships. We begin by identifying quantities that characterize the state of the body (5 we hope) that we will call state variables. A convenient choice is,

Solids: ϱ, \mathbf{u} ($v = \dot{u}$), θ (or e)

Fluids: $\varrho, \mathbf{v}, \theta$ (or e)

We need to formulate relationships to determine remaining quantities from these. These relationships are:

- determined by properties of the material
- models of phenomena (atomic) not represented in the continuum approximation

We need constitutive models for:

- T – inner atomic forces and/or momentum diffusion
- \mathbf{q} – heat conduction
- e – thermodynamics and inner atomic forces
- η – thermodynamics

In addition to dimensional and coordinate invariance, we will require that our constitutive models follow these principles:

- 1.) **Determinism** – no dependence on the future state.
- 2.) **Material Frame Indifference (MFI)** – invariance to changes in the reference frame, i.e. no dependence on rigid motions (uniform velocity or rotation)
- 3.) **Physical Consistency** – cannot violate conservation laws or Clausius Duhem (2nd law)
- 4.) **Material Symmetry** – If the material is invariant to a group of "unimodular" transformations (e.g. rotations; reflections; continuous rotation group or isotropy) then constitutive model must also be invariant to these transformations
- 5.) **Local Action** – Constitutive model depends only on "local" state; essentially depends on the state and finite number of spatial derivatives at a point.
- 6.) **Dimensional Consistency** – dimensional and coordinate invariance
- 7.) **Other Considerations** – well-posedness, Equipresence

6.1 Application of MFI to T for Solids.

Two observers can be moving or rotating differently. The motion observed by one observer, $x(t)$ is related to the motion observed by the second observer, $x^*(t)$ through,

$$x^*(t) = \underbrace{Q(t)}_{\text{rotation}} x(t) + \underbrace{c(t)}_{\text{translation}}$$

where Q and c do not depend on x . Since rotations are unitary transformations, it must be that

$$Q^\top(t) = Q^{-1}(t).$$

We expect (hypothesize) the Cauchy stress tensor T to depend on x , X , and t (or equivalently u, t). But, note, to have T independent of $c(t)$ we instead need to consider derivatives of u , i.e. F and other higher order derivatives.

Suppose $Q = I$, $c \neq 0$, $c(0) = 0$. Then MFI and locality principles require that

$$T = \mathcal{T}(F, \text{higher derivatives}, X)$$

and where we have assumed that T depends only on the first derivative of u , F . Thus we have,

$$T(X, \tau) = \mathcal{T}(F(X, \tau), X) \quad \text{for } \tau \geq t$$

Now **Suppose $Q \neq I$.** Observe,

$$F^* = QF \quad \text{and} \quad \det F^* = \det F$$

with

$$n^* = Qn \implies \sigma^*(n^*) = Q\sigma(n).$$

We know that $\sigma(n) = Tn$ and $\sigma^*(n^*) = T^*n^*$ thus we get that

$$\sigma^*(n^*) = T^*n^* = Q\sigma(n) = QTn = QT \underbrace{Q^\top n^*}_n$$

which implies that,

MFI Constraint

$$T^* = QTQ^\top. \tag{6.1.1}$$

But note, this implies that the relationship

$$T = \mathcal{T}(F)$$

must have the property that

MFI Constraint (Solids)

$$\mathcal{T}(\underbrace{QF}_{F^*}) = Q \mathcal{T}(F) Q^\top \tag{6.1.2}$$

for all $Q(t)$.

Notice that for $C = F^\top F$,

$$C^* = \underbrace{F^\top Q^\top}_{F^{*\top}} \underbrace{Q F}_{F^*} = C.$$

We need to satisfy (6.1.2) and will thus make an educated guess that the relation takes the form,

$$\mathcal{T}(F) = F \hat{\mathcal{T}}(C) F^\top. \quad (6.1.3)$$

Then

$$\mathcal{T}(QF) = QF \hat{\mathcal{T}}(C) F^\top Q^\top = Q \mathcal{T}(F) Q^\top$$

satisfies (6.1.2) so is consistent with MFI.

Finally, because of the polar decomposition theorem, $F = RU$ where R is orthonormal and U is symmetric and positive definite. So MFI implies,

$$\mathcal{T}(U) = \mathcal{T}(R^\top F) = R^\top \mathcal{T}(F) R \quad \implies \quad \mathcal{T}(F) = R^\top \mathcal{T}(U) R$$

Recalling that $C = U^2$, $R = FU^{-1}$, and $R^\top = U^{-1}F^\top$ we have that

$$\mathcal{T}(F) = F \underbrace{U^{-1} \mathcal{T}(C^{\frac{1}{2}}) U^{-1}}_{\hat{\mathcal{T}}(C)} F^\top = F \hat{\mathcal{T}}(C) F^\top$$

where the last equality follows from (6.1.3) and since $U = C^{\frac{1}{2}}$ we can write,

$$\hat{\mathcal{T}}(C) = \left(C^{\frac{1}{2}}\right)^{-1} \mathcal{T}\left(C^{\frac{1}{2}}\right) \left(C^{\frac{1}{2}}\right)^{-1}$$

so this is the ONLY form satisfying MFI.

In this last expression we have expressed U as $C^{\frac{1}{2}}$ to make the dependence on C explicit as hypothesized in (6.1.3). We have finally that any constitutive relationship for T depending on F must take the form (expressed now in terms of U) of

MFI Solids

$$T = \mathcal{T}(F) = FU^{-1}\mathcal{T}(U)U^{-1}F^\top \quad (6.1.4)$$

6.2 Restricted Classes of Constitutive Relations

Thermo-elastic. For this class of constitutive relations we assume that quantities T, \mathbf{q}, e, η at a point depend only on present values of the following state variables at the point:

Solids: $F, \theta, \nabla \theta$

Fluids: $D, \theta, \nabla \theta, \rho$

Homogeneity. For this class of constitutive relations we assume we have the same material everywhere (reference configuration uniform) so no explicit dependence on X .

6.3 Application of MFI to Fluids

A fluid is a material for which dependence of the Cauchy stress tensor T on deformation limited to dependence on $\det F$. As a consequence, the reference configuration becomes irrelevant. Also since $\varrho_0 = \det F \varrho$, the dependence on $\det F$ is already included with dependence on ϱ in an Eulerian representation.

An **ideal fluid** (also inviscid fluid) is one in which T is isotropic, i.e.

Stress in Ideal Fluid

$$T = -P(x, t)I \quad (6.3.1)$$

which clearly satisfies MFI requirements since I commutes with tensor operations (note dependence of ϱ, θ suppressed here). From thermodynamics, a constitutive model for P in terms of ϱ at θ is an "equation of state."

In a viscous fluid, $T = -PI + \mathcal{F}(L)$ where $L = \text{grad } v$ and dependence on ϱ, θ has been suppressed.

MFI Considerations. Observe,

$$L = \dot{F}F^{-1} = \text{grad } v$$

and

$$L^* = \dot{F}^*(F^*)^{-1} = QLQ^T + \Omega$$

where $\Omega = \dot{Q}Q^T$ is the *rotation rate*. But note,

$$\Omega + \Omega^T = \dot{Q}Q^T + Q\dot{Q}^T = \overline{\dot{Q}Q^T} = \dot{I} = 0$$

thus,

$$\Omega = -\Omega^T. \quad (\text{Antisymmetric})$$

Considering MFI we can see that

MFI Fluid

$$\mathcal{F}(\overbrace{QLQ^T + \Omega}^{L^*}) = Q \mathcal{F}(L) Q^T$$

Thus, by eliminating Ω in the equation above and (and recalling $D = \frac{1}{2}(L + L^T)$) we see (this follows because applying MFI to $D = \frac{1}{2}(L + L^T)$ gives $D^* = QDQ^T$)

$$\mathcal{F}(L) = \mathcal{F}(D)$$

We then need that

MFI Fluid (Stokes)

$$\mathcal{F}(\overbrace{QDQ^T}^{D^*}) = Q \mathcal{F}(D) Q^T$$

for an isotropic fluid—this is a Stokes fluid.

If we further assume that \mathcal{F} is *linear* in D we obtain the general form for a Newtonian Fluid,

Stress in Newtonian Fluid

$$T = -PI + 2\mu \left(D - \frac{1}{3} \text{tr}[D]I \right) + \kappa \text{tr}[D]I$$

where,

P – pressure

μ – viscosity (shear viscosity)

κ – bulk viscosity (commonly assumed to be 0)

and P, μ, κ can depend on ϱ, θ . The bulk viscosity represents the irreversibility of volume change and is typically small for fluids at reasonable conditions (thus assumed to be zero).

Special case of incompressible flow. Valid in the limit as $\frac{\|v\|}{c} \rightarrow 0$ with c is the speed of sound. In this limit:

$$\text{div } \mathbf{v} = 0 \quad \implies \quad \text{tr}[D] = 0$$

$$\varrho = \text{constant}$$

$$T = -P_{in} + T_{\text{visc}} \quad \text{where} \quad \text{tr}[T_{\text{visc}}] = 0$$

where constitutive relation needed for T_{visc} , and P_{in} determined by $\text{div } v = 0$. For a Newtonian fluid, $T_{\text{visc}} = 2\mu D$.

6.4 2nd Law Consistency

We must adhere to the second law of thermodynamics, i.e. satisfy the Clausius-Duhem inequality,

$$\varrho \frac{\partial \eta}{\partial t} + \operatorname{div} \frac{\mathbf{q}}{\theta} - \frac{r}{\theta} \geq 0$$

For solids (Coleman-Noll). One approach to enforce the 2nd law in solids is the so-called Coleman-Noll approach.

Consider the Helmholtz free energy

$$\Psi = e - \theta \eta.$$

We develop a relation for Ψ by subtracting the conservation equation for e from θ times the conservation constraint (Clausius-Duhem) for η :

$$\begin{aligned} & - \left(\varrho \frac{\partial e}{\partial t} - T : D - r + \operatorname{div} \mathbf{q} = 0 \right) \\ & + \left(\varrho \theta \frac{\partial \eta}{\partial t} + \theta \operatorname{div} \frac{\mathbf{q}}{\theta} - r \geq 0 \right) \\ \hline & - \varrho \frac{d\Psi}{dt} - \varrho \eta \frac{d\theta}{dt} + T : D - \frac{\mathbf{q}}{\theta} \cdot \operatorname{grad} \theta \geq 0 \end{aligned}$$

or, in expressed in the reference configuration

$$- \varrho_0 \dot{\Psi}_0 - \varrho_0 \eta_0 \dot{\theta} + S : \dot{E} - \frac{\mathbf{q}_0}{\theta} \cdot \nabla \theta \geq 0. \quad (6.4.1)$$

Suppose,

$$\Psi_0 = \psi(E, \theta, \nabla \theta)$$

Then

$$\dot{\Psi}_0 = \frac{\partial \psi}{\partial E} : \dot{E} + \frac{\partial \psi}{\partial \theta} \dot{\theta} + \frac{\partial \psi}{\partial \nabla \theta} \cdot \nabla \dot{\theta}.$$

Substituting this into (6.4.1) we obtain that,

$$\left(S - \varrho_0 \frac{\partial \psi}{\partial E} \right) : \dot{E} - \varrho_0 \left(\frac{\partial \psi}{\partial \theta} + \eta_0 \right) \dot{\theta} - \varrho_0 \frac{\partial \psi}{\partial \nabla \theta} \cdot \nabla \dot{\theta} - \frac{\mathbf{q}_0}{\theta} \cdot \nabla \theta \geq 0$$

However, notice that this must hold for arbitrary sign of $\dot{E}, \dot{\theta}, \nabla \dot{\theta}$, so it is convenient to let coefficient on these be zero,

$$S = \varrho_0 \frac{\partial \psi}{\partial E}, \quad \eta_0 = - \frac{\partial \psi}{\partial \theta}, \quad \frac{\partial \psi}{\partial \nabla \theta} = 0.$$

Then we have simply that

$$- \frac{\mathbf{q}_0}{\theta} \cdot \nabla \theta \geq 0.$$

Rearranging and switching back to the current configuration we see,

$$-\mathbf{q} \cdot \operatorname{grad} \theta \geq 0$$

Thus heat must flow from hot to cold.

In this case where $S = \varrho_0 \frac{\partial \psi}{\partial E}$ the strain is non-dissipative—i.e. work done on the body to affect deformation can be recovered.

More generally S may depend also on \dot{E} , which will lead to dissipation. Then,

$$S = \mathcal{F}(E) + \mathcal{I}(\dot{E})$$

where

$$\mathcal{F}(E) = \varrho_0 \frac{\partial \psi}{\partial E}$$

and

$$\mathcal{I}(\dot{E}) : \dot{E} - \frac{\mathbf{q}_0}{\theta} \cdot \nabla \theta \geq 0.$$

Before considering 2nd law consistency for fluids we note that the above assumption that

$$S = \varrho_0 \frac{\partial \psi}{\partial E}$$

implies that instead of identifying a tensor function to model S we may instead be able to identify scalar function. In particular the previous equality assumes that material deformations are reversible. In the case of material deformations in the elastic regime this is fair assumption. We will return to this later.

For fluids. For a viscous fluid in the Eulerian representation, we assume ψ depends on $\theta, \text{grad } \theta$ only. That is, $\psi(\theta, \text{grad } \theta)$. We have,

$$T = \mathcal{I}(D)$$

where,

$$\mathcal{I}(D) : D - \frac{q}{\theta} \cdot \text{grad } \theta \geq 0.$$

Then

$$\eta = -\frac{\psi}{\theta} \qquad \frac{\partial \psi}{\partial \text{grad } \theta} = 0$$

leads to $\mu, \kappa > 0$.

We return to the governing equation for a Newtonian fluid. The Newtonian constitutive law for T is

$$T = -P I + 2\mu \tilde{D} + \kappa \text{tr}[D]$$

where $\tilde{D} = D - \frac{1}{3} \text{tr}[D] I$ is the deviatoric part of D (constructed such that $\text{tr}[\tilde{D}] = 0$). Most often we have the following dependencies,

$$\begin{array}{lll} \mu = \mu(\theta) & \kappa = 0 & P = P(\varrho, \theta) \\ & & \text{e.g. } P = \varrho R \theta, \text{ (ideal gas)} \end{array}$$

We also have the Fourier heat conduction law,

Fourier's Law

$$\mathbf{q} = -k \text{grad } \theta \tag{6.4.2}$$

where $k = k(\theta)$ is the *thermal conductivity* of the material.

For a *thermally perfect* material, we have that $e = e(\theta)$ and

Specific Heat

$$\frac{de}{d\theta} = C_\nu \quad (6.4.3)$$

If we additionally assume the material is *Calorically perfect*, i.e.

$$C_\nu = \text{constant}$$

we get

$$e = C_\nu \theta.$$

Substituting these into the Eulerian equations we get the Navier-Stokes equations,

Compressible Navier-Stokes Equations

$$\begin{aligned} \frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho \mathbf{v}) &= 0 \\ \varrho \frac{\partial \mathbf{v}}{\partial t} + \varrho \mathbf{v} \cdot \operatorname{grad} \mathbf{v} &= -\operatorname{grad} P + \operatorname{div}(2\mu \tilde{D}) + \operatorname{grad}(\kappa \operatorname{div} \mathbf{v}) \\ \varrho C_\nu \frac{\partial \theta}{\partial t} + \varrho C_\nu \mathbf{v} \cdot \operatorname{grad} \theta &= -P \operatorname{div} \mathbf{v} + 2\mu \tilde{D} : D \\ &\quad + \kappa (\operatorname{div} \mathbf{v})^2 + \operatorname{div}(k \operatorname{grad} \theta) \end{aligned} \quad (6.4.4)$$

This system is now closed if we have constitutive relations for parameters μ, κ, k and an equation of state (e.g. ideal gas law) for P . The following dependencies are commonly used: $P(\theta, \varrho), \mu(\theta), k(\theta), \kappa = 0$.

Additionally, to be able to solve this set of differential equations we need initial and boundary conditions.

Initial Conditions include:

$$\varrho(x, 0) = \varrho_0(x) \quad \mathbf{v}(x, 0) = \mathbf{v}_0(x) \quad \theta(x, 0) = \theta_0(x)$$

Boundary Conditions include:

$$\left. \begin{aligned} (PI + \kappa \operatorname{div} \mathbf{v} I + 2\mu \tilde{D}) n &= \mathbf{g}(x, t) \\ \mathbf{v} &= \mathbf{v}(x, t) \end{aligned} \right\} \text{or} \quad \left. \begin{aligned} -\kappa \nabla \theta \cdot n &= g(x, t) \\ \theta &= \theta(x, t) \end{aligned} \right\} \text{or}$$

In many situations it has been observed that *no-slip* boundary conditions, enforcing zero velocity at walls, are roughly ideal. This can be seen by considering how particles might

bounce off of a rough surface—when the surface is rough there will be some plane close to the wall where the average tangential velocity of molecules will be zero.

Another boundary condition, considered more physically accurate than the no-slip boundary condition is the *robin* boundary condition. This is given by,

$$\mathbf{v}_{\text{tan,wall}} = \underbrace{\ell_s}_{\text{slip-length}} \frac{\partial \mathbf{v}_{\text{tan}}}{\partial n} \approx 0.$$

Equations (6.4.4) can be written in many other equivalent forms. For example:

Conservative form of momentum equation

$$\frac{\partial(\varrho \mathbf{v})}{\partial t} + \text{div}(\varrho \mathbf{v} \otimes \mathbf{v}) = -\text{grad } P + 2 \text{div}(\mu \tilde{D}) + \text{grad}(\kappa \text{div } \mathbf{v})$$

Conservative form of energy equation

$$\frac{\partial(\varrho E)}{\partial t} + \text{div}(\varrho \mathbf{v} E) = -\text{div}(P \mathbf{v}) + 2 \text{div}(\tilde{D} \mathbf{v}) + \text{div}(\kappa \text{grad } \theta)$$

where, $E = \frac{1}{2} \mathbf{v} \cdot \mathbf{v}$. These forms are convenient in numerical simulations for example, where these forms may be better suited to ensure conservative laws.

