General Math Tricks and Other Notes

Lynn B. Wilson III

November 13, 2015

Table of Contents

1 Conversion Factors

The following are useful conversion factors:

$$1 \frac{km\ T}{s} = 10^{-3}\ \frac{mV}{m} = 1 \frac{\mu V}{m} \tag{1.1a}$$

$$1 \frac{eV}{cm^3} = 1.60217646 \times 10^{-4}\ \frac{J}{km^3} \tag{1.1b}$$

$$1 \frac{\mu W}{m^2} = 6.24150974 \times 10^{0}\ \frac{keV\ km}{s\ cm^3} \tag{1.1c}$$

$$1 \frac{keV\ km}{s\ cm^3} = 1.60217646 \times 10^{-7}\ \frac{erg}{s\ cm^2} \tag{1.1d}$$

$$1 \frac{erg}{s\ cm^2} = 10^{3}\ \frac{\mu W}{m^2} \tag{1.1e}$$

$$1 \frac{s^3}{km^3\ cm^3} = 10^{-3}\ \frac{s^3}{m^6} \tag{1.1f}$$

$$1 \frac{\mu V}{cm} = 10^{-1}\ \frac{mV}{m} \tag{1.1g}$$

Let us define q_s = Z e [≡ charge of particle species s], Z ≡ number of unit charges, $m_e(M_s)$ ≡ mass of electron(ion species s), $\mu \equiv M_i/M_p$, n_s ≡ number density of particle species s, T_s ≡ average temperature of particle species s, and B_o ≡ magnitude of the quasi-static magnetic field. In addition, let us define $\omega_{ps} = 2\pi f_{ps} = \sqrt{n_s {q_s}^2/(m_s \varepsilon_o)}$ [≡ plasma frequency of particle species s], $\Omega_{cs} = 2\pi f_{cs} = q_s B_o/m_s$ [≡ cyclotron frequency of particle species s], $V_{Ts} = \sqrt{(2 k_B T_s)/m_s}$ [≡ average thermal speed of particle species s], $\lambda_s = c/\omega_{ps}$ [≡ inertial length (or skin depth) of particle species s], $\lambda_{Ds} = \sqrt{(\varepsilon_o k_B T_s)/(n_s {q_s}^2)}$ [≡ Debye length of particle species s], $\rho_{cs} = V_{Ts}/\omega_{ps}$ [≡ thermal gyroradius of particle species s], and $V_A = \sqrt{{B_o}^2/(\mu_o M_i n_i)}$ [≡ Alfvén speed][1].

Below, the units are defined as follows: all frequencies are in Hz; distances in meters; speeds in km/s; temperatures in eV; magnetic fields in nT; and densities in cm^{-3}. The approximate factors are:

$$f_{pe} \cong 8.9787 \times 10^{3} \sqrt{n_e} \tag{1.2}$$

$$f_{pi} \cong 209.5353 \sqrt{\frac{Z^2 n_i}{\mu}} \tag{1.3}$$

$$f_{ce} \cong 27.9925\ B_o \tag{1.4}$$

$$f_{ci} \cong 1.5245 \times 10^{-2}\ \frac{Z}{\mu} B_o \tag{1.5}$$

[1] we also refer to an electron Alfvén speed, V_{Ae}, on occasion, but it does not have the same physical significance as V_A

$$\rho_{ce} \cong 3.3721 \, \frac{\sqrt{T_e}}{B_o} \tag{1.6}$$

$$\rho_{ci} \cong 144.4970 \, \frac{\sqrt{\mu T_i}}{Z B_o} \tag{1.7}$$

$$\lambda_e \cong 5.3141 \times 10^3 \, {n_e}^{-1/2} \tag{1.8}$$

$$\lambda_i \cong 2.2771 \times 10^5 \, \sqrt{\frac{\mu}{Z^2 n_i}} \tag{1.9}$$

$$\tilde{\lambda}_{De} = \frac{V_{Te}}{\sqrt{2}\omega_{pe}} \cong 7.4339 \, \sqrt{\frac{T_e}{n_e}} \tag{1.10}$$