The Navier-Stokes (N-S) equations describe a wide variety of observed phenomena in fluids including: turbulence, sound waves—compression waves traveling at speed a , where $a^2 = \left. \frac{\partial P}{\partial \varrho} \right|_s$. It has been common practice in engineering to use the solution of these equations to design aircraft, etc.

Incompressible Formulation. These equations can be simplified for an approximately incompressible fluid, but before we dive into that analysis, consider what it means for a fluid to be incompressible. In fact no fluid is incompressible, there are simply conditions under which $\varrho \approx c$ (c is a constant) and this approximation can be made. For example air can be considered incompressible in most analysis of land vehicles where maximum velocities are not close to the speed of sound. This condition, having maximum velocities significantly less than the speed of sound is the condition for which incompressibility is considered a valid assumption.

We will use asymptotic analysis to formulate the incompressible N-S equations. But first, to get a handle on relative scales of variables we need to non-dimensionalize the equations. If we non-dimensionalize these equations the natural scaling is,

$$\begin{aligned} \hat{x} &= \frac{x}{\delta} & \hat{\mathbf{v}} &= \frac{\mathbf{v}}{\mathbf{v}_0} & \hat{\varrho} &= \frac{\varrho}{\varrho_0} \\ \hat{P} &= \frac{P}{\varrho_0 a_0^2} & \hat{t} &= \frac{t \mathbf{v}_0}{\delta} & \hat{\theta} &= \frac{\theta}{\theta_0} \end{aligned} \tag{6.4.5}$$

where δ is the length scale and a is the speed of sound, and superscript 0 indicates that property is taken at some reference thermodynamic state. Note that typically we would non-dimensionalize using v_0 instead of a_0 as the units are the same however a_0 is more convenient due to the ideal gas law and since,

$$P = \varrho R \theta \propto \varrho a^2 \quad \text{since,} \quad a^2 = \gamma R \theta.$$

Rewriting the momentum equation we have,

$$\hat{\varrho} \frac{\partial \hat{\mathbf{v}}}{\partial \hat{t}} + \hat{\varrho} \hat{\mathbf{v}} \cdot \widehat{\text{grad}} \hat{\mathbf{v}} = \frac{1}{M^2} \widehat{\text{grad}} \hat{\mathbf{P}} + \frac{1}{\text{Re}} \widehat{\text{div}} \hat{T}_{\text{visc}} \quad (6.4.6)$$

where $M = \frac{\mathbf{v}_0}{a_0}$ (Mach number) and $\text{Re} = \frac{\varrho_0 \mathbf{v}_0 \delta}{\mu_0}$ (Reynold's number). In the limit as $M \rightarrow 0$, for the pressure gradient term to be finite we must have

$$\hat{\mathbf{P}} = \underbrace{\hat{\mathbf{P}}_0(t)}_{\text{ord. 0 \& indepnt. of } x} + \underbrace{M^2 \hat{\mathbf{P}}'(x, t)}_{\text{ord. 1 pressure fluctuates}}$$

Note that because the first term is independant of x , it vanishes when the gradient is applied. Equations of state $\mathbf{P} = \mathbf{P}(\varrho, \theta)$ (e.g. ideal gas $\mathbf{P} = \varrho R \theta$) then suggest that

$$\begin{aligned} \hat{\theta} &= 1 + M^2 \theta'(x, t) \\ \hat{\varrho} &= 1 + M^2 \varrho'(x, t) \end{aligned}$$

though other behaviors are possible—leading to different equations.

Keeping terms of order 1 only in (6.4.6) gives,

$$\hat{\varrho} \frac{\partial \hat{\mathbf{v}}}{\partial \hat{t}} + \hat{\varrho} \hat{\mathbf{v}} \cdot \widehat{\text{grad}} \hat{\mathbf{v}} = -\widehat{\text{grad}} \hat{\mathbf{P}}' + \frac{1}{\text{Re}} \widehat{\text{div}} \hat{T}_{\text{visc}}.$$

For mass conservation,

$$\frac{\partial \hat{\varrho}}{\partial \hat{t}} + \hat{\mathbf{v}} \cdot \widehat{\text{grad}} \hat{\varrho} = -\hat{\varrho} \widehat{\text{div}} \hat{\mathbf{v}} \quad \implies \quad \widehat{\text{div}} \hat{\mathbf{v}} = 0$$

Thus,

$$\begin{aligned} \hat{T}_{\text{visc}} &= 2\hat{\mu} \hat{\hat{D}} + \hat{\kappa} \widehat{\text{div}} \hat{\mathbf{v}} I \\ &= \hat{\mu} \left(\widehat{\text{grad}} \hat{\mathbf{v}} + \widehat{\text{grad}} \hat{\mathbf{v}}^\top \right) + \left(\hat{\kappa} - \frac{2}{3} \hat{\mu} \right) \widehat{\text{div}} \hat{\mathbf{v}} I \end{aligned}$$

Further $\hat{\mu} \rightarrow 1$ and $\widehat{\text{div}}(\widehat{\text{grad}} \hat{\mathbf{v}}^\top) = \widehat{\text{grad}}(\widehat{\text{div}} \hat{\mathbf{v}}) = 0$, therefore

$$\widehat{\text{div}} \hat{T}_{\text{visc}} = \widehat{\text{div}}(\widehat{\text{grad}} \hat{\mathbf{v}}) = \hat{\Delta} \hat{\mathbf{v}},$$

the vector Laplacian.

$$\hat{\varrho} \frac{\partial \hat{\mathbf{v}}}{\partial \hat{t}} + \hat{\varrho} \hat{\mathbf{v}} \cdot \widehat{\text{grad}} \hat{\mathbf{v}} = -\widehat{\text{grad}} \hat{\mathbf{P}}' + \frac{1}{\text{Re}} \hat{\Delta} \hat{\mathbf{v}}.$$

Note that

$$\begin{aligned} \mathbf{P} &= \mathbf{P}_0(t) + \mathbf{P}'(x, t) \\ \implies \hat{\mathbf{P}} &= \frac{\mathbf{P}_0(t)}{\varrho_0 a_0^2} + \frac{\mathbf{P}'_0}{\varrho_0 a^2} \\ \implies \frac{\mathbf{P}'}{\varrho a^2} &= M^2 \hat{\mathbf{P}}' \\ \implies \hat{\mathbf{P}}' &= \frac{\mathbf{P}'}{\varrho v_0^2} \end{aligned}$$

Expressing everything back in dimensional variables and noting that the energy equation is now redundant we have,

Incompressible Navier-Stokes

$$\begin{aligned} \varrho_0 \frac{\partial \mathbf{v}}{\partial t} + \varrho_0 \mathbf{v} \cdot \text{grad } \mathbf{v} &= -\text{grad } P' + \mu_0 \Delta \mathbf{v} + \mathbf{f} \\ \text{div } \mathbf{v} &= 0 \end{aligned} \tag{6.4.7}$$

These are the incompressible Navier-Stokes Equations—note that no constitutive relations for P' are needed. In fact, counting independent variables we can see that pressure P' is more of a reaction to velocity. In fact, when viewed in the context of a mixed method, pressure can be considered a lagrange multiplier term. We still need boundary and initial conditions however.

Initial Conditions:

$$\text{need } \mathbf{v}(x, 0) = \mathbf{v}_0(x) \quad \text{with} \quad \text{div } \mathbf{v}_0 = 0$$

Boundary Conditions: several types for example

1. Inflow – \mathbf{v} is specified
2. Walls – $\mathbf{v} = 0$ (no-slip condition)
3. Free stream – $T_{\text{visc}} n = 0$ (no stress)
4. Outflow – $T_{\text{visc}} n = 0$

6.5 Heat Equation

Take what we did with energy and heat flux, i.e.

$$de = C_\nu d\theta \qquad \mathbf{q} = -k \operatorname{grad} \theta,$$

then if there is no deformation we get,

Heat Equation

$$\varrho C_\nu \frac{\partial \theta}{\partial t} = \operatorname{div}(\kappa \operatorname{grad} \theta) + r \tag{6.5.1}$$

with $k > 0$ required by the 2nd law.

6.6 Elasticity

For a deformable body (solid) with uniform θ , no heat flux ($\mathbf{q} = 0$), homogeneous, and isotropic.

The free energy constitutive dependence simplifies to

$$\Psi = \psi(E).$$

We then call Ψ the *stored energy function* or *strain energy function*. This is a *hyperelastic* constitutive relation.

Since ψ is an isotropic scalar function of a tensor it must take the following form:

$$\psi = W(\overbrace{I_E, II_E, III_E}^{\text{invariants}})$$

where W is called the work function. Then,

$$S = \frac{\partial \psi}{\partial E} = \frac{\partial W}{\partial I_E} \frac{\partial I_E}{\partial E} + \frac{\partial W}{\partial II_E} \frac{\partial II_E}{\partial E} + \frac{\partial W}{\partial III_E} \frac{\partial III_E}{\partial E}.$$

For example, $\frac{\partial E_{ii}}{\partial E_{jk}} = \delta_{jk}$ implies that $\frac{\partial I_E}{\partial E} = I$.

Similarly we should get

$$\frac{\partial II_E}{\partial E} = \text{tr}[E] I - E$$

$$\frac{\partial III_E}{\partial E} = \frac{1}{2} (\text{tr}[E]^2 - \text{tr}[E^2]) I + (E^2 - \text{tr}[E]E)$$

Equations then become,

Elasticity Equations

$$\varrho_0 \frac{\partial^2 u}{\partial t^2} = \text{Div} \left((I + \nabla u) \frac{\partial \psi}{\partial E} \right) + F_0$$

$$E = \frac{1}{2} (\nabla u + \nabla u^\top + \nabla u^\top \nabla u).$$

(6.6.1)

6.7 Linear Elasticity

Assume all displacements are small, then non-linear terms become negligible,

$$E = \frac{1}{2} (\nabla u + \nabla u^\top + \nabla u^\top \nabla u) \approx e = \frac{1}{2} (\nabla u + \nabla u^\top).$$

If $\frac{\partial^2 u}{\partial t^2}$ is zero these are the Lamé equations,

Lamé Equations

$$(\lambda + \mu) \frac{\partial^2 u_k}{\partial X_i \partial X_k} + \mu \frac{\partial^2 u_i}{\partial X_k \partial X_k} + \mathbf{f}_{0i} = \varrho_0 \frac{\partial^2 u_i}{\partial t^2} \quad (6.7.1)$$

These describe many important phenomena such as elastic wave propagation.

Boundary conditions—no displacement, specified traction.

Furthermore, if we assume S is linear in E , then S depends only on the first and second invariants of E only. In general,

$$W = \frac{1}{2} \varepsilon_{ijkl} e_{kl} e_{ij}$$

and

$$S_{ij} = \frac{\partial W}{\partial e_{ij}} = \varepsilon_{ijkl} e_{kl}$$

and symmetry requires that,

$$\varepsilon_{ijkl} = \varepsilon_{jikl} = \varepsilon_{ijlk} = \varepsilon_{klij}$$

thus,

$$S_{ij} = \varepsilon_{ijk\ell} \frac{\partial u_k}{\partial X_\ell}$$

So we have finally that,

Linear Elasticity Equations

$$\varrho_0 \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial X_j} \left(\varepsilon_{ijk\ell} \frac{\partial u_k}{\partial X_\ell} \right) + \mathbf{f}_{0,i} \quad (6.7.2)$$

For an isotropic material, most general form is

$$\varepsilon_{ijk\ell} = \lambda \delta_{ij} \delta_{k\ell} + \mu (\delta_{ik} \delta_{j\ell} + \delta_{i\ell} \delta_{jk})$$

and μ, λ and the Lamé constants

$$\lambda = \frac{\nu E}{(1 + \nu)(1 - 2\nu)} \quad \mu = \frac{E}{2(1 + \nu)}$$

E is Young's Modulus, ν is Poisson's ratio.

Then

$$\begin{aligned} S &= \lambda \operatorname{tr}[E] I + 2\mu E \\ &= \lambda \operatorname{Div} u + \mu (\nabla u + \nabla u^T) \end{aligned}$$

Finally, notice that for $\nu \rightarrow \frac{1}{2}$, then $\lambda \rightarrow \infty$; so for S to remain finite, $\operatorname{tr}[E] \rightarrow 0$ and $\operatorname{Div} u \rightarrow 0$.

What do Young's Modulus and Poisson's ratio mean? Consider a cylinder under uni-axial loading. In the central region, far from clamped ends must have,

$$S_{11} = \frac{F}{A}, \quad S_{22} = S_{33} = 0.$$

Also,

$$\frac{\partial u_2}{\partial x_2} = \frac{\partial u_3}{\partial x_3}$$

due to isotropy. Then,

$$\begin{aligned} S_{11} &= (2\mu + \lambda) \frac{\partial u_1}{\partial X_1} + 2\lambda \frac{\partial u_2}{\partial X_2} \\ S_{22} &= \lambda \frac{\partial u_1}{\partial X_1} + (2\mu + 2\lambda) \frac{\partial u_2}{\partial X_2} = 0 \\ \implies \frac{\partial u_2}{\partial X_2} &= \frac{\lambda}{2\mu + 2\lambda} \frac{\partial u_1}{\partial X_1} = -\nu \frac{\partial u_1}{\partial X_1} \end{aligned}$$

So we can see that Poisson's ratio ν represents the transverse strain, or the amount a material thins in directions perpendicular to the direction of stretching.

We also have,

$$S_{11} = (2\mu + \lambda(1 - 2\nu)) \frac{\partial u_1}{\partial X_1} = E \frac{\partial u_1}{\partial X_1}$$

thus Young's modulus E (different from strain tensor) represents the scaling between stress and strain in linear elasticity.

6.8 Summary

Constitutive Model Constraints:

- 1) **Determinism** – No dependence on the future state.
- 2) **Material Frame Indifference (MFI)** – Invariance to changes in the frame of an observer. In terms of the stress tensor in conservation of momentum this requires,

$$T^* = QTQ^\top$$

Thus the following constraints must hold in their respective settings:

$$\mathcal{T}(F) = FU^{-1}\mathcal{T}(U)U^{-1}F^\top \quad (\text{Solids})$$

$$\mathcal{F}(\overbrace{QLQ^\top}^{L^*} + \Omega) = Q\mathcal{F}(L)Q^\top \quad (\text{Fluids})$$

$$\mathcal{F}(\overbrace{QDQ^\top}^{D^*}) = Q\mathcal{F}(D)Q^\top \quad (\text{Stokes Fluids})$$

MFI is important because the phenomenon we are modeling should not be affected by simply looking at it different.

- 3) **Physical Consistency** – Model cannot violate conservation laws or Clausius Duhem (2nd law). This implies Fourier's law of heat conduction:

$$\mathbf{q} = -k \text{grad } \theta$$

- 4) **Material Symmetry** – If the material is invariant to a group of "unimodular" transformations (e.g. rotations; reflections; continuous rotation group or isotropy) then constitutive model must also be invariant to these transformations.
- 5) **Local Action** – Constitutive model depends only on "local" state; essentially depends on the state and finite number of spatial derivatives at a point.
- 6) **Dimensional Consistency** – Dimensional and coordinate invariance. Important because definition of dimensional units and coordinate basis is arbitrary and our model cannot depend on arbitrary choices.
- 7) **Other Considerations** – Well-posedness, Equipresence.

Navier-Stokes Equations (fluids).

Elasticity (solids).

Exercise

A liquid of constant density ρ is flowing in a straight horizontal rectangular channel in a gravitational field with acceleration of gravity g . Let x be the coordinate along the length of the channel, let $h(x, t)$ be the height of the liquid about the flow of the channel at the point x , which does not vary across the width of the channel, and let $V(x, t)$ be the velocity in the x -direction, which does not vary over the width or over the depth of the liquid. Because nothing varies over the width, the equations to be derived below can be written in terms of the mass and momentum per unit width, in which case the width of the channel never enters.

- (a) What is the total mass of liquid per unit width in the interval $x \in [a, b]$?
- (b) Derive an integral relation expressing conservation of mass (per unit width) in the interval $x \in [a, b]$.
- (c) Use the fact that $[a, b]$ is arbitrary to write a PDE in x and t expressing conservation of mass.
- (d) What is the total x -momentum of the liquid in the interval $x \in [a, b]$?
- (e) Derive an integral relation expressing conservation of x -momentum (per unit width) in the interval $x \in [a, b]$. What quantity needs to be modeled in this relation?
- (f) Use the fact that $[a, b]$ is arbitrary to write a PDE in x and t expressing conservation of x -momentum.
- (g)) In a momentum conservation equation, we expect an internal force to appear. In this case, it would be the x -direction force per unit width f . Determine a model for f , assuming that f does not depend on V . Are there any undetermined constants? Does this model allow you to close the momentum equations derived above?

Solution:

- (a)
- (b)
- (c)
- (d)

7. Electromagnetism

Electromagnetism is a physical phenomenon that is central to much of our understanding of the behavior of matter and energy. It is central to:

- Structure of atoms
- Interaction between atoms (chemistry, material strength, etc)
- Lights, radio, etc
- Modern power technology
- All electronic technology

It is one of the four fundamental forces (electromagnetic, weak, strong, and gravity)

NOTE: Modern theory unifies the first 3 - grand unification. Professor Weinberg in physics received the Nobel prize in 1979 for unification of electromagnetic and weak forces, which has been confirmed experimentally. Electroweak theory led to the predictions of the Higgs Boson, which was recently discovered at the Large Hadron Collider.

Stoke's Theorem Before we continue with the development of electromagnetic theory we note a calculus identity known as Stoke's Theorem that will be critical to our derivations. Stoke's Theorem relates the line integral of a vector field quantity F around a simply closed path Γ to an area integral of the curl of the field. That is,

Stoke's Theorem

$$\oint_{\Gamma} F \cdot d\ell = \int_A (\nabla \times F) \cdot n dA. \quad (7.0.1)$$

7.1 Electricity (Coulomb and Gauss)

Before considering electromagnetism in all its glory we develop ideas of electricity. We will show how the electro- part of electromagnetism deals with the forces between charges. In particular we will develop the idea of an "electric field". While we will later see that this electric field is coupled to a so-called "magnetic field", we can develop some important relationships without consideration of magnetic fields. Of fundamental importance to this section is the development of Gauss' law, one of four equations in the famous Maxwell equations that describe electromagnetism.

Coulomb's Law. The formalization of electromagnetism began with the work of Charles-Augustin de Coulomb. Electromagnetic forces had been observed prior to the work of Coulomb—for example in "charging" an amber rod by rubbing it with fur—but no formal theory existed to explain the effect. Coulomb posited the first of the mathematical description of electromagnetic forces with the following theorem:

Theorem 7.1.1 (Coulomb's Law). *Two point particles with charges q_1 and q_2 separated by a distance r experience a repulsive force*

$$F = K \frac{q_1 q_2}{r^2} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2} \quad (7.1.1)$$

A new physical quantity was introduced here—*charge*—measured in Coulombs which are now defined as the charge transmitted in one second in a current of one Ampere ($1C = 1\text{Amp} \cdot s$)

<p>Question 1: <i>Where does Coulomb's Law come from?</i></p>

To better understand this new physical quantity or charge we consider a few experimental observations:

- Charge is "quantized" The charge of an electron

$$e = 1.60 \times 10^{-19}$$

so $q = \pm ne$

- Charge is conserved - a body occupying region Ω can only change due to flux of charge through $\partial\Omega$

$$\frac{\partial q_\Omega}{\partial t} = \int_{\partial\Omega} j \cdot ndA$$

where j is the *current density*.

If we treat the electromagnetic media with a continuum approach, we can introduce the charge density ρ to obtain a continuum representation of conservation of charge:

$$\frac{\partial}{\partial t} \int_\Omega \rho dx = \int_{\partial\Omega} j \cdot ndA = \int_\Omega \nabla \cdot j dx$$

Conservation of Charge

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot j \quad (7.1.2)$$

Electric Field. We can now introduce the concept of "electric field", a vector field representing the force/unit charge on a stationary point charge. If a charge q_1 is at the origin and q_2 is at a point x , the force on q_2 is

Coulomb's Law (stationary charges)

$$F = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{|x|^3} x \quad (7.1.3)$$

where the additional factor $\frac{x}{|x|}$ has been multiplied to (7.1.1) to represent the direction of the force. The electric field due to a point charge at the origin is then given by:

$$E = \frac{1}{4\pi\epsilon_0} \frac{q_1}{|x|^3} x$$

and has units of force/charge. Note that this implies that $F = Eq$.

Gauss' Law. The electric field is linear in the charge and so the contribution from different charges can be added up. In particular the electric field due to a charge density can be completed:

$$E(x) = \frac{1}{4\pi\epsilon_0} \int_{\Omega} \rho \frac{x-s}{|x-s|^3} ds.$$

With some effort, this expression of Coulomb's law can be used to show

$$q_{\Omega} = \int_{\Omega} \rho dx = \int \epsilon_0 E \cdot n dA$$

which implies that,

Gauss's Law

$$\epsilon_0 \nabla \cdot E = \rho \quad (7.1.4)$$

This is derived from electrostatics (without moving charges) but it turns out, as we will see, it is true more generally (i.e. for moving charges). This is Gauss' Law. Coulomb's Law is a special case of Gauss' Law.

Electric potential (voltage). In electrostatics the electric field is conservative, this means that the circuit integral:

$$\oint_C E \cdot ds = 0$$

$$\implies \nabla \times E = 0 \quad (7.1.5)$$

where C is any closed curve. Since the force on a particle with charge q is qE , this is the work required to move a unit charge around the loop. Conservative vector fields can be described by a potential, in this case an energy potential. That is, there is a scalar field, V , such that

Electric Potential (Voltage)

$$E = -\nabla V \quad (7.1.6)$$

where V is the electric potential or voltage.

The work required to move a particle with charge q from a point x to a point y is simply

$$\begin{aligned} -q \int_x^y E \cdot dl &= q \int_x^y \nabla V \cdot dl \\ &= q(V(y) - V(x)) \end{aligned}$$

V is thus the potential energy/unit charge - with SI units: joules/coulomb = volts relative to some (arbitrary) reference. Note in particular that this work required to move a particle is independent of the path, the work is simply the charge multiplied by the voltage difference at the starting and endpoints.

The potential due to a point charge of charge q is then given by

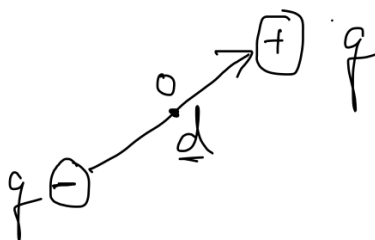
$$V = \frac{1}{4\pi\epsilon_o} \frac{q}{r} \quad (7.1.7)$$

where r is the distance to the point.

7.2 Magnetism (Ampere)

Our formalism of electricity in the previous section only considered the forces stationary charges exert on one another. While it was claimed that Gauss' Law holds for more general fields we have not yet considered how the motion of charges might affect the electric field. In order to consider the effect of moving charges we need to introduce the concept of magnetism. Similar to how forces between charges was the fundamental to our development of ideas of electricity, the motion of charges will be fundamental to our development of magnetism. In particular, in this section we will derive two important relationships, Ampere's law and another constraint known as the "absence of magnetic monopoles".

Magnetic dipole. Consider a pair of particles with equal but opposite charges, separated by a vector \mathbf{d} and centered at the origin. This system is said to have a dipole moment



$\mathbf{m} = q\mathbf{d}$ (Note: the \mathbf{d} points from the negative to the positive charge). Then an electric field perpendicular to \mathbf{d} will exert forces in opposite directions on the $+$ and $-$ particles, causing a torque. The torque exerted on this dipole couple in an electric field E is then given by

$$\tau = m \times E$$

Due to the linearity of our formulations so far we can superimpose the forces exerted these particles. The potential due to such a dipole is then

$$V = \frac{q}{4\pi\epsilon_0} \left(\frac{1}{|x + \frac{d}{2}|} - \frac{1}{|x - \frac{d}{2}|} \right).$$

When $x \gg d$, this can be simplified to

$$\begin{aligned} V &= \frac{1}{4\pi\epsilon_0} \frac{q(d \cdot x)}{|x|^3} = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{m} \cdot \mathbf{x}}{|x|^3} \\ \Rightarrow E &= \frac{1}{4\pi\epsilon_0} \left(\frac{3\mathbf{m} \cdot \mathbf{x}}{|x|^5} \mathbf{x} - \frac{1}{|x|^3} \mathbf{m} \right) \end{aligned}$$

NOTE: the potential dies off like $\frac{1}{|x|^2}$, instead of $\frac{1}{|x|}$ for a single charge.

Forces on moving charges. When a charge is moving, in addition to the force exerted by an electric field, there will be a force due to magnetic fields (this force is called the Lorentz force). That is, when charges are allowed to move forces take the form,

Force on a Charge (moving charges)

$$F = q(E + v \times B) \quad (7.2.1)$$

where B is the magnetic field and v is the velocity. The magnetic field B therefore can be understood to describe the forces acting on a charge when the charge is in motion. Note in particular that coordinate system invariance requires that B be a pseudovector, depending on the handed-ness of the coordinate system.

Question 2: *Is B a vector field? What about frame invariance?*

No, it is a pseudovector field. We have to specify the handed-ness of B to remove the ambiguity.

B has units of force/charge/velocity - in SI: Telsa = $\frac{N \cdot s}{C \cdot m}$.

Question 3: *What gives rise to magnetic fields*

The motion of charges.

History of Magnetism We digress briefly from our formalization of magnetism to consider the historical development of this theory and the ideas that lead to its development. This historical development can help to better understand magnetism and highlight some of its consequences.

Add in historical notes (and pictures) here...

Ampere's Law. Ampere's description of magnets were simply current loops was formalized by Maxwell as,

$$\oint_c B \cdot dl = \mu_0 i_{\text{enclosed}}$$

where μ_0 is the permeability constant = $4\pi \times 10^{-7} \frac{T \cdot m}{A}$ and i_{enclosed} is the current flowing through the loop C (units of Amperes = Coulombs/sec).

Question 4: *Where did Ampere's Law come from?*

Maxwell. Just kidding, it comes from the representation of dipoles as current in a closed loop.

In our continuum setting, the current can be represented as

$$i = \int_A j \cdot ndA$$

where j is the current density and A is the area of a surface (any surface) enclosed by C . Then,

$$\oint_C B \cdot ds = \int_A (\nabla \times B) \cdot ndA = \mu_0 \int_A j \cdot ndA$$

and, since A is arbitrary,

Ampere's Law (for steady E)

$$\nabla \times B = \mu_0 j \quad (7.2.2)$$

This is only true if the electric field E is steady (not varying in time), but we will get to that more later. After modifying this relationship to account for time variations in E this relationship will form the second of four Maxwell equations.

Absence of Magnetic Monopoles. Gauss' Law for electric fields related the flux of E to charges - the analog for magnetic fields is also true, except there are no magnetic charges or "monopoles"

Question 5: *Is this consistent with your experience of magnetism?*

Yes, a magnet always has both north and a south polarizations. If you cut a magnet you'll just end up with two smaller magnets with the same polarization as the original magnet.

Because magnets come in dipole pairs, the magnetic field can never have a source of positive or negative magnetism alone. This leads to the following relation:

$$\int_{\partial\Omega} B \cdot n dA = 0 \quad (7.2.3)$$

which leads to the following fundamental equation by applying the divergence theorem and noting the domain is arbitrary,

Absence of Magnetic Monopoles

$$\nabla \cdot B = 0 \quad (7.2.4)$$

7.3 Electromagnetism (Faraday and Maxwell)

Faraday's law of magnetic induction. Ampere's formulation of magnetic dipoles as current loops, reveals that we can generate a magnetic field by applying a current. Faraday and others wondered if there was a symmetric mechanism for generating current through a magnetic field. This is indeed the case as we will now see and the equation describing this mechanism is called the Law of Magnetic Induction or Faraday's Law.

Recall that in electrostatics E is conservative so one cannot extract energy by running a charge around a loop. But this is exactly what is needed to have electric power. Faraday's Law to the rescue:

$$\oint_c E \cdot ds = -\frac{d}{dt} \int_A B \cdot nds$$

Or using Stokes Theorem:

$$\int_A (\nabla \times E) \cdot dA = -\frac{d}{dt} \int_A B \cdot nds$$

Faraday's Law

$$\nabla \times E = -\frac{\partial B}{\partial t} \quad (7.3.1)$$

Question 6: <i>Where does Faraday's Law come from?</i>

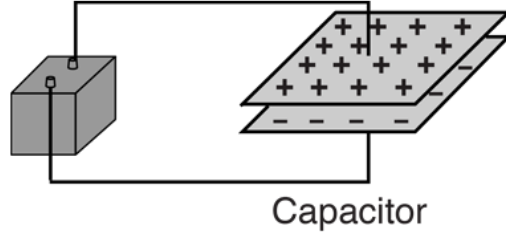
Ampere-Maxwell Law, correcting for time varying electric fields. By the mid 19th century Coulomb's, Gauss', Ampere's, and Faraday's laws were all known in one form or another from numerous experimental observations but did not use modern vector field theory and their formulations were not compact or clear. James Clerk Maxwell set out to change this and formalize electromagnetics. All of previously noted laws of electromagnetism were formalized and condensed by Maxwell. However, in addition to improving the book keeping of these laws Maxwell noticed that Ampere's Law was incomplete because magnetic fields were known to occur in the gap of capacitors when charging. Maxwell corrected this inconsistency as we will now demonstrate, and the resulting (fixed) law is known as the Ampere-Maxwell Law.

To understand the inconsistency of Ampere's law consider a charging capacitor, where current is flowing in one wire, charges are accumulating on the plates, and current is flowing out the other end. Consider Ampere's Law in a loop in the plane between the capacitor plates.

$$\oint_C B \cdot ds = \mu_0 \int_A j \cdot ndA$$

No current is flowing through this loop so the right-hand-side is 0.

Now consider a loop going around the capacitor plates. Then current is flowing so the right-hand-side is non-zero. This is clearly an inconsistency as the magnetic field should be the same for both loops. We need to account for this by introducing another term to account for the changing electric field.



Clearly as charges accumulate on the capacitor plates the electric field between the plates is changing. Observe,

$$q_{\Omega} = \int_{\Omega} \epsilon_0 E \cdot n \, dA = \epsilon_0 E \delta A$$

but since the electric field acts vertically (can be seen due to symmetry and by considering the capacitor plates as infinitely long relative to a charge) we see,

$$q_{\Omega} = -Q(t) \frac{\delta A}{A}$$

where Q is the total charge in the capacitor. Thus,

$$\begin{aligned} \epsilon_0 E \delta A &= -Q(t) \frac{\delta A}{A} \\ \Rightarrow E &= -\frac{Q(t)}{\epsilon_0 A} \end{aligned}$$

then,

$$\frac{dE}{dt} = -\frac{1}{\epsilon_0 A} \underbrace{\frac{dQ}{dt}}_{-I}$$

which implies that,

$$\frac{dE}{dt} = \frac{1}{\epsilon_0 A} I$$

and finally that,

$$\int_A \underbrace{\epsilon_0 \mu_0 \frac{dE}{dt}}_{\text{displacement current}} = \mu_0 I.$$

Thus we see that a changing (in time) electric field induces a current.

Maxwell added this term to Ampere's law along with the standard representation of current:

$$\int_A (\nabla \times B) \cdot n \, dA = \mu_0 \int_A j \cdot n \, dA + \mu_0 \epsilon_0 \frac{d}{dt} \int_A E \cdot n \, dA$$

and since A is arbitrary we get,

Ampere-Maxwell Law

$$\nabla \times B = \mu_0 \left(j + \epsilon_0 \frac{\partial E}{\partial t} \right) \quad (7.3.2)$$

Which states that a magnetic field is induced when the electric field changes. This form of Ampere's law is now consistent and forms the fourth and final equation in Maxwell's equations.

7.4 Maxwell's Equations

The combined result are the celebrated Maxwell's equations:

Maxwell's Equations

$$\begin{aligned}
 \epsilon_0 \nabla \cdot E &= \rho && \text{(Gauss's Law)} \\
 \nabla \cdot B &= 0 && \text{(No Magnetic Monopoles)} \\
 \nabla \times E &= -\frac{\partial B}{\partial t} && \text{(Faraday's Law)} \\
 \nabla \times B &= \mu_0 \left(j + \epsilon_0 \frac{\partial E}{\partial t} \right) && \text{(Ampere-Maxwell Law)}
 \end{aligned} \tag{7.4.1}$$

together with the definition of forces on charges as $F = q(E + v \times B)$, these equations describe electromagnetism in a vacuum.

Magnetic dipoles: a magnet produces a magnetic field like that of a dipole.

Consider a current loop with current i in a loop around an area A with normal \mathbf{n} . The magnetic moment is $m = iA\mathbf{n}$. For $|x| \gg \sqrt{A}$, the magnetic field is given by

$$B(x) = \frac{\mu_0}{4\pi} \left(\frac{3m \cdot x}{|x|^5} \mathbf{x} - \frac{m}{|x|^3} \right) \tag{7.4.2}$$

Electrons (+protons) have angular momentum (spin) and charge - which produces a magnetic moment $\mu = \frac{e}{m} \mathbf{s}$ where $s = \pm \frac{1}{2} \hbar$ and m is the mass of the particle - this electrons have much larger magnetic moments. In ferromagnets it is the alignment of the electronic dipoles that creates the magnetic field.



7.5 Electromagnetic Waves

Maxwell noticed a remarkable fact about his equations—they support waves! Assume $\rho = 0$ and $j = 0$, then take the curl of Faraday's Law

$$\nabla \times \nabla \times E = -\frac{\partial \nabla \times B}{\partial t} = -\mu_0 \epsilon_0 \frac{\partial^2 E}{\partial t^2} \quad (7.5.1)$$

But $\nabla \times \nabla \times E = \nabla(\nabla \cdot E) - \Delta E = -\Delta E$ so

$$\frac{\partial^2 E}{\partial t^2} = \frac{1}{\mu_0 \epsilon_0} \Delta E \quad (7.5.2)$$

This is the wave equation! It describes waves that travel at speed

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \approx 3.0 \times 10^8 \frac{m}{s} \quad (7.5.3)$$

which is the speed of light. Thus Maxwell identified light as an electromagnetic wave. Similarly, B satisfies the same equation

$$\frac{\partial^2 B}{\partial t^2} = \frac{1}{\mu_0 \epsilon_0} \Delta B \quad (7.5.4)$$

Solutions of the wave equation are well studied. Individually, these equations have the form

$$E(\mathbf{x}, t) = \mathbf{G}(\mathbf{x} \cdot \mathbf{e} - ct) \quad (7.5.5)$$

for any vector function \mathbf{G} (twice differentiable) and any unit vector \mathbf{e} . \mathbf{e} is the propagation direction of this plane wave. Since the equation is linear, superpositions of such solutions are also solutions. But the E and B solutions are coupled through Maxwell's equations.