$$\lambda_{Ds} = \frac{V_{Ts}}{\omega_{pe}} \cong 10.5132 \, \sqrt{\frac{T_s}{{Z_s}^2 n_s}} \tag{1.11}$$

$$V_{Te} \cong 593.0970 \, \sqrt{T_e} \tag{1.12}$$

$$V_{Ti} \cong 13.8411 \, \sqrt{\frac{T_i}{\mu}} \tag{1.13}$$

$$\frac{\omega_{pe}}{\Omega_{ce}} \cong 3.21 \times 10^2 \, \frac{\sqrt{n_e(cm^{-3})}}{B(nT)} \tag{1.14}$$

$$\beta_s \cong 0.403 \, \frac{n_s T_s}{{B_o}^2} \tag{1.15}$$

The following are useful relationships:

$$\omega_{pe} = \left(\frac{c}{V_{Ae}} \right) \Omega_{ce} \tag{1.16a}$$

$$\eta = \frac{\nu}{\varepsilon_o {\omega_{pe}}^2} \tag{1.16b}$$

2 General Mathematical Rules

2.1 The Dirac Delta Function

Definition = a mathematically *improper* function having the properties :

1. $\delta\big(x - a\big) = 0$ for $x \neq a$
2. $\int \delta\big(x - a\big)\, dx = 1$ (*if region includes x = a which we'll assume from here on, otherwise it is zero*)
3. $\int dx\, f(x)\, \delta\big(x - a\big) = f(a)$
4. $\int dx\, f(x)\, \delta'\big(x - a\big) = -f'(a)$
5. The delta function transforms according to the rule seen in Equation 2.8, assuming f(x) only has simple zeros located at x = x_i.
6. In more than one dimension, the delta function can be written as seen in Equation 2.9
7. The delta function has the inverse units of whatever the delta function happens to be a function of ⇒ the delta function in Equation 2.9 has the units of an inverse volume
8. One can expand a delta function in a Taylor series according to the rules defined in Equations (2.11a - 2.11d)
9. Typically one assumes that $\nabla^2(1/r) = 0$, assuming $r \neq 0$ and its volume integral is equal to -4π. One can then use the properties of the delta function to say $\nabla^2(1/r) = -4\pi\, \delta(\mathbf{x})$. A more general version can be seen in Equation 2.10.

$$\delta\left(x' - x\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{ik(x'-x)} \tag{2.1}$$

$$\frac{1}{k}\delta\left(k - k'\right) = \int_{0}^{\infty} d\rho\rho J_{\nu}\big(k\rho\big) J_{\nu}\big(k'\rho\big) \tag{2.2}$$

where J_ν are *Bessel Functions* and $\mathrm{Re}\{\nu\} > -1$.

$$\frac{d^n \delta\left(x' - x\right)}{dx^n} = \delta\left(x' - x\right)\frac{d^n}{dx^n} \tag{2.3}$$

$$\delta\left(xa\right) = \frac{\delta\left(x\right)}{|a|} \tag{2.4}$$

$$\delta'\left(x' - x\right) = \frac{d}{dx}\delta\left(x' - x\right) = -\frac{d}{dx'}\delta\left(x' - x\right) \tag{2.5}$$

$$\delta\left(x' - x\right) = \frac{d}{dx}\Theta\left(x' - x\right) \tag{2.6}$$

where $\Theta\big($x' - x$\big)$ is the *Theta Function* which has the properties:

$$\Theta\big(x' - x\big) = \begin{cases} 0 & \text{if (x' - x)} < 0, \\ 1 & \text{if (x' - x)} > 0. \end{cases} \tag{2.7}$$

$$\delta\big(f(x)\big) = \sum_i \frac{1}{|\frac{df}{dx_i}|} \delta\left(x - x_i\right) \ \text{[x}_i \text{ are the zeros of f(x)]} \tag{2.8}$$

$$\delta\left(\mathbf{x} - \mathbf{x}'\right) = \delta\left(x_1 - x'_1\right)\delta\left(x_2 - x'_2\right)\delta\left(x_3 - x'_3\right) \tag{2.9}$$