By Fourier analysis, \mathbf{G} can be represented by a superposition of sinusoidal functions

$$E = E_0 e^{i(k \cdot x - \omega t)} \quad B = B_0 e^{i(k \cdot x - \omega t)}$$

where

\mathbf{k} = wave vector or wave number vector

ω = angular frequency

$|E_0|$ = amplitude

Clearly, $\omega/|k|$ is the propagation speed c , and $k/|k|$ is the direction of the propagation. Also define,

$$\begin{aligned} \lambda &= \frac{2\pi}{|k|} = \text{wavelength} \\ T &= \frac{2\pi}{\omega} = \text{period} \\ \nu &= \frac{1}{T} = \frac{\omega}{2\pi} = \text{frequency} \end{aligned}$$

It is easy to show that for such solutions, Maxwell's equations require

$$k \cdot E = 0 \implies k \cdot E_0 = 0 \quad (7.5.6)$$

$$k \cdot B = 0 \implies k \cdot B_0 = 0 \quad (7.5.7)$$

$$k \times E = \omega B \implies \hat{k} \times E = cB \quad (7.5.8)$$

$$k \times B = -\frac{1}{c^2}\omega E \implies \hat{k} \times B = -\frac{1}{c}E \quad (7.5.9)$$

where $\hat{k} = \frac{k}{|k|}$

The last two imply:

$$E \cdot B = 0$$

$$|E_0| = c|B_0|$$

$$E \times B = \frac{k}{\omega}|E|^2 = \frac{1}{c}|E|^2\hat{k}$$

Thus, a sinusoidal plane wave has the following structure:

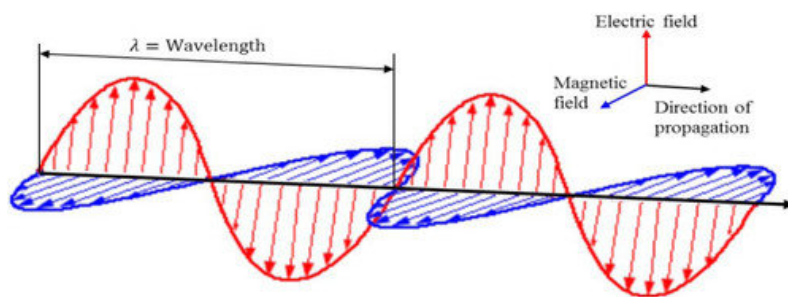


Figure 7.5.1

and an arbitrary solution can be composed of such solutions.

Familiar forms of E-M radiation differ only in frequency/ wavelength

light	$\lambda = 400 - 700nm$	$\nu = 4.3 - 7.5 \times 10^{14}Hz$
microwaves	$\lambda = 1mm - 1m$	$\nu = 0.3 - 300GHz$
radio	$\lambda = 1m - 1km$	$\nu = 300kHz - 300MHz$
xrays	$\lambda = 0.01 - 10nm$	$\nu = 3 \times 10^{16-19}Hz$
microwave oven	$\lambda = 12cm$	$\nu = 2.45GHz$

7.6 Summary

Stoke's Theorem. Relates a contour integral to an integral over the area enclosed by the contour.

$$\oint_{\Gamma} F \cdot d\ell = \int_A (\nabla \times F) \cdot n \, dA.$$

Conservation of Charge. This simply expresses that the rate of change of charge density (ρ) inside a volume is equal to the flux of charge (j is current density or ρv).

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot j$$

Coulomb's Law. Two point particles with charges q_1 and q_2 separated by a distance r experience a repulsive force

$$F = K \frac{q_1 q_2}{r^2} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2}$$

For two stationary charges (with q_1 at origin) then we have then that

$$F = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{|x|^3} x$$

and,

$$E = \frac{1}{4\pi\epsilon_0} \frac{q_1}{|x|^3} x \quad (\text{since } F = Eq)$$

Force on a Charge (moving charges) Moving charges experience both an electrostatic force due to Coulomb's law, as well as a force due to the magnetic field their motion induces called a Lorentz force.

$$F = \underbrace{qE}_{\text{Coulomb force}} + \underbrace{qv \times B}_{\text{Lorentz force}}$$

Electric Field. A vector field representing the force/unit charge on a stationary point charge.

$$\text{i.e.} \quad F = Eq$$

Because the electric field is conservative ($\nabla \times E = 0$), Helmholtz decomposition tells us it can be represented as the gradient of a (scalar) potential,

$$E = -\nabla V \quad (\text{Electric Potential or Voltage})$$

Magnetic Field. Just as an electric field was defined by the (Coulomb) force that it induces on a stationary charge, the magnetic field is defined by (Lorentz) force that it induces on a moving charge. The magnetic field isn't technically a vector field, it's a pseudo-vector field (due to cross product in Lorentz force) so we must agree on a coordinate system (define handedness of the field).

Maxwell Equations.

- **Gauss's Law.** This represents the effect of charge density on the electric field—it relates electric flux in a volume to the amount of charge in that volume. (Recall that when we have a divergence in a differential relationship it almost always arises through a flux term)

$$\epsilon_0 \nabla \cdot E = \rho$$

- **Absence of Magnetic Monopoles (Gauss' Law of Magnetism).** This is Gauss' Law but for magnetic flux. However, since there is no magnetic monopole (positive polarization always comes with negative polarization) there can be no source term inside a volume, thus the flux is zero.

$$\nabla \cdot B = 0$$

- **Faraday's Law (Law of Magnetic Induction).** Ampere's law says we can make a magnetic field using an electric field, Faraday's law is the opposite effect: it allows an electric field to be generated by a changing magnetic field. A common device that takes advantage of this is an electric generator.

$$\nabla \times E = -\frac{\partial B}{\partial t}$$

- **Ampere-Maxwell Law.** Represents the formulation of magnetic fields by currents. A common device that takes advantage of this is an electromagnet.

$$\nabla \times B = \mu_0 \left(j + \epsilon_0 \frac{\partial E}{\partial t} \right)$$

- **Definition of force on a charge.** Moving charges induce a magnetic field, thus our definition of force on a charge needs to consider both Coulomb electrostatic forces and so-called Lorentz forces.

$$F = \underbrace{qE}_{\text{Coulomb force}} + \underbrace{qv \times B}_{\text{Lorentz Force}}$$

Exercise

Compute the total force on a point charge q at $(0, r)$ from a line charge distribution with a constant charge density q_l along the line at $y = 0$. (Hint: To compute the resulting integral, use the symmetry about the x -axis and that $\int_{-\infty}^{\infty} \frac{1}{(1 + \xi^2)^{3/2}} = 2$.)

Solution:

8. Quantum Mechanics

Quantum Mechanics is the theory underlying our understanding of physical phenomena at the atomic scale. It is well-known for its "strangeness". It is spectacularly successful theory, which is mathematically rich.

8.1 History

Question 1: *Where did Quantum Mechanics come from?*

Max Planck - Black body radiation spectrum observed to be inconsistent with Maxwell's equations and statistical mechanics! He fit the spectrum to a functional form that was consistent with energy in **electromagnetic (EM)** waves being "quantized".

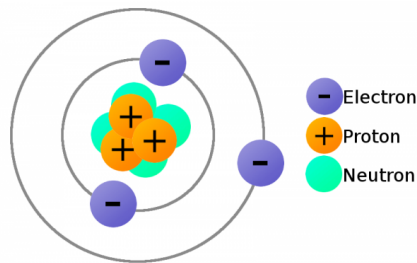
$$E = nh\nu \quad (8.1.1)$$

ν is frequency of electromagnetic wave, $h = 6.6 \cdot 10^{-34} Js$ - **Planck's constant** - a fitting parameter (Planck didn't believe it)

1905: Einstein's description of the photoelectric effect relies on the concept of a **"photon"** - a discrete particle of EM energy, with energy $E = h\nu = \hbar\omega$ (ω is angular frequency $\rightarrow \hbar = \frac{h}{2\pi}$)

Photo-electric effect: when illuminated by EM radiation, some materials emit free electrons. **The energy of the liberated electrons is $h\nu - l_0$.** ν is the frequency of illuminating radiation, l_0 is some minimum energy required to liberate an electron. *How does this suggest a photon?*

Rutherford (1911): shows (experimentally) that an atom has a heavy + charged nucleus "orbited" by (light) electrons. *What is wrong with this picture?*



Kirchoff (1860): It was well known that when heated a body can produce electromagnetic radiation. Kirchoff studied the concept of a "black-body" or a material which perfectly absorbs and emits radiation (no reflection, etc.). Black bodies approximated using cavities with small holes. In particular Kirchoff noted that under equilibrium conditions (uniform temperature) the density of radiation must be a function of frequency and temperature only $\rho(\nu, T)$ otherwise would violate 2nd Law of Thermodynamics.

Wien (1886): Did a thermodynamic analysis of black-body cavity resulting in the following relationships:

$$\rho(\nu, T) = \nu^3 g\left(\frac{\nu}{T}\right) \qquad \left. \frac{\partial \rho}{\partial \nu} \right|_T = f\left(\frac{\nu}{T}\right)$$

which implies,

Wien Displacement Law

$$\nu_{max} \propto T \qquad I \sim T^4$$

Wien (1896): Based on experimental measurements proposed the model,

$$\rho(\nu, T) = \alpha \nu^3 e^{-\beta \nu / T}$$

Plank (1899): Did his PhD in entropy, this was his tool of choice. Proposed a field of charged harmonic oscillators. This led to the relationship,

$$\begin{aligned} \rho(\nu, T) &= \frac{8\pi\nu^2}{c^3} U(\nu, T) \\ \implies U &= \frac{c^3}{8\pi\nu^2} \rho(\nu, T) = \frac{\alpha c^3 \nu}{8\pi} e^{-\beta \nu / T} \end{aligned}$$

Using the relationship $TdS = -dU$

$$\begin{aligned} \implies S &= \frac{U}{8\pi\nu^2} \left(\ln \frac{8\pi U}{\alpha \nu c^2} - 1 \right) \\ \implies \frac{\partial^2 S}{\partial U^2} &= -\frac{1}{\beta U} \end{aligned}$$

Plank thought this was a fundamental relationship at first but would soon learn it would need to be modified.

Ruebens and Kurlbaum (1900): Measured low frequency emission and proposed instead

$$\rho(\nu, T) \sim \nu^2 T.$$

This implies instead that,

$$U \sim T \quad \text{independent of } \nu$$

This meant that Plank's model was inconsistent (indeed it was known to be inconsistent in some regime due to the — phenomenon) and instead,

$$\frac{\partial^2 S}{\partial U^2} = -\frac{\gamma}{U^2}.$$

To make his previous model consistent with these low frequency observations, Plank proposed a model that 'blended' these relationships,

$$\frac{\partial^2 S}{\partial U^2} = -\frac{\nu}{U\left(\frac{\beta}{\gamma}\nu + U\right)}$$

$$= -\frac{k}{U(h\nu + U)}$$

where k ($1.38 \times 10^{-23} \frac{\text{m}^2 \text{kg}}{\text{s}^2 \text{K}}$) and h ($6.6 \times 10^{-34} \text{Js}$) are the Boltzman and Plank constants respectively. Integrating gives,

$$S = k \left[\left(1 + \frac{U}{h\nu}\right) \ln \left(1 + \frac{U}{h\nu}\right) - \frac{U}{h\nu} \ln \left(\frac{U}{h\nu}\right) \right] \quad U = \frac{h\nu}{(e^{\frac{h\nu}{kT}} - 1)}$$

which gives,

$$\rho(\nu, T) = \frac{8\pi h\nu^3}{c^3 (e^{\frac{h\nu}{kT}} - 1)}$$

Before all this, in 1877, Boltzman derived the following equation for entropy (Boltzman Law),

$$S = k \ln W$$

where W is a measure on how many microstates can be consistent with the macroscopic phenomenon. Plank tried to apply this relationship, assuming discrete energy levels in an atom to limit the possible distribution, i.e. energy/atom = ϵ . For N atoms with energy $u = j\epsilon$ the total energy is NU . Then we have,

$$W = {}_N C_M = \frac{(N + M - 1)!}{M!(N - 1)!}$$

Applying Sterling's formula ($\ln N! \approx N \ln N - N$) we get,

$$\ln W = N \left[\left(1 + \frac{M}{N}\right) \ln \left(1 + \frac{M}{N}\right) - \frac{M}{N} \ln \left(\frac{M}{N}\right) \right].$$

Notice that this is the same as Plank's previous relationship for S provided $\frac{M}{N} = \frac{U}{h\nu}$. This seemed too good to be true but started to spark the idea that energy levels might be quantized.

Reighley (1900): Reighley wasn't buying these models and proposed a wrong one instead, but did hypothesize that Plank's harmonic oscillators were electromagnetic waves with some sort of resonance.

Einstein (1905): While studying the photoelectric effect, proposed energy in EM radiation quantized. This discrete behavior was more like a particle than a wave and led to the idea of a photon.

Bohr (1913): Bohr proposes that the electrons orbiting the nucleus have quantized angular momentum: $L = n\hbar$ - which defines a discrete set of orbits. With energy E_n . If an electron transitions from one orbit (n) to a lower orbit (m), the energy difference is emitted as a photon with

$$h\nu = E_n - E_m \quad n > m$$

This is ad hoc, but the resulting predictions of the spectrum are right on.

de Broglie (1922): generalized wave-particle duality. EM waves also act like particles (photons) and particles (e.g. electrons) also act like waves. Photons (γ) associated with waves of frequency ν : $E_\gamma = h\nu = \hbar\omega = \hbar ck$ where k is the wave number. Special relativity + Maxwell's equations \implies EM radiation of energy E have momentum $p = \frac{E}{c}$, thus it makes sense to assume:

$$p_\nu = \frac{E_\nu}{c} = \hbar k$$

de Broglie asked, what if the analogous were true - a particle (e.g. electron) is associated with a wave with k and ω given by:

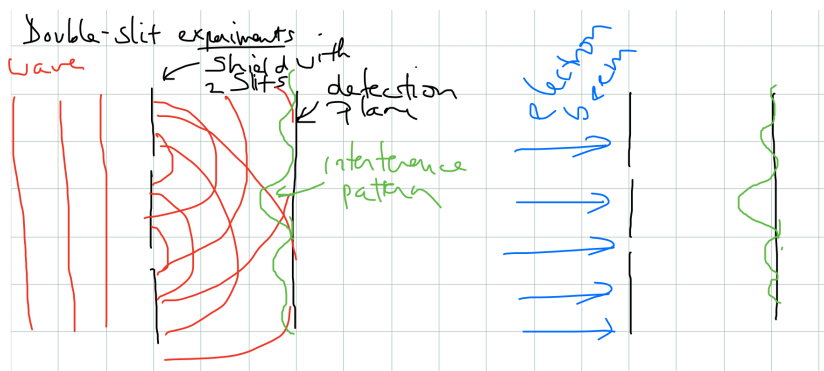
$$\omega = \frac{E}{\hbar}, \quad k = \frac{p}{\hbar} \quad (8.1.2)$$

Where E, p are energy and momentum of the particle (e.g. $\frac{mv^2}{2}, mv$). This gives a nice interpretation for Bohr's atomic model - *what might that be?*

The circumference of an electron's orbit is an integer multiple of the wavelength.

Compton (1923): discovers the "Compton Effect" (change in wavelength of γ -rays upon interaction with electrons) - confirming existence of photons with energy and momentum $E_\gamma = \hbar\omega, p = \hbar k$.

Davisson and Germer (1925): Diffraction—a wave phenomenon—when electrons were scattered from a single-crystal surface of Nickel was observed. This was consistent with the de Broglie hypothesis (wave-particle duality)! Inferred wavelength of electrons agreed with de Broglie to 1%.



For both EM waves and electrons:

1. Interference patterns occur - consistent with wave propagation
2. Discrete particles can be detected independently passing through one slit or the other, and discretely landing on the detector. The interference patterns arise due to the rate at which particles land at each point on the detector.

8.2 Schrödinger's Equation

Schrödinger (1927): While giving a seminar, Schrödinger was asked why, if these particles behaved like waves, is there no wave equation? Schrödinger went home over Christmas and made a wave equation up to describe de Broglie's waves. His reasoning is outlined below:

1. Seek a "wave function" of the form (in 1-D)

$$\Psi(x, t) = A \exp(i(kx - \omega t))$$

2. The kinetic energy of a particle:

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m} \quad (p = mv)$$

3. From de Broglie: $E = \hbar\omega$, $p = \hbar k$ so

$$\hbar\omega = \frac{(\hbar k)^2}{2m}$$

4. We want to find a wave equation that implied this relation. Note:

$$\frac{\partial \Psi}{\partial t} = -i\omega \Psi$$

$$\frac{\partial^2 \Psi}{\partial x^2} = -k^2 \Psi$$

So if

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} \implies \hbar\omega = \frac{\hbar^2 k^2}{2m}$$

This is Schrödinger's equation for a particle moving in free space.

Notice that the LHS is a representation of the total energy of the particle. The RHS represents the kinetic energy. This equation says that the particle only has kinetic energy. Suppose that there is also a potential energy described by a potential $V(x)$ then

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x)\Psi$$

In 3-D, $\Psi = A \exp(i(x \cdot k - \omega t))$ and the same argument yields

Schrödinger's Equation

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V(x)\Psi \quad (8.2.1)$$

Schrödinger developed this as a plausible equation for the wave function, but is it right? How could we tell? He used the equation to compute the energy levels for the hydrogen atom, and got results consistent with the spectrum, and the Bohr model. We'll see about this later.

$$i\hbar \frac{\partial \Psi}{\partial t} \approx E\Psi$$

NOTE: This is an eigenvalue problem where energies are eigenvalues of the RHS.