$$\nabla^2\left(\frac{1}{|\mathbf{x} - \mathbf{x}'|}\right) = -4\pi\delta\big(\mathbf{x} - \mathbf{x}'\big) \tag{2.10}$$

$$\vec{r}_{i+\delta} = \vec{r}_i + \vec{r}_{\delta i} \tag{2.11a}$$

$$\frac{|\vec{r}_{\delta i}|}{|\vec{r}_{i+\delta}|} \ll 1 \tag{2.11b}$$

$$\delta\left(\vec{r} - \vec{r}_{i+\delta}\right) \to \delta\left(\vec{r} - \vec{r}_i - \vec{r}_{\delta i}\right) \tag{2.11c}$$

$$\approx \delta\left(\vec{r} - \vec{r}_i\right) - \vec{r}_{\delta i} \cdot \nabla_{\vec{r}}\Big(\delta\left(\vec{r} - \vec{r}_i\right)\Big) \tag{2.11d}$$

2.2 Vector and Tensor Calculus

If we have an arbitrary vector that is not coplanar with a plane that has a unit normal $\hat{\mathbf{n}}$, we can define the vector along the normal (subscript n) and transverse to the normal (subscript t) as:

$$Q_n = \mathbf{Q} \cdot \hat{\mathbf{n}} \tag{2.12a}$$
$$\mathbf{Q}_t = (\hat{\mathbf{n}} \times \mathbf{Q}) \times \hat{\mathbf{n}} \tag{2.12b}$$
$$= \mathbf{Q} \cdot (\mathbb{I} - \hat{\mathbf{n}}\hat{\mathbf{n}}) \tag{2.12c}$$

Note that advection is not necessarily the same as convection. Convection is the sum of the advective and diffusive effects of a fluid flow. Diffusion describes the spread of particles through random motion from regions of higher concentration to regions of lower concentration. Mathematically, this can be shown by considering the advection term, $\nabla \times \mathbf{V}$, separately from the convection term, $\mathbf{V} \cdot (\nabla\ \mathbf{V})$, because the convection term can be rewritten in the following form:

$$\mathbf{V} \cdot (\nabla \mathbf{V}) = \left[\nabla \frac{\mid \mathbf{V} \mid^2}{2} - (\nabla \mathbf{V}) \cdot \mathbf{V} \right] + (\mathbf{V} \cdot \nabla)\, \mathbf{V} \tag{2.13}$$

where we have used the vector identity:

$$\mathbf{A} \times (\nabla \times \mathbf{B}) = (\nabla \mathbf{B}) \cdot \mathbf{A} - (\mathbf{A} \cdot \nabla)\, \mathbf{B} \tag{2.14}$$

1. Assume that the vector $\mathbf{A}$ and the scalars, ψ and ϕ, are well behaved vector functions
2. V $\equiv$ 3D volume with volume element $\mathrm{d}^3\mathrm{x}$
3. S $\equiv$ is a closed 2D surface bounding volume V, with area element da
4. $\mathbf{n} \equiv$ unit *outward* normal vector at surface element da

$$\int_V d^3x\ \nabla \cdot \mathbf{A} = \int_S da\ \mathbf{n} \cdot \mathbf{A} \tag{2.15a}$$
$$\int_V d^3x\ \nabla\psi = \int_S da\ \mathbf{n}\psi \tag{2.15b}$$
$$\int_V d^3x\ \nabla \times \mathbf{A} = \int_S da\ \mathbf{n} \times \mathbf{A} \tag{2.15c}$$
$$\int_V d^3x \left[\mathbf{A} \cdot \Big(\nabla \times (\nabla \times \mathbf{B}) - \mathbf{B} \cdot \Big(\nabla \times (\nabla \times \mathbf{A}) \Big) \right] = \int_S da\ \mathbf{n} \cdot \Big[\mathbf{B} \times \Big(\nabla \times \mathbf{A} \Big) - \mathbf{A} \times \Big(\nabla \times \mathbf{B} \Big) \Big] \tag{2.15d}$$
$$\int_V d^3x \left(\phi\nabla^2\psi + \nabla\phi \cdot \nabla\psi \right) = \int_S da\ \phi\big(\mathbf{n} \cdot \nabla\psi\big)\ (\text{Green's } 1^{st} \text{ Identity}) \tag{2.15e}$$
$$\int_V d^3x \left(\phi\nabla^2\psi + \psi\nabla^2\phi \right) = \int_S da\ \phi\big(\phi\nabla\psi - \psi\nabla\phi\big)\ (\text{Green's Theorem}) \tag{2.15f}$$

1. In the following equations, we define $S \equiv$ open surface

2. $C \equiv$ contour bounding the open surface S, with line element d**l**
3. **n** $\equiv$ normal to the surface S with the direction defined by the *right-hand-screw rule* in relation to the direction of d**l** (i.e. the line integral around contour C)

$$\int_S da \left(\nabla \times \mathbf{A}\right) \cdot \mathbf{n} = \oint_C \mathbf{A} \cdot d\mathbf{l} \text{ (Stokes's Theorem)} \tag{2.16a}$$
$$\int_S da \left(\mathbf{n} \times \nabla\right)\psi = \oint_C \psi d\mathbf{l} \tag{2.16b}$$
$$\int_S da \left(\mathbf{n} \times \nabla\right) \times \mathbf{A} = \oint_C d\mathbf{l} \times \mathbf{A} \tag{2.16c}$$
$$\int_S da\, \mathbf{n} \cdot \left(\nabla f \times \nabla g\right) = \oint_C dg\, f = -\oint_C df\, g \tag{2.16d}$$

1. In the following equations, we define **x** $\equiv$ coordinate of some point with respect to some origin
2. r $\equiv$ the magnitude of **x** (= |**x**|)
3. **k** $\equiv$ **x**/r = unit radial vector
4. f(r) $\equiv$ a well-behaved function of r
5. **a** $\equiv$ an arbitrary vector
6. **L** $\equiv$ the angular momentum operator defined in Equation 2.17g

$$\nabla \cdot \mathbf{x} = 3 \tag{2.17a}$$
$$\nabla \times \mathbf{x} = 0 \tag{2.17b}$$
$$\nabla \cdot \left[\mathbf{n} f(r)\right] = \frac{2}{r} f(r) + \frac{\partial f}{\partial r} \tag{2.17c}$$
$$\nabla \times \left[\mathbf{n} f(r)\right] = 0 \tag{2.17d}$$
$$\left(\mathbf{a} \cdot \nabla\right)\mathbf{n} f(r) = \frac{f(r)}{r}\left[\mathbf{a} - \mathbf{n}\left(\mathbf{a} \cdot \mathbf{n}\right)\right] + \mathbf{n}\left(\mathbf{a} \cdot \mathbf{n}\right)\frac{\partial f}{\partial r} \tag{2.17e}$$
$$\nabla\left(\mathbf{x} \cdot \mathbf{a}\right) = \mathbf{a} + \mathbf{x}\left(\nabla \cdot \mathbf{a}\right) + i\left(\mathbf{L} \times \mathbf{a}\right) \tag{2.17f}$$
$$\mathbf{L} = -i\left(\mathbf{x} \times \nabla\right) \tag{2.17g}$$