Schrödinger's equation has been found to be a remarkably accurate description of atomic-scale physics. *But what does Ψ mean?*

Wave Functions and Observables. The wave function Ψ is generally not given physical meaning itself. Instead, its magnitude is given a meaning in terms of probability. This is the Born rule: the probability that the particle is in a small volume dV centered at x at time t is

$$|\Psi(x, t)|^2 dV = \Psi^*(x, t)\Psi(x, t)dV$$

Thus $|\Psi|^2$ is taken to be a probability density. The normalization condition for probability densities implies:

Born Rule

$$\int |\Psi|^2 dx = 1 \quad (8.2.2)$$

Which sets the arbitrary amplitude of Ψ . However, it turns out that the wave function Ψ encodes a lot more than just the probability distribution of the position of a particle. To see this consider de Broglie's relations: $E = \hbar\omega$, $p = \hbar k$. This means we can express the wave function as:

$$\Psi(x, t) = \Psi_0 e^{\frac{i}{\hbar}(p \cdot x - Et)}$$

Then

$$\nabla \Psi = i \frac{p}{\hbar} \Psi \quad \text{and} \quad \frac{\partial \Psi}{\partial t} = -i \frac{E}{\hbar} \Psi$$

or equivalently,

$$p\Psi = \left(\frac{\hbar}{i} \nabla \right) \Psi \quad \text{and} \quad E\Psi = \left(-\frac{\hbar}{i} \frac{\partial}{\partial t} \right) \Psi.$$

The momentum p can thus be thought of as the operator $\frac{\hbar}{i} \nabla$ and the energy E as the operator $-\frac{\hbar}{i} \frac{\partial}{\partial t}$. We will return to the concept of observables as operators on the wave function but for now list these two important operators.

Two Observable Operators

$$\begin{aligned} p\Psi &= \left(\frac{\hbar}{i} \nabla \right) \Psi && \text{(Momentum Operator)} \\ E\Psi &= \left(-\frac{\hbar}{i} \frac{\partial}{\partial t} \right) \Psi && \text{(Energy Operator)} \end{aligned} \quad (8.2.3)$$

8.3 Uncertainty Quantification and Function Spaces

In order to quantify uncertainty we need to know both what the expected value (mean) should be, and how much measurements might deviate from that expected value (standard deviation). We recall the definition of these quantities for probabilities of continuous variables.

Let γ be a random variable with probability density $\rho(\gamma)$ such that

$$\int_{\gamma} \rho(\gamma) d\gamma = 1. \quad (8.3.1)$$

Then we have the following,

Expected Value (mean):

$$\langle \gamma \rangle = \mathbb{E}_{\rho(\gamma)} = \int_{\gamma} \gamma \rho(\gamma) d\gamma \quad (8.3.2)$$

Variance

$$\sigma_{\gamma}^2 = \int_{\gamma} (\gamma - \langle \gamma \rangle)^2 \rho(\gamma) d\gamma = \langle \gamma^2 \rangle - \langle \gamma \rangle^2 \quad (8.3.3)$$

The square root of the variance is the standard deviation σ_{γ} and is a measure of how likely observations of variable γ are expected to deviate from $\langle \gamma \rangle$. This is our measure of uncertainty.

Recall that Schrödinger's wave function Ψ was given the interpretation of a probability density function through the Born rule,

$$\int |\Psi|^2 dx = 1.$$

Recalling the definition of modulus operator for complex variables, i.e.

$$|\Psi|^2 = \Psi^* \Psi$$

we can express probability densities in terms of the wave function and its complex conjugate. Taking x to be the position variable, the expected position can then be computed as,

$$\langle x \rangle = \mathbb{E}_{\rho(\gamma)} = \int_{\mathbb{R}^3} x \Psi^* \Psi dx.$$

8.4 Function Spaces

We briefly review the notion of function spaces and in particular Hilbert spaces (complete inner product spaces) that will allow us to write expressions such as the expected value of x in the form of an inner product,

$$\langle x \rangle = (\Psi, x\Psi)_{L^2}.$$

Recall that the wave function is normalized s.t. $|\Psi|^2$ is the probability density and is thus integrable over \mathbb{R}^d (where d is the dimension of the space). The wave function Ψ must then belong to the class of square integrable functions— $L^2(\mathbb{R}^d)$. Informally, this space $L^2(\mathbb{R}^d)$ is the space of functions u where

$$\int_{\mathbb{R}^d} u^* u dx < \infty$$

and the integral is the so-called Lebesgue integral. See pg. 116 in the textbook for a formal definition.

$L^2(\mathbb{R}^d)$ is a complete inner-product space, or Hilbert space where the inner product is defined as

$$(\Psi, \phi)_{L^2} = \int_{\mathbb{R}^d} \Psi^* \phi dx.$$

This space is complete, which means that all Cauchy sequences in L^2 converge to an element in L^2 . The norm associated with the inner product on L^2 is then

$$\|\Psi\| = \sqrt{(\Psi, \Psi)_{L^2}} = \left(\int_{\mathbb{R}^d} \Psi^* \Psi dx \right)^{1/2}$$

In general, operators that we care about often involve partial derivatives and we can't expect that the derivatives are as well behaved as the original function so we also must consider **Sobolev spaces**, spaces that require that the function—as well as its first m derivatives (understood in the sense of distributions)—are square integrable (L^2 integrable). For example, $H_1(\mathbb{R}^d)$ is the set of functions v s.t.

$$v \in L^2(\mathbb{R}^d), \quad \frac{\partial v}{\partial x_i} \in L^2(\mathbb{R}^d) \quad i = 1, 2, \dots, d$$

Sobolev spaces are also a Hilbert Spaces with the inner product defined as:

$$(\Psi, \phi)_{H^1} = \int_{\mathbb{R}^d} \nabla \Psi^* \cdot \nabla \phi + \alpha \Psi^* \phi dx$$

where $\alpha > 0$ is an arbitrary scale factor. H^m is the analogous space with m derivatives. We will consider L^2 inner products involving operators. Making use of the proper functional spaces will allow us to manipulate operators, for example through integration by parts when derivatives are passed from one argument of the inner product to the other. If the function spaces are not properly defined these types of manipulations are illegal since if the derivative of an L^2 function is not L^2 itself then the L^2 inner product becomes meaningless on the differentiated function.

Where no distinction is made we will write $(\cdot, \cdot)_{L^2}$ as simply (\cdot, \cdot) .

8.5 Wave Function Operators and Observables

In 8.2.3 we made the rather bold claim that we can express the momentum (or rather its expected value) and energy in terms of operators on the wave function Ψ . In this section we hope to motivate this claim. As we will see later, momentum and energy are not the only operators we can define on our wave function.

Position and Momentum Operators. The most simple example of a wave function operator is that for computing the expected value of position $\langle x \rangle$. From the Born rule we have,

$$\langle x \rangle = (\Psi, x\Psi).$$

We then say that the position operator $\tilde{x} = x$ which is simply the identity operator. To see an example of a less trivial operator consider $m \frac{d\langle x \rangle}{dt}$. In Newtonian mechanics we have that $m \frac{dx}{dt} = p$, the momentum, should a similar result hold for the expected value of p ?

Observe,

$$\begin{aligned} \frac{d\langle x \rangle}{dt} &= \frac{d}{dt} \int_{\mathbb{R}^3} x \Psi^* \Psi dx \\ &= \int_{\mathbb{R}^3} x \underbrace{\frac{\partial}{\partial t} (\Psi^* \Psi)} dx \end{aligned}$$

For a moment let's focus on the underbraced term only,

$$\frac{\partial}{\partial t} (\Psi^* \Psi) = \Psi^* \frac{\partial \Psi}{\partial t} + \Psi \frac{\partial \Psi^*}{\partial t}$$

but, using Schrödinger's equation we can substitute for $\frac{\partial \Psi}{\partial t}$ and $\frac{\partial \Psi^*}{\partial t}$ to get,

$$\begin{aligned} &= \frac{i\hbar}{2m} \left(\Psi^* \nabla^2 \Psi - \Psi^* \frac{2m}{\hbar^2} V(x) \Psi - \nabla^2 \Psi^* \Psi + \frac{2m}{\hbar^2} V(x) \Psi^* \Psi \right) \\ &= \frac{i\hbar}{2m} \left(\nabla \cdot (\Psi^* \nabla \Psi - \nabla \Psi^* \Psi) \right) \end{aligned}$$

Plugging this back into our expression for $\frac{d\langle x \rangle}{dt}$ and integrating by parts (note that x disappears as $\nabla x = \delta(x)$ and the flux term vanishes by assumptions on field) we get,

$$\begin{aligned} \frac{d\langle x \rangle}{dt} &= -\frac{i\hbar}{2m} \int_{\mathbb{R}^3} (\Psi^* \nabla \Psi - \nabla \Psi^* \Psi) dx \\ &= -\frac{i\hbar}{2m} \int_{\mathbb{R}^3} (\Psi^* \nabla \Psi + \Psi^* \nabla \Psi) dx \quad \text{applying IBP to second term} \\ &= -\frac{i\hbar}{m} \int_{\mathbb{R}^3} \Psi^* \nabla \Psi dx. \end{aligned}$$

Multiplying by mass m we then obtain,

$$\begin{aligned} m \frac{d\langle x \rangle}{dt} &= -i\hbar \int_{\mathbb{R}^3} \Psi^* \nabla \Psi dx \\ &= \int_{\mathbb{R}^3} \Psi^* \frac{\hbar}{i} \nabla \Psi dx \\ &= (\Psi, \frac{\hbar}{i} \nabla \Psi) = \langle p \rangle. \end{aligned}$$

It thus follows that the momentum operator defined previously in 8.2.3 returns the expected value of the momentum when operating on Ψ . To be clear, p is an observable quantity—for a given particle we can measure its momentum. The momentum operator defined (abusing notation)

$$\tilde{p} := \frac{\hbar}{i} \nabla$$

is an operator that, when applied in the manner $(\Psi, \tilde{p}\Psi)$ returns the expected value of p ,

$$\langle p \rangle = (\Psi, \tilde{p}\Psi).$$

As we continue to define wave function operators for various observable quantities this pattern will hold thus it is important to understand the connection between the quantity, its wave function operator, and the expected value.

Similar to our previous analysis for $\frac{d\langle x \rangle}{dt}$, we can show that the time derivative of the expected value of momentum $\frac{d\langle p \rangle}{dt}$ corresponds to a force (congruent with ideas from Newtonian mechanics), but to do this it will first be convenient to introduce the idea of the Hamiltonian.

Hamiltonian Operator. The Hamiltonian of a system is just the function of position and momentum that yields energy. For a particle then the Hamiltonian is simply the sum of kinetic and potential energies,

$$H(x, p) = \underbrace{\frac{p \cdot p}{2m}}_{\text{Kinetic}} + \underbrace{V(x)}_{\text{Potential}}.$$

In quantum mechanics however we formulate the Hamiltonian as an operator. Since we already have operators for position and momentum constructing such an operator is simple to obtain,

$$\tilde{H} = \frac{\tilde{p} \circ \tilde{p}}{2m} + V(x)$$

or,

$$\tilde{H}(x, \frac{\hbar}{i} \frac{\partial}{\partial x}) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x).$$

Recall however from 8.2.3 that E also has a corresponding wave function operator,

$$\tilde{E} = i\hbar \frac{\partial}{\partial t},$$

thus Schrödinger's equation simply becomes,

$$\begin{aligned} H(x, p) &= E \\ \implies \tilde{H}(x, \frac{\hbar}{i} \frac{\partial}{\partial x}) \Psi &= i\hbar \frac{\partial}{\partial t} \Psi \end{aligned} \tag{8.5.1}$$

More generally, systems may have generalized "momentum" variables, p , and generalized position variables q (e.g. angular momentum and angular position) also, there can be many (N) particles, with N different positions and momenta. Thus, $H(q_1, q_2, \dots, q_N, p_1, p_2, \dots, p_N)$ and Schrödinger becomes:

$$i\hbar \frac{\partial \Psi}{\partial t} = H(q_1, q_2, \dots, q_N, \frac{\hbar}{i} \frac{\partial}{\partial q_1}, \dots, \frac{\hbar}{i} \frac{\partial}{\partial q_N}) \Psi$$

where $\Psi = \Psi(q_1, \dots, q_N, t)$.

The Correspondence Principle. We are now equipped to show that $\frac{d\langle p \rangle}{dt} = \langle -\nabla V \rangle = \langle F \rangle$, which is congruent with the concept of conservation of momentum in classical mechanics. This is an example of the correspondence principle—since classical mechanics is accurate at very large scales (compared to the scales we are now dealing with), it must be recovered in the appropriate sense.

Observe,

$$\begin{aligned} \frac{d}{dt} \langle p \rangle &= \frac{d}{dt} (\Psi, \tilde{p} \Psi) \\ &= \left(\frac{\partial}{\partial t} \Psi, \tilde{p} \Psi \right) + (\Psi, \tilde{p} \frac{\partial}{\partial t} \Psi) \\ &= -\frac{1}{i\hbar} \left[(\tilde{H} \Psi, \tilde{p} \Psi) - (\tilde{p} \Psi, \tilde{H} \Psi) \right] \quad \text{substituting (8.5.1)} \\ &= -\frac{1}{i\hbar} (\Psi, (\tilde{H} \tilde{p} - \tilde{p} \tilde{H}) \Psi) \end{aligned}$$

but, since ∇^2 in \tilde{H} and \tilde{p} commute we have,

$$\begin{aligned} &= -\frac{1}{i\hbar} (\Psi, V \tilde{p} \Psi - \tilde{p} (V \Psi)) \\ &= -\frac{1}{i\hbar} (\Psi, \cancel{V \tilde{p} \Psi} - (\tilde{p} V) \Psi - \cancel{V \tilde{p} \Psi}) \\ &= \frac{1}{i\hbar} (\Psi, \left(\frac{\hbar}{i} \nabla V \right) \Psi) \\ &= (\Psi, (-\nabla V) \Psi), \end{aligned}$$

thus $\tilde{F} = -\nabla V$ is the force operator.

Time Invariant Schrödinger's Equation and Eigenfunction Expansions.

Suppose the wave function Ψ can be written as the product of a spatial function and a temporal function, i.e.,

$$\Psi(x, t) = \psi(x) \phi(t)$$

then we can perform separation of variables and obtain the following,

$$\phi(t) = e^{-iEt/\hbar},$$

and

Time Invariant Schrödinger's Equation

$$\tilde{H} \psi = E \psi \tag{8.5.2}$$

where ψ is a function of x only (note that in the future we will do our best to stick to the convention $\Psi(x, t)$ and $\psi(x)$), and \tilde{H} is the Hamiltonian operator and E is the energy.

Note that the operator \tilde{E} involves a time derivative so in this case would return 0 since ψ is independent of time, thus the distinction should be clear that E here is a scalar representing the energy.

In fact, this equation can be viewed as an eigenvalue problem with eigenfunctions corresponding to permissible quantum states and eigenvalues E_i corresponding to the energy levels of those quantum states. We will return to this concept of eigenvalue problems in the next section but note that since (8.5.2) is linear, eigenfunctions can be superimposed and since the operator \tilde{H} is Hermitian (more about that later), the eigenvalues are real, and eigenfunctions are orthogonal and form a complete basis for the space of permissible quantum states.

Heisenberg Uncertainty Principle—a Precursor. We now use our understanding of wave function operators and uncertainty quantification to derive the famous Heisenberg Uncertainty Principle—that the momentum and position of a particle may not be known simultaneously. At this point however we do not give a fully detailed derivation, as this derivation will be repeated in more detail in the following section for general dynamical variables (not just position and momentum).

For simplicity suppose that the expected value of $\langle x \rangle = 0$ and $\langle p \rangle = 0$, indeed if this were not the case a change of variables may be performed and the following argument would hold (in the derivation in the next section we will not make this assumption). Then,

$$\begin{aligned}\sigma_x^2 = \langle x^2 \rangle &= \int_x |x|^2 \psi^* \psi dx = \overbrace{(x\psi)^*}^f, x\psi = (f, f) \\ \sigma_p^2 = \langle p^2 \rangle &= \int_x \psi^* \left(\frac{\hbar}{i} \nabla \right)^2 \psi dx = -\overbrace{\left(\frac{\hbar}{i} \nabla \psi \right)^*}^g, \frac{\hbar}{i} \nabla \psi = (g, g)\end{aligned}$$

Where the last step follows from integration by parts. Letting $f = x\psi$ and $g = \frac{\hbar}{i} \nabla \psi$. By Cauchy Schwartz,

$$(f, f)(g, g) \geq |(f, g)|^2 \geq (\text{Im}(f, g))^2 = \left(\frac{(f, g) - (g, f)}{2i} \right)^2$$

but,

$$\begin{aligned}(f, g) - (g, f) &= \int \left[(\Psi x)^* \cdot \left(\frac{\hbar}{i} \nabla \Psi \right) - \left(\frac{\hbar}{i} \nabla \Psi \right)^* \cdot x \Psi \right] dx \\ &= \frac{\hbar}{i} \int \Psi^* \left[x \cdot \nabla \Psi - \nabla \cdot (x \Psi) \right] dx \\ &= \frac{\hbar}{i} \int \Psi^* \left[\cancel{x \cdot \nabla \Psi} - \Psi - \cancel{x \cdot \nabla \Psi} \right] dx \\ &= i\hbar \int \Psi^* \Psi dx = i\hbar.\end{aligned}$$

It thus follows that,

$$\sigma_x^2 \sigma_p^2 \geq \frac{\hbar^2}{4}$$

which gives the famous result,

Heisenberg Uncertainty Principle

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}. \quad (8.5.3)$$

This implies that the product of the variation is bounded below, so we can never have a 0 variation (perfect knowledge) in both variables simultaneously. Indeed, if $\sigma_x = 0$, then $\sigma_p = \infty$ and we know absolutely nothing about momentum.

8.6 Dynamical Variables, Hermitian Operators, and Observables.

Dynamical Variables and Eigenproblem. We can generalize the concept of wave function operators using a form similar to how the Hamiltonian was defined. A dynamical variable is some physical feature of the quantum system that depends upon the physical state of the system. Again in the context of quantum mechanics we will define a wave function operator that returns probabilistic information about the dynamical variable. In general a dynamical variable function will take the form,

$$Q = Q(q_1, \dots, q_N, p_1, \dots, p_N)$$

And the associated operator (recall $\tilde{p} = -i\hbar \frac{\partial}{\partial q}$) will take the form,

$$\tilde{Q}(q_1, \dots, q_N, -i\hbar \frac{\partial}{\partial q_1}, \dots, -i\hbar \frac{\partial}{\partial q_N}).$$

The expected value of the dynamical variable for a quantum state Ψ is then again give by appropriate application of the operator to the wave function,

$$\langle Q \rangle = (\Psi, \tilde{Q}\Psi).$$

Hermitian Operators and Observables. Note however that if we hope to measure or observe Q its value should be real. This puts constraints on our operator \tilde{Q} to be a so-called Hermitian operator. Notice that,

$$\langle Q \rangle = \langle Q \rangle^* \quad (\text{i.e. } \langle Q \rangle \text{ is real})$$

implies that,

$$(\Psi, \tilde{Q}\Psi) = (\tilde{Q}\Psi, \Psi)$$

An operator is Hermetian if and only if this condition holds for all possible functions in our Hilbert space. Thus observable dynamical variables are characterized by Hermitian operators. It can easily be verified that all of the wave function operators introduced so far (e.g. position, energy, momentum, and force) are Hermitian.

Eigenvalue Problem for Hermitian Operators. Recall that the variance of Q about $\langle Q \rangle$ is given by

$$\begin{aligned} \sigma_Q^2 &= \langle (\tilde{Q} - \langle Q \rangle)^2 \rangle \\ &= (\Psi, (\tilde{Q} - \langle Q \rangle)^2 \Psi) \\ &= ((\tilde{Q} - \langle Q \rangle)\Psi, (\tilde{Q} - \langle Q \rangle)\Psi) \\ &= \|(\tilde{Q} - \langle Q \rangle)\Psi\|^2. \end{aligned}$$

It therefore follows that the value of Q is certain (i.e. $\sigma_Q = 0$) if and only if Ψ is an eigenfunction of \tilde{Q} , in which case $\langle Q \rangle$ is the eigenvalue.

Eigenproblem for Hermitian operator \tilde{Q}

$$\tilde{Q}\Psi = \langle Q \rangle \Psi \quad (8.6.1)$$

EXAMPLE: Time independent Schrödinger : $\tilde{H}\Psi = E\Psi$

For operators \tilde{Q} with a discrete set of eigenvalues and eigenvectors, we have the following three important properties:

1. Eigenvalues are real
2. Eigenfunctions are (or can be made to be) orthonormal: $(\phi_k, \phi_m) = \delta_{km}$
3. Eigenfunctions are complete
 For any $\Psi(q) = \sum_{k=1}^{\infty} C_k \phi_k(q)$ then $C_k = (\phi_k, \Psi)$
 i.e $\lim_{m \rightarrow \infty} \|\Psi - \sum_{k=1}^m C_k \phi_k\| = 0$

For any such operator, $(\tilde{Q})^r$ has eigenvalues λ_k^r with eigenfunctions ϕ_k .

The fact that the set of eigenfunctions—denoted $\{\phi_k\}$ —form a complete basis for the space L^2 implies that any function ψ can be represented as a linear combination of these functions, (see property 3 above). We then that for an operator \tilde{Q} ,

$$\tilde{Q}\psi = \sum_{k=1}^{\infty} c_k \lambda_k, \phi_k.$$

This Fourier representation of a wave function in terms of the eigenfunctions of some operator Q will be important in the following analysis.

Characteristic Functions and Discrete Probability Distributions. We define the characteristic function for the probability density ρ of some random variable x as

Characteristic Function

$$a(\xi) := \langle e^{i\xi x} \rangle = \int_{-\infty}^{\infty} \rho(x) e^{i\xi x} dx. \quad (8.6.2)$$

Let \tilde{Q} be an operator for some dynamical variable Q of discrete spetrum, then

$$a(\xi) := \langle e^{i\xi Q} \rangle = (\Psi, e^{i\xi \tilde{Q}} \Psi)$$

but what does the exponential function of an operator even mean? The answer lies in the Taylor series expansion of analytic functions.

Let $F[\tilde{Q}](u)$ denote some analytic function of an operator, which is then operating on some ϕ . This notation is a bit awkward at first but we emphasize that $F[\tilde{Q}]$ is an operator, not a value and is subsequently applied to some operand u .

Taylor expanding F we obtain,

$$F[\tilde{Q}] = F_0 + \tilde{Q}F'(0) + \frac{\tilde{Q}^2}{2}F''(0) + \dots$$

but since $\tilde{Q}\phi_k = \lambda_k\phi_k$ we have,

$$\begin{aligned} F[\tilde{Q}](\phi_k) &= \left(F_0 + \tilde{Q}F'(0) + \frac{\tilde{Q}^2}{2}F''(0) + \dots \right) \phi_k \\ &= F(\lambda_k)\phi_k. \end{aligned}$$

Then for some operand u we can take advantage of the Fourier representation of u in terms of eigenfunctions of \tilde{Q} to get,

$$F[\tilde{Q}](u) = \sum b_k F(\lambda_k)\phi_k,$$

where b_k are the Fourier coefficients for representing the function u .

It then follows that the characteristic function $a(\xi)$ can be expressed,

$$\begin{aligned} a(\xi) &= (\Psi, e^{i\xi\tilde{Q}}\Psi) = \sum_{k=1}^{\infty} (\Psi, e^{i\xi\lambda_k} c_k \phi_k) \\ &= \sum_{k=1}^{\infty} e^{i\xi\lambda_k} c_k \overbrace{(\Psi, \phi_k)}^{c_k^*}. \end{aligned}$$

Recalling our original definition of characteristic function $a(\xi)$ however we have,

$$a(\xi) = \sum_{k=1}^{\infty} e^{i\xi\lambda_k} c_k \overbrace{(\Psi, \phi_k)}^{c_k^*} = \int_Q \rho(Q) e^{i\xi Q} dQ$$

we can thus see that the probability density function must take discrete values, non-zero only at λ_k with a relative probability of $|c_k|^2$. Thus $\rho(Q)$ is essentially the superposition of δ_{λ_k} and

$$\rho = \sum_{k=1}^{\infty} |c_k|^2 \delta(Q - \lambda_k).$$

The key point here is that energy is a DISCRETE dynamic variable. Not all dynamic variables corresponding to Hermitian operators are discrete (for example consider p).

Let Ψ be a wave function with an eigenfunction representation for some observable M , then

$$\Psi = \sum_{k=1}^{\infty} c_k^M \phi_k^M$$

essentially lists all possible quantum states of our system and coefficients c_k encode the probability of each state.

The Generalized Uncertainty Principle. We return briefly to uncertainty quantification and consider two dynamical variables Q and M with associated Hermitian operators \tilde{Q} , \tilde{M} . We can compute the variance of these variables as follows,

$$\begin{aligned} \sigma_Q^2 &= \langle (Q - \langle Q \rangle)^2 \rangle \\ &= ((\tilde{Q} - \langle Q \rangle)\Psi, (\tilde{Q} - \langle Q \rangle)\Psi) \end{aligned}$$

and similarly,

$$\sigma_M^2 = ((\tilde{M} - \langle M \rangle)\Psi, (\tilde{M} - \langle M \rangle)\Psi)$$

Define $\tilde{Q}' = (\tilde{Q} - \langle Q \rangle)$ and $\tilde{M}' = (\tilde{M} - \langle M \rangle)$. Then,

$$\begin{aligned}\sigma_M^2 \sigma_Q^2 &\geq \left(\frac{1}{2i} (\tilde{Q}'\Psi, \tilde{M}'\Psi) - (\tilde{M}'\Psi, \tilde{Q}'\Psi) \right)^2 \\ &\geq \left(\frac{1}{2i} (\Psi, (\tilde{Q}'\tilde{M}' - \tilde{M}'\tilde{Q}')\Psi) \right)^2 \\ &\geq \left(\frac{1}{2i} (\Psi, (\tilde{Q}\tilde{M} - \tilde{M}\tilde{Q})\Psi) \right)^2\end{aligned}$$

Note that the primes have been dropped in the last line. If operators \tilde{Q} , \tilde{M} commute then this last expression is zero. If this is the case for two observable quantities then the quantities are said to be *compatible* since they can be observed (known) simultaneously. Note that in general, operators commute if and only if their eigenfunctions coincide. If operators don't commute we can instead define a new operator—called the *commutator* of Q and M —as,

Commutator

$$[\tilde{Q}, \tilde{M}] = (\tilde{Q}\tilde{M} - \tilde{M}\tilde{Q}) \quad (8.6.3)$$

We can now rewrite the expression for $\sigma_M^2 \sigma_Q^2$ as,

$$\sigma_M^2 \sigma_Q^2 \geq \left(\frac{1}{2i} \langle [\tilde{Q}, \tilde{M}] \rangle \right)^2.$$

Suppose now that x and p are generalized momentum operators, we can show that the commutator of x and p is in fact a constant operator. Indeed,

$$\begin{aligned}[\tilde{x}, \tilde{p}]\phi &= x \frac{\hbar}{i} \frac{\partial \phi}{\partial x} - \frac{\hbar}{i} \frac{\partial}{\partial x} (x\phi) \\ &= x \cancel{\frac{\hbar}{i} \frac{\partial \phi}{\partial x}} - \frac{\hbar}{i} \frac{\partial \phi}{\partial x} - x \cancel{\frac{\hbar}{i} \frac{\partial \phi}{\partial x}} \\ &= -\frac{\hbar}{i} \frac{\partial \phi}{\partial x}\end{aligned}$$

Thus,

$$[\tilde{x}, \tilde{p}] = \hbar i. \quad (8.6.4)$$

These quantities are called incompatible since they cannot be observed (known) simultaneously.

8.7 Harmonic Oscillators and Ladder Operators

Consider a harmonic oscillator. The governing equation for this system is

$$m \frac{d^2 x}{dt^2} = -kx$$

where k is the linear spring constant. The solution then takes the form,

$$x = A \sin\left(\sqrt{\frac{k}{m}}t + \phi\right) \quad \text{and} \quad \omega = \sqrt{\frac{k}{m}}$$

How can we translate this to a quantum system? We begin by writing the Hamiltonian. We then have,

$$q = x \quad \text{and} \quad p = mv = m \frac{dx}{dt}$$

so,

$$H(q, p) = \frac{p^2}{2m} + \frac{kq^2}{2} \quad \text{and} \quad E = \frac{A^2 k}{2} = \frac{A^2 m \omega^2}{2}$$

Our operator \tilde{H} then takes the form

$$\begin{aligned} \tilde{H} &= \frac{\tilde{p}^2}{2m} + \frac{k\tilde{x}^2}{2} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial q^2} + \frac{1}{2} m \omega^2 q^2 \\ &= \frac{1}{2m} (\tilde{p}^2 + (m\omega\tilde{q})^2) \end{aligned} \quad (8.7.1)$$

Recall the time invariant Schrödinger's equation,

$$\tilde{H}\psi = E\psi,$$

our goal now is to find the eigenfunctions and eigenvalues of this system. While we could proceed by various methods, one that is particularly enlightening essentially involves 'factoring' the operators. If \tilde{p} and \tilde{q} were scalars we could simply factor (8.7.1) as,

$$H = \frac{1}{2m} (ip + m\omega q)(ip - m\omega q) \quad (8.7.2)$$

However if we were to multiply this out without reordering any terms we would see that we left with a commutator-like term. Indeed, modifying this idea slightly we have,

$$\tilde{H} = \frac{1}{2m} (i\tilde{p} + m\omega\tilde{q})(i\tilde{p} - m\omega\tilde{q}) + \underbrace{[\tilde{p}, \tilde{q}]}_{\text{Constant}}$$

but recall from the previous section that this commutator term is simply a constant. If we then define these 'factors' of \tilde{H} as,

$$\tilde{a}_{\pm} = \frac{1}{\sqrt{2\hbar m \omega}} (\mp i\tilde{p} + m\omega\tilde{q}) \quad (8.7.3)$$

then we have,

$$\tilde{H} \propto (\tilde{a}_+)(\tilde{a}_-) + C$$

As we will soon see, these operators \tilde{a}_+ and \tilde{a}_- have a surprising utility.

Ladder Operators. As it turns out, operators \tilde{a}_+ and \tilde{a}_- can be used to increment or decrement (in terms of energy) the quantum state of a system. We will show this now.

Observe,

$$\begin{aligned}\tilde{a}_- \tilde{a}_+ &= \frac{1}{\sqrt{2\hbar m\omega}} \left[\tilde{p}^2 + (m\omega\tilde{q})^2 - im\omega \underbrace{(\tilde{q}\tilde{p} - \tilde{p}\tilde{q})}_{=[\tilde{q}, \tilde{p}] = i\hbar} \right] \\ &= \frac{1}{\sqrt{2\hbar m\omega}} \left[\tilde{p}^2 + (m\omega\tilde{q})^2 + m\omega\hbar \right] \\ &= \frac{1}{\hbar\omega} \tilde{H} + \frac{1}{2}\end{aligned}$$

A similar result can be shown for $\tilde{a}_+ \tilde{a}_-$ but with a the sign changed on the $\frac{1}{2}$. We rewrite these relationships again,

Ladder Operators

$$\begin{aligned}\tilde{a}_- \tilde{a}_+ &= \frac{1}{\hbar\omega} \tilde{H} + \frac{1}{2} & \tilde{a}_+ \tilde{a}_- &= \frac{1}{\hbar\omega} \tilde{H} - \frac{1}{2} \\ [\tilde{a}_-, \tilde{a}_+] &= 1\end{aligned}\tag{8.7.4}$$

Schrödinger's equation can now be rewritten,

$$\tilde{H}\psi = \hbar\omega \left(\tilde{a}_\pm \tilde{a}_\mp \pm \frac{1}{2} \right) \psi = E\psi.$$

Now consider,

$$\begin{aligned}\tilde{a}_+ E\psi &= \hbar\omega \tilde{a}_+ \left(\tilde{a}_- \tilde{a}_+ + \frac{1}{2} \right) \psi \\ &= \hbar\omega \tilde{a}_+ \left(\underbrace{\tilde{a}_- \tilde{a}_+}_{=E/\hbar\omega} - \frac{1}{2} + 1 \right) \psi \\ &= \hbar\omega \tilde{a}_+ \left(\frac{E}{\hbar\omega} + 1 \right) \psi \\ &= (E + \hbar\omega) \tilde{a}_+ \psi\end{aligned}$$

Thus we have that applying \tilde{a}_+ to an eigenfunction ψ of the time invariant Schrödinger's equation returns another eigenfunction corresponding to a quantum state of energy one level higher. In general we have that

$$\begin{aligned}\psi' &= \tilde{a}_+^n \psi \text{ (where } \psi \text{ an eigenfunction with energy } E) \\ &\text{is another eigenfunction corresponding to energy } E + n\hbar\omega\end{aligned}$$

A similar result holds for the operator \tilde{a}_- , but where the energy is decremented with each application. It should thus be clear why \tilde{a}_+ and \tilde{a}_- are called ladder operators, each application either steps up or steps down the quantum state.

A natural question to ask now is if there is a limit to the number of times these operators can be applied. For example, intuition tells us that there is a lowest energy level (or ground state) and we should therefore not be permitted to apply \tilde{a}_- to this ground state. This is indeed the case as applying \tilde{a}_- to this ground state gives the zero eigenfunction $\phi = 0$ which conveys no meaningful information about our system. Indeed the lowest energy this system can have is

$$E_0 = \frac{1}{2}\hbar\omega$$

and

$$E_0 = 0$$

requires that $k = 0$ (no spring constant) or $m = \infty$ which is clearly not permitted.

It turns out however that there is no highest energy state as can be shown by assuming a highest state exists and showing that the resulting wave function ψ_∞ could not be normalized.

8.8 Summary

Photoelectric effect. Incoming EM radiation (photons) can free electrons from material. The energy of the liberated electron is then $h\nu - l_0$ where ν is the frequency of illuminating radiation, and l_0 is the ionization energy (energy required to strip an electron off an atom).

Quantized Emission and Absorption. Energy states of the electrons in an atom are discrete or quantized, and when electrons transition from a higher energy state to a lower energy state they emit a photon at a frequency characteristic of the change in energy.

Wave particle duality. Particles behave like waves and vice versa. An experiment that demonstrated this was the double slit experiment by Davisson and Gremmer where electrons sent through double slit showed interference pattern (wave) but discrete impulses were detected by the sensor (particle).

Statistics.

$$\begin{aligned}\langle \gamma \rangle &= \mathbb{E}_{\rho(\gamma)} = \int_{\gamma} \gamma \rho(\gamma) d\gamma && \text{(Expected Value)} \\ \sigma_{\gamma}^2 &= \int_{\gamma} (\gamma - \langle \gamma \rangle)^2 \rho(\gamma) d\gamma = \langle \gamma^2 \rangle - \langle \gamma \rangle^2 && \text{(Variance)}\end{aligned}$$

Schrödinger's Equation. Wave equation to describe wave-particle duality.

$$\underbrace{i\hbar \frac{\partial \Psi}{\partial t}}_{\text{Energy}} = - \underbrace{\frac{\hbar^2}{2m} \nabla^2 \Psi}_{\text{Kinetic}} + \underbrace{V(x)\Psi}_{\text{Potential}}$$

Ψ normalized through Born Rule,

$$\int |\Psi|^2 dx = 1 \quad (8.8.1)$$

and is interpreted as the probability that the particle is within a certain differential volume.

Observable Operators. For a dynamical variable to be observable its operator must be Hermitian, i.e. $(\psi, A\phi) = (A\psi, \phi) \quad \forall \psi, \phi \in H^m$.

$$\begin{aligned}x\Psi &= x\Psi && \text{(Position)} \\ p\Psi &= \left(\frac{\hbar}{i} \nabla\right) \Psi && \text{(Momentum)} \\ E\Psi &= \left(-\frac{\hbar}{i} \frac{\partial}{\partial t}\right) \Psi && \text{(Total Energy)} \\ K\Psi &= \left(-\frac{\hbar^2}{2m} \nabla^2\right) \Psi && \text{(Kinetic Energy)} \\ V\Psi &= V\Psi && \text{(Potential Energy)}\end{aligned} \quad (8.8.2)$$

Hamiltonian Operator. The Hamiltonian of a system is just the function of position and momentum that yields energy. For a particle then the Hamiltonian is simply the sum of kinetic and potential energies,

$$H(x, p) = \underbrace{\frac{p \cdot p}{2m}}_{\text{Kinetic}} + \underbrace{V(x)}_{\text{Potential}}.$$

Thus operator form is,

$$\tilde{H}(x, \frac{\hbar}{i} \frac{\partial}{\partial x}) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x).$$

The Correspondence Principle. The idea that since classical mechanics very accurate at large scales, it must be recovered in the appropriate sense:

$$\underbrace{m \frac{d\langle x \rangle}{dt}}_{mv} = \underbrace{\langle p \rangle}_{\text{Momentum}} \quad \text{and} \quad \underbrace{\frac{d\langle p \rangle}{dt}}_{ma} = \langle -\nabla V \rangle = \underbrace{\langle F \rangle}_{\text{Force}}$$

Time Invariant Schrödinger's Equation. Derived using separation of variables on $\Psi(x, t) = \psi(x)\phi(t)$. For time independent part energy must be constant leading to the following eigenvalue problem:

$$\tilde{H}\psi = E\psi$$

Since \tilde{H} Hermitian eigenfunctions are complete in L^2 .

Heisenberg Uncertainty Principle.

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}.$$

Dynamical Variables Dynamical variable functions take the general form,

$$\tilde{Q}(q_1, \dots, q_N, \underbrace{-i\hbar \frac{\partial}{\partial q_1}, \dots, -i\hbar \frac{\partial}{\partial q_N}}_{p_1}).$$

The eigenvalue problem for a dynamical variable is:

$$\tilde{Q}\psi = \langle Q \rangle \psi$$

For two dynamical variables Q and M we can define the *commutator* as the operator

$$[\tilde{Q}, \tilde{M}] = (\tilde{Q}\tilde{M} - \tilde{M}\tilde{Q}).$$

Note that $[\tilde{Q}, \tilde{M}] = 0$ if and only if \tilde{Q}, \tilde{M} have the same eigenfunctions. In this case Q and M are said to be *compatible* (i.e. they can be observed simultaneously). Otherwise dynamical variables are said to be *incompatible*.

Characteristic Functions and Eigenfunction Expansions. This is a bit detailed but the important thing to know is that if ψ can be represented as an eigenfunction expansion

$$\psi = \sum_0^{\infty} c_k \phi_k$$

where ϕ_k are eigenfunctions of a dynamical variable Q , then ϕ_k are the possible quantum states of variable Q and $|c_k|^2$ are the probabilities that state ϕ_k will be observed if measured.

Ladder Operators Ladder operators are defined as,

$$\tilde{a}_{\pm} = \frac{1}{\sqrt{2\hbar m\omega}} \left(\mp i\tilde{p} + m\omega\tilde{q} \right)$$

and can be used to increment or decrement the eigenfunction (quantum state):

$$\psi_{n+1} = \tilde{a}_+ \psi_n \qquad \psi_{n-1} = \tilde{a}_- \psi_n.$$

The following identities hold,

$$\tilde{a}_- \tilde{a}_+ = \frac{1}{\hbar\omega} \tilde{H} + \frac{1}{2} \qquad \tilde{a}_+ \tilde{a}_- = \frac{1}{\hbar\omega} \tilde{H} - \frac{1}{2}$$

$$[\tilde{a}_-, \tilde{a}_+] = 1.$$

Schrödinger's equation can then be expressed,

$$\tilde{H}\psi = \hbar\omega \left(\tilde{a}_{\pm} \tilde{a}_{\mp} \pm \frac{1}{2} \right) \psi = E\psi.$$

Exercise

Consider a particle of mass m constrained to move only along the x axis in a potential given by:

$$V(x) = \begin{cases} 0 & 0 \leq x \leq a \\ \infty & \text{otherwise.} \end{cases}$$

- (a) Write Schrodinger equation (time independent) describing the motion of the particle (with boundary conditions).

Solution: We begin by writing the Hamiltonian of the system. Clearly the particle is confined to the interval $(0, a)$ and $V(x) = 0$ in this region so the Hamiltonian will simply be the kinetic energy:

$$\begin{aligned} H &= \frac{p^2}{2m} \\ \Rightarrow \tilde{H} &= \frac{\tilde{p}^2}{2m} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} \end{aligned}$$

Thus we get the time harmonic Schrodinger's equation as,

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} = E\psi \quad \psi(0) = \psi(a) = 0$$

- (b) Determine the energy levels and associated wave functions.

Solution: We need to solve the eigenvalue problem in part (a). Two methods could work here, we could simply guess a solution form and plug it in (since we know the solution will be given by sinusoids) or we can simply solve the ODE (in higher-dimensions we would proceed by separation of variables).

Substitute μ^2 for E (this enforces positivity of the RHS). Then,

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{2m\mu^2}{\hbar^2} \psi$$

implies,

$$\psi(x) = C_1 \sin\left(\frac{\sqrt{2m\mu}}{\hbar} x\right) + C_2 \cos\left(\frac{\sqrt{2m\mu}}{\hbar} x\right)$$

Applying boundary conditions we substitute $x = 0$ and see,

$$\psi(0) = C_2 = 0.$$

Substituting now $x = a$ now requires that

$$C_1 \sin\left(\frac{\sqrt{2m\mu}}{\hbar} a\right) = 0$$

which implies that,

$$\frac{\sqrt{2m\mu}}{\hbar} = \frac{n\pi}{a} \quad \text{for } n \in \mathbb{Z}$$

$$\implies E = \mu^2 = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$

and finally we must have $||\psi||^2 = 1$, thus

$$\begin{aligned} C_1^2 \int_0^a \sin^2 \left(\frac{n\pi x}{a} \right) dx &= 1 \\ \implies C_1 &= \sqrt{\frac{2}{a}} \end{aligned}$$

thus we have,

$$\psi_n = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}, \quad E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$

(c) Now let the quantum state of the particle be defined by the wave function,

$$\psi = Ax(a - x).$$

- (i) What is the value of A ?
 - (ii) What are the expected values of the momentum and energy of the particle?
 - (iii) A measurement of the energy of the particle will be performed. What is the probability that the result of the measurement will be the ground state energy from part (b)?
-

Solution:

- (i) We must have that $||\psi||^2 = 1$, thus

$$\begin{aligned} A^2 \int_0^a (x(a - x))^2 dx &= 1 \\ \implies A &= \sqrt{\frac{30}{a^5}} \end{aligned}$$

- (ii) Observe,

$$\begin{aligned} \langle x \rangle &= \int_0^a \psi^* x \psi dx \\ &= \frac{30}{a^5} \int_0^a x^3 (a - x)^2 dx = \frac{a}{2} \\ \langle p \rangle &= \int_0^a \psi^* \frac{\hbar}{i} \frac{d}{dx} \psi dx \\ &= \frac{30}{a^5} \frac{\hbar}{i} \int_0^a x(a - x)(a - 2x) dx = 0 \end{aligned}$$

- (iii) We make use of the result on Fourier representations of wave functions that $P_n = |c_n|^2$. Recall from the definition of the generalized Fourier transform that,

$$c_n = (\psi, \phi_n).$$

Using the eigenfunctions solved for in part (b) we get,

$$c_n = \sqrt{\frac{60}{a^6}} \int_0^a x(a-x) \sin\left(\frac{n\pi x}{a}\right) dx$$

$$= \begin{cases} \frac{8\sqrt{15}}{n^3\pi^3} & n \text{ odd} \\ 0 & n \text{ even} \end{cases}$$

Thus,

$$P_1 = c_1^2 = \frac{960}{\pi^6} = 0.9986$$

(d) Now suppose there are three identical, non-interacting particles.

- (i) Suppose all the particles are electrons (fermions with spin 1/2). By Pauli's exclusion principle no more than 2 electrons can be in the first energy level. What is the ground state of this system?
- (ii) Suppose all the particles are photons (bosons with spin 1). The Pauli exclusion principle doesn't apply to bosons. What is the ground state energy of the system now?

Solution:

- (i) Simply add up the different energy levels corresponding to each state. Since we can only have 2 electrons in the first orbital we have energy states (1,1,2) and

$$E = \frac{(1^2 + 1^2 + 2^2)\pi^2\hbar^2}{2ma^2} = \frac{3\pi^2\hbar^2}{ma^2}$$

- (ii) Now we have energy states (1,1,1) and,

$$E = \frac{(1^2 + 1^2 + 1^2)\pi^2\hbar^2}{2ma^2} = \frac{3\pi^2\hbar^2}{2ma^2}$$

(e) Suppose now we have a single particle in the infinite well in its ground state,

$$\psi = \sqrt{\frac{2}{a}} \sin \frac{\pi}{a} x,$$

and that the well suddenly expands to twice its original size (right boundary moves from a to $2a$). The energy of the system is now measured.

- (i) The most probable result for the energy is E_2 , what is the probability of this result?
- (ii) How would you determine the probability of getting the next most probable result for the energy?

Solution:

- (i) Note in particular that the domain of integration changed and we must renormalize our eigenfunctions to satisfy the Born rule:

$$\begin{aligned} \|\psi_n\|^2 &= C^2 \int_0^{2a} \sin^2 \left(\frac{n\pi x}{a} \right) = 1 \\ \Rightarrow \quad \psi_n &= \sqrt{\frac{1}{a}} \sin \frac{n\pi x}{a} \end{aligned}$$

but note that we are actually leaving out an eigenfunctions that would correspond to multiples of $1/2$ so we also change our frequency to fix this,

$$\psi_n = \sqrt{\frac{1}{a}} \sin \frac{n\pi x}{2a}$$

This could have been done rigorously by resolving eigenvalue problem. Taking now (note integral only to a since ψ before expansion not defined on larger interval),

$$P_2 = (\psi, \psi_2)^2 = \left(\frac{\sqrt{2}}{a} \int_0^a \sin^2 \frac{\pi x}{a} \right)^2 = \frac{1}{2}.$$

- (ii) In general,

$$P_n = (\psi, \psi_n)^2 = \left(\frac{\sqrt{2}}{a} \int_0^a \sin \frac{\pi x}{a} \sin \frac{n\pi x}{2a} \right)^2.$$

But note that for $n > 2$ even this will be 0 by orthogonality of sine basis. For $n = 1$ we get

$$P_1 = \frac{32}{9\pi^2} \approx 0.3603.$$

and the energy corresponding to this state is

$$E_1 = \frac{\pi^2 \hbar^2}{2m(2a)^2} = \frac{\pi^2 \hbar^2}{8m(a)^2}$$

- (f) Suppose now that the particle is in the quantum state described by the wave function

$$\psi = c\sqrt{\frac{2}{a}} \sin \left(\frac{2\pi x}{a} \right) + d\sqrt{\frac{2}{a}} \sin \left(\frac{3\pi x}{a} \right).$$

- (i) How must c and d be related?
- (ii) What is the expected value of the energy
- (iii) A measurement of the energy is to be made, what are the possible outcomes of such a measurement, and what are the probabilities of obtaining each?

Solution:

(i) The Born rule (normalization) implies the relationship:

$$\begin{aligned} 1 = (\psi, \psi) &= \frac{2}{a} \left(c \sin \left(\frac{2\pi x}{a} \right) + d \sin \left(\frac{3\pi x}{a} \right), c \sin \left(\frac{2\pi x}{a} \right) + d \sin \left(\frac{3\pi x}{a} \right) \right) \\ &= c^2 + d^2 \end{aligned}$$

by orthogonality of sine basis and evaluating inner products. Thus, $c^2 + d^2 = 1$.

(ii) Clearly ψ is just a linear combination of $\psi_2 + \psi_3$, so $\langle E \rangle = c^2 E_2 + d^2 E_3$

(iii) The possible outcomes are simply E_2 and E_3 with probabilities c^2 and d^2 respectively.

9. Atomic Structure

9.1 The Hydrogen Atom

We wish to find the quantum states of the hydrogen atom composed of one electron and one proton. First we consider the properties of each of these particles,

Electron:

- (-) charge, $e = 1.6 \times 10^{-19}$ C
- $m_e = 9.1 \times 10^{-31}$ kg

Proton:

- $m_p = 1.7 \times 10^{-27}$ kg
- $\epsilon_0 = 8.9 \times 10^{-12}$ C²/Nm

Notice in particular that the proton is roughly 1000x heavier. This will simplify our analysis as we can assume that the proton and center of mass of the atom remain fixed. Note that this is an assumption, a typical analysis of an n particle system would require us to solve for a $3n$ dimensional wave function. Without this assumption our wave function for the hydrogen atom would already require a 6-D analysis.

We begin as for the harmonic oscillator by determining the Hamiltonian,

$$H = \frac{|p|^2}{2m} + V(\mathbf{r}) \quad (9.1.1)$$

We can obtain the force acting on the electron using Coulomb's law,

$$F = -\frac{e^2 \mathbf{r}}{4\pi\epsilon_0 r^3}$$

but by the correspondence principle,

$$F = -\nabla V = -\nabla \left(-\frac{e^2}{4\pi\epsilon_0 r} \right)$$

which implies that,

$$V = -\frac{e^2}{4\pi\epsilon_0 r}. \quad (9.1.2)$$

We substitute this potential into the Hamiltonian (9.1.1) and convert to operator form to obtain,

$$H = \frac{|p|^2}{2m} - \frac{e^2}{4\pi\epsilon_0 r}.$$

We can now consider the time independent Schrödinger's equation,

$$\tilde{H}\psi = \left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r} \right)\psi = E\psi \quad (9.1.3)$$

Switching to polar coordinates to simplify our analysis (see Appendix for coordinate transforms). Then,

$$-\frac{\hbar^2}{2m}\nabla^2\psi = (E - V)\psi.$$

becomes,

$$-\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} \right] + V\psi = E\psi$$

We can attempt to solve this beast of an equation using separation of variables by letting $\psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$. We neglect to show the process of separation by variables here and skip to the two independent radial and angular equations.

Radial Equation. The radial portion of our equation then becomes,

$$\frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{2mr^2}{\hbar^2} (V - E)R - kR = 0. \quad (9.1.4)$$

We can assume $R \propto e^{-\alpha r}$ which gives,

$$\left(\alpha^2 + \frac{2mE}{\hbar^2} - \frac{2\alpha}{r} + \frac{2me^2}{4\pi\epsilon_0\hbar^2 r} - \frac{k}{r^2} \right) R = 0.$$

To solve for these various coefficients we can equate like powers of r (since the relation must hold for all r). This gives,

$$\begin{aligned} k &= 0 \\ \alpha &= \frac{me^2}{4\pi\epsilon_0\hbar^2} = \frac{1}{a_0} & a_0 &\approx 0.53 \times 10^{-10} \text{ (Bohr radius)} \\ E_1 &= -\frac{\hbar^2\alpha^2}{2m} = -\frac{me^4}{(4\pi\epsilon_0)^2 2\hbar} & &\approx 2.17 \times 10^{-18} \text{ J } (-13.6 \text{ eV}) \end{aligned}$$

To obtain additional eigenvalues we can assume $R \propto e^{-\alpha_n r}$, which leads to,

$$E_n = \frac{E_1}{n^2} \quad \left(\text{with } E_1 = -\frac{me^4}{(4\pi\epsilon_0)^2 2\hbar} \right)$$

with corresponding eigenfunctions. Now consider the case where $R \propto P_m(r)e^{-\alpha_n r}$ where $P_m(r)$ is a polynomial of r . Then we can derive the following results:

$$\begin{aligned} \alpha_n &= \frac{1}{na_0} & P_m &= \left(\frac{2r}{na_0} \right)^l \underbrace{L_{n-l}^{1+2l} \left(\frac{2r}{na_0} \right)}_{\text{associated Laguerre poly}} & E_n &= \frac{E_1}{n^2} \\ n &= 1, \dots, \infty & l &= 0, 1, \dots, n-1 & l &= l(l+1) \end{aligned}$$

$$\Delta E = E_1 \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right) \quad (9.1.5)$$

A. Calculus Identities

The following reference pages were created by my good friend and mentor, Dr. Vladimir Solovjov and are available with many other engineering and math resources on his webpage, <http://www.et.byu.edu/~vps/ME505>.

Let $\mathbf{a}(\mathbf{r}), \mathbf{b}(\mathbf{r}), \mathbf{F}(\mathbf{r}) : \mathbb{R}_3 \rightarrow \mathbb{R}_3$ be vector fields, $\varphi(\mathbf{r}), u(\mathbf{r}) : \mathbb{R}_3 \rightarrow \mathbb{R}$ be scalar fields, $c \in \mathbb{R}$ [see B&T, p.168]

$\nabla \varphi$	$= \text{grad } \varphi$	$\frac{\partial \varphi}{\partial x_j}$
$\nabla \cdot \mathbf{a}$	$= \text{div } \mathbf{a}$	$\delta_{kj} \frac{\partial a_k}{\partial x_j} = \frac{\partial a_k}{\partial x_k}$
$\nabla \times \mathbf{a}$	$= \text{curl } \mathbf{a}$	$\varepsilon_{ijk} \frac{\partial a_k}{\partial x_j}$
$\nabla \cdot \nabla \varphi$	$= \text{div grad } \varphi \equiv \Delta \varphi \equiv \nabla^2 \varphi$	$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_i} \varphi = \partial_i \partial_i \varphi$ <i>Laplacian operator</i>
$\nabla \times \nabla \varphi$	$= \text{curl grad } \varphi = \mathbf{0}$	<i>vanishes identically</i>
$\nabla (\nabla \cdot \mathbf{a})$	$= \text{grad div } \mathbf{a}$	$\delta_{jk} \frac{\partial^2 a_j}{\partial x_i \partial x_k} = \frac{\partial}{\partial x_i} \left(\frac{\partial a_k}{\partial x_k} \right)$
$\nabla \cdot (\nabla \times \mathbf{a})$	$= \text{div curl } \mathbf{a} = 0$	<i>vanishes identically</i>
$\nabla \times (\nabla \times \mathbf{a})$	$= \text{curl curl } \mathbf{a}$	$\varepsilon_{ijk} \frac{\partial}{\partial x_j} \left(\varepsilon_{klm} \frac{\partial a_m}{\partial x_l} \right) = \varepsilon_{ijk} \varepsilon_{klm} \frac{\partial^2 a_m}{\partial x_j \partial x_l}$
1.	$\Delta(\varphi + \psi) = \Delta \varphi + \Delta \psi$	
2.	$\nabla(\varphi + \psi) = \nabla \varphi + \nabla \psi$	$\text{grad}(\varphi + \psi) = \text{grad } \varphi + \text{grad } \psi$
3.	$\nabla[\nabla \cdot (\mathbf{a} + \mathbf{b})] = \nabla(\nabla \cdot \mathbf{a}) + \nabla(\nabla \cdot \mathbf{b})$	$\text{grad div}(\mathbf{a} + \mathbf{b}) = \text{grad div } \mathbf{a} + \text{grad div } \mathbf{b}$
4.	$\nabla \cdot (\mathbf{a} + \mathbf{b}) = \nabla \cdot \mathbf{a} + \nabla \cdot \mathbf{b}$	$\text{div}(\mathbf{a} + \mathbf{b}) = \text{div } \mathbf{a} + \text{div } \mathbf{b}$
5.	$\nabla \cdot (c\mathbf{a}) = c \nabla \cdot \mathbf{a}$	$\text{div}(c\mathbf{a}) = c \text{div } \mathbf{a}$
6.	$\nabla \times (\mathbf{a} + \mathbf{b}) = \nabla \times \mathbf{a} + \nabla \times \mathbf{b}$	$\text{curl}(\mathbf{a} + \mathbf{b}) = \text{curl } \mathbf{a} + \text{curl } \mathbf{b}$
7.	$\nabla \times (\nabla \times (\mathbf{a} + \mathbf{b})) = \nabla \times (\nabla \times \mathbf{a}) + \nabla \times (\nabla \times \mathbf{b})$	$\text{curl curl}(\mathbf{a} + \mathbf{b}) = \text{curl curl } \mathbf{a} + \text{curl curl } \mathbf{b}$
8.	$\nabla \times (\nabla \times \mathbf{a} + \nabla \varphi) = \nabla \times (\nabla \times \mathbf{a})$	$\text{curl}(\text{curl } \mathbf{a} + \text{grad } \varphi) = \text{curl}(\text{curl } \mathbf{a})$
9.	$\nabla(\varphi \psi) = \psi \nabla \varphi + \varphi \nabla \psi$	$\text{grad}(\varphi \psi) = \psi \text{grad } \varphi + \varphi \text{grad } \psi$
10.	$\nabla \cdot (\varphi \mathbf{a}) = \varphi \nabla \cdot \mathbf{a} + \mathbf{a} \cdot \nabla \varphi$	$\text{div}(\varphi \mathbf{a}) = \varphi \text{div } \mathbf{a} + \mathbf{a} \cdot \text{grad } \varphi$
11.	$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b})$	$\text{div}(\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot \text{curl } \mathbf{a} - \mathbf{a} \cdot \text{curl } \mathbf{b}$
12.	$\nabla \times (\varphi \mathbf{a}) = \varphi (\nabla \times \mathbf{a}) + \nabla \varphi \times \mathbf{a}$	$\text{curl}(\varphi \mathbf{a}) = \varphi \text{curl } \mathbf{a} + \text{grad } \varphi \times \mathbf{a}$
13.*	$\nabla \times (\nabla \times \mathbf{a}) = \nabla(\nabla \cdot \mathbf{a}) - \overbrace{(\nabla \cdot \nabla)}^{\nabla^2 \text{ Laplacian}} \mathbf{a}$	$\text{curl}(\text{curl } \mathbf{a}) = \text{grad}(\text{div } \mathbf{a}) - \overbrace{\Delta \mathbf{a}}^{\text{Laplacian}}$
14.	$\nabla(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) + (\mathbf{a} \cdot \nabla) \mathbf{b} + (\mathbf{b} \cdot \nabla) \mathbf{a}$	$= \overbrace{\nabla \mathbf{a} \mathbf{b}}^{\text{dyadic}} + \overbrace{\nabla \mathbf{b} \mathbf{a}}^{\text{dyadic}}$
15.	$\nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a})$	$\text{curl}(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \text{div } \mathbf{b} - \mathbf{b} \text{div } \mathbf{a}$

For composite functions $\varphi[f(\mathbf{r})]$ and $\mathbf{a}[f(\mathbf{r})]$, the chain rule is applied

16.	$\nabla \varphi[f(\mathbf{r})] = \frac{d\varphi}{df} \nabla f$	$\text{grad } \varphi[f(\mathbf{r})] = \frac{d\varphi}{df} \text{grad } f$
17.	$\nabla \cdot \mathbf{a}[f(\mathbf{r})] = (\nabla f) \cdot \frac{d\mathbf{a}}{df}$	$\text{div } \mathbf{a}[f(\mathbf{r})] = \text{grad } f \cdot \frac{d\mathbf{a}}{df}$
18.	$\nabla \times \mathbf{a}[f(\mathbf{r})] = (\nabla f) \times \frac{d\mathbf{a}}{df}$	$\text{curl } \mathbf{a}[f(\mathbf{r})] = \text{grad } f \times \frac{d\mathbf{a}}{df}$

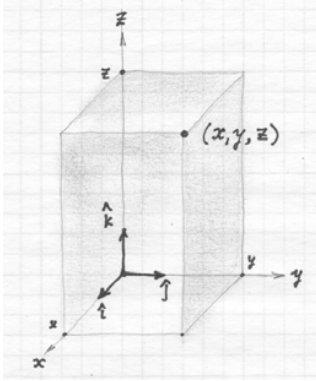
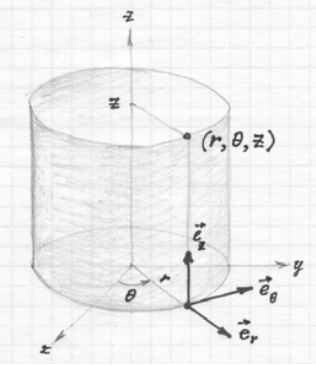
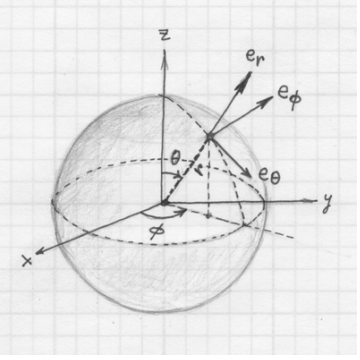
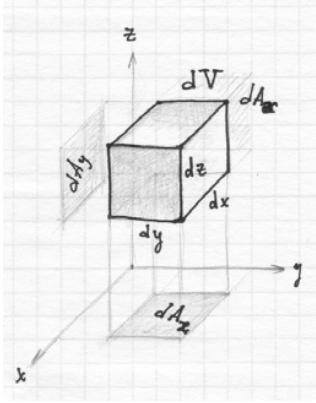
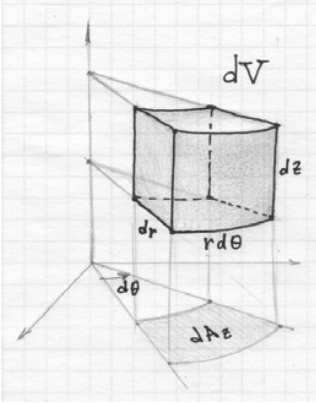
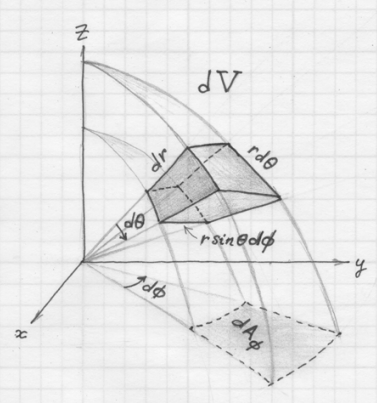
* $[\nabla \times (\nabla \times \mathbf{a})]_i = \varepsilon_{ijk} \varepsilon_{ilm} \frac{\partial}{\partial x_j} \left(\frac{\partial a_m}{\partial x_l} \right) = [\nabla(\nabla \cdot \mathbf{a})]_i - [(\nabla \cdot \nabla) \mathbf{a}]_i$ tensor notations for #13.

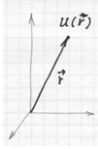
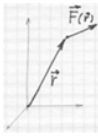
Laplacian of a vector field can be calculated $(\nabla \cdot \nabla) \mathbf{a} = \nabla(\nabla \cdot \mathbf{a}) - \nabla \times (\nabla \times \mathbf{a})$ using gradient, divergence and curl operators

B. Coordinate Systems

The following reference pages were created by my good friend and mentor, Dr. Vladimir Solovjov and are available with many other engineering and math resources on his webpage, <http://www.et.byu.edu/~vps/ME505>.

Coordinate Systems

<p>Cartesian coordinates (x, y, z)</p> 	<p>Cylindrical coordinates (r, θ, z)</p>  <p> $x = r \cos \theta$ $y = r \sin \theta$ $z = z$ $r^2 = x^2 + y^2$ $\tan \theta = \frac{y}{x}$ $z = z$ </p>	<p>Spherical coordinates (r, ϕ, θ)</p>  <p> $x = r \cos \phi \sin \theta$ $y = r \sin \phi \sin \theta$ $z = r \cos \theta$ $r^2 = x^2 + y^2 + z^2$ $\tan \phi = \frac{y}{x}$ $\tan \theta = \frac{z}{r} = \frac{z}{\sqrt{x^2 + y^2 + z^2}}$ </p>
<p>Basic vectors</p> <p> $\mathbf{i} = (1, 0, 0)$ $\mathbf{j} = (0, 1, 0)$ $\mathbf{k} = (0, 0, 1)$ </p>	<p> $\mathbf{e}_r = \mathbf{i} \cos \theta + \mathbf{j} \sin \theta$ $\mathbf{e}_\theta = -\mathbf{i} \sin \theta + \mathbf{j} \cos \theta$ $\mathbf{e}_z = \mathbf{k}$ </p>	<p> $\mathbf{e}_r = \mathbf{i} \cos \phi \sin \theta + \mathbf{j} \sin \phi \sin \theta + \mathbf{k} \cos \theta$ $\mathbf{e}_\theta = -\mathbf{i} \sin \theta + \mathbf{j} \cos \theta$ $\mathbf{e}_\phi = \mathbf{i} \cos \phi \cos \theta + \mathbf{j} \sin \phi \cos \theta - \mathbf{k} \sin \theta$ </p>
 <p>Line elements dx, dy, dz</p> <p>Differential areas</p> <p> $dA_x = dydz$ $dA_y = dx dz$ $dA_z = dx dy$ </p> <p>Differential volume</p> <p>$dV = dx dy dz$</p> <p>Arc length</p> <p>$ds^2 = dx^2 + dy^2 + dz^2$</p>	 <p> $dr, r d\theta, dz$ </p> <p> $dA_r = r d\theta dz$ $dA_\theta = dr dz$ $dA_z = r d\theta dr$ </p> <p>$dV = r dr d\theta dz$</p> <p>$ds^2 = dr^2 + r^2 d\theta^2 + dz^2$</p>	 <p> $dr, r \sin \theta d\phi, r d\theta$ </p> <p> $dA_r = r^2 \sin \theta d\phi d\theta$ $dA_\theta = r \sin \theta d\phi dr$ $dA_\phi = r d\theta dr$ </p> <p>$dV = r^2 \sin \theta d\phi d\theta dr$</p> <p>$ds^2 = dr^2 + r^2 \sin^2 \theta d\phi^2 + r^2 d\theta^2$</p>

<p>scalar field $u(\mathbf{r})$</p>  <p>Gradient ∇u</p> <p>Laplacian $\nabla^2 u$</p>	<p>$u(x, y, z)$</p> $\nabla u = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right)$ $= \frac{\partial u}{\partial x} \mathbf{i} + \frac{\partial u}{\partial y} \mathbf{j} + \frac{\partial u}{\partial z} \mathbf{k}$ $\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$	<p>$u(r, \theta, z)$</p> $\nabla u = \left(\frac{\partial u}{\partial r}, \frac{1}{r} \frac{\partial u}{\partial \theta}, \frac{\partial u}{\partial z} \right)$ $= \frac{\partial u}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial u}{\partial \theta} \mathbf{e}_\theta + \frac{\partial u}{\partial z} \mathbf{e}_z$ $\nabla^2 u = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2}$	<p>$u(r, \phi, \theta)$</p> $\nabla u = \left(\frac{\partial u}{\partial r}, \frac{1}{r \sin \theta} \frac{\partial u}{\partial \phi}, \frac{1}{r} \frac{\partial u}{\partial \theta} \right)$ $= \frac{\partial u}{\partial r} \mathbf{e}_r + \frac{1}{r \sin \theta} \frac{\partial u}{\partial \phi} \mathbf{e}_\phi + \frac{1}{r} \frac{\partial u}{\partial \theta} \mathbf{e}_\theta$ $\nabla^2 u = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 u}{\partial \phi^2} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial u}{\partial \theta} \right)$
<p>vector field $\mathbf{F}(\mathbf{r})$</p>  <p>Divergence $\text{div} \mathbf{F} = \nabla \cdot \mathbf{F}$</p> <p>Curl $\text{curl} \mathbf{F} = \nabla \times \mathbf{F}$</p>	<p>(F_x, F_y, F_z)</p> $\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$ $\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathbf{k}$	<p>(F_r, F_θ, F_z)</p> $F_r = F_x \cos \theta + F_y \sin \theta$ $F_\theta = F_x \sin \theta - F_y \cos \theta$ $F_z = F_z$ $F_x = F_r \cos \theta - F_\theta \sin \theta$ $F_y = F_r \sin \theta + F_\theta \cos \theta$ $F_z = F_z$ $\frac{1}{r} \frac{\partial}{\partial r} (r F_r) + \frac{1}{r} \frac{\partial F_\theta}{\partial \theta} + \frac{\partial F_z}{\partial z}$ $\frac{1}{r} \frac{\partial}{\partial r} (r F_r) + \frac{1}{r} \frac{\partial F_\theta}{\partial \theta} + \frac{\partial F_z}{\partial z}$ $\begin{vmatrix} \mathbf{e}_r & r \mathbf{e}_\theta & \mathbf{e}_z \\ \frac{1}{r} \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial z} \\ F_r & r F_\theta & F_z \end{vmatrix} = \left(\frac{1}{r} \frac{\partial F_z}{\partial \theta} - \frac{\partial F_\theta}{\partial z} \right) \mathbf{e}_r + \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r} \right) \mathbf{e}_\theta + \frac{1}{r} \left[\frac{\partial (r F_\theta)}{\partial r} - \frac{\partial F_r}{\partial \theta} \right] \mathbf{e}_z$	<p>(F_r, F_ϕ, F_θ)</p> $F_r = F_x \cos \phi \sin \theta + F_y \sin \phi \sin \theta + F_z \cos \theta$ $F_\phi = -F_x \sin \phi + F_y \cos \phi$ $F_\theta = F_x \cos \phi \cos \theta + F_y \sin \phi \cos \theta - F_z \sin \theta$ $F_x = F_r \cos \phi \sin \theta - F_\phi \sin \phi + F_\theta \cos \phi \cos \theta$ $F_y = F_r \sin \phi \sin \theta + F_\phi \cos \phi + F_\theta \sin \phi \cos \theta$ $F_z = F_r \cos \theta - F_\theta \sin \theta$ $\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 F_r) + \frac{1}{r \sin \theta} \frac{\partial F_\phi}{\partial \phi} + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} [(\sin \theta) F_\theta]$ $\frac{1}{r^2 \sin \theta} \begin{vmatrix} \mathbf{e}_r & r \mathbf{e}_\theta & r \sin \theta \mathbf{e}_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ F_r & r F_\theta & r \sin \theta F_\phi \end{vmatrix} = \frac{1}{r \sin \theta} \left[\frac{\partial (F_\phi \sin \theta)}{\partial \theta} - \frac{\partial F_\theta}{\partial \phi} \right] \mathbf{e}_r + \frac{1}{r} \left[\frac{\partial (r F_\theta)}{\partial r} - \frac{\partial F_r}{\partial \theta} \right] \mathbf{e}_\phi + \frac{1}{r \sin \theta} \left[\frac{\partial F_r}{\partial \phi} - \sin \theta \frac{\partial (r F_\phi)}{\partial r} \right] \mathbf{e}_\theta$