$$\mathbf{a} \cdot \left(\mathbf{b} \times \mathbf{c}\right) = \mathbf{b} \cdot \left(\mathbf{c} \times \mathbf{a}\right) = \mathbf{c} \cdot \left(\mathbf{a} \times \mathbf{b}\right) \tag{2.18a}$$
$$\mathbf{a} \times \left(\mathbf{b} \times \mathbf{c}\right) = \left(\mathbf{a} \cdot \mathbf{c}\right)\mathbf{b} - \left(\mathbf{a} \cdot \mathbf{b}\right)\mathbf{c} \tag{2.18b}$$
$$\left(\mathbf{a} \times \mathbf{b}\right) \cdot \left(\mathbf{c} \times \mathbf{d}\right) = \left(\mathbf{a} \cdot \mathbf{c}\right)\left(\mathbf{b} \cdot \mathbf{d}\right) - \left(\mathbf{a} \cdot \mathbf{d}\right)\left(\mathbf{b} \cdot \mathbf{c}\right) \tag{2.18c}$$
$$\nabla \times \nabla\psi = 0 \tag{2.18d}$$
$$\nabla \cdot \left(\nabla \times \mathbf{a}\right) = 0 \tag{2.18e}$$
$$\nabla \times \left(\nabla \times \mathbf{a}\right) = \nabla\left(\nabla \cdot \mathbf{a}\right) - \nabla^2\mathbf{a} \tag{2.18f}$$
$$\nabla \cdot \left(\psi\mathbf{a}\right) = \mathbf{a} \cdot \nabla\psi + \psi\nabla \cdot \mathbf{a} \tag{2.18g}$$
$$\nabla \times \left(\psi\mathbf{a}\right) = \nabla\psi \times \mathbf{a} + \psi\nabla \times \mathbf{a} \tag{2.18h}$$
$$\nabla\left(\mathbf{a} \cdot \mathbf{b}\right) = \left(\mathbf{a} \cdot \nabla\right)\mathbf{b} + \left(\mathbf{b} \cdot \nabla\right)\mathbf{a} + \mathbf{a} \times \left(\nabla \times \mathbf{b}\right) + \mathbf{b} \times \left(\nabla \times \mathbf{a}\right) \tag{2.18i}$$
$$\nabla \cdot \left(\mathbf{a} \times \mathbf{b}\right) = \mathbf{b} \cdot \left(\nabla \times \mathbf{a}\right) - \mathbf{a} \cdot \left(\nabla \times \mathbf{b}\right) \tag{2.18j}$$
$$\nabla \times \left(\mathbf{a} \times \mathbf{b}\right) = \mathbf{a}\left(\nabla \cdot \mathbf{b}\right) - \mathbf{b}\left(\nabla \cdot \mathbf{a}\right) + \left(\mathbf{b} \cdot \nabla\right)\mathbf{a} - \left(\mathbf{a} \cdot \nabla\right)\mathbf{b} \tag{2.18k}$$
$$\left(\nabla\mathbf{b}\right) \cdot \mathbf{a} = \mathbf{a} \times \left(\nabla \times \mathbf{b}\right) + \left(\mathbf{a} \cdot \nabla\right)\mathbf{b} \tag{2.18l}$$
$$\nabla^2\mathbf{a} = \nabla\left(\nabla \cdot \mathbf{a}\right) - \nabla \times \left(\nabla \times \mathbf{a}\right) \tag{2.18m}$$
$$\tag{2.18n}$$

$$\nabla^2 \equiv \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial}{\partial u_1}\right) + \frac{\partial}{\partial u_2}\left(\frac{h_1 h_3}{h_2}\frac{\partial}{\partial u_2}\right) + \frac{\partial}{\partial u_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial}{\partial u_3}\right)\right] \tag{2.19}$$

Table 1: Scale factors of the Laplacian

Coord.	u_1	u_2	u_3	h_1	h_2	h_3
Cartesian	x	y	z	1	1	1
Cylindrical	r	ϕ	z	1	r	1
Spherical	z	θ	ϕ	1	r	r$\sin\phi$
Oblate Sph.	ξ	η	ϕ	a$\sqrt{\sinh^2\xi + \sin^2\eta}$	a$\sqrt{\sinh^2\xi + \sin^2\eta}$	a$\cosh\xi\cos\eta$
Elliptic Cyl.	u	ν	z	a$\sqrt{\sinh^2 u + \sin^2\nu}$	a$\sqrt{\sinh^2 u + \sin^2\nu}$	1

1. In the following equations, $dA \equiv$ unit surface area
2. $dV \equiv$ unit volume
3. $ds^2 \equiv 1^{st}$ Fundamental Form of a Line Element or Geodesic Equation of Free Motion
4. $h_i \equiv$ scale factors in the coordinate system metric
5. $g_{\mu\nu} \equiv$ coordinate system metric
6. $\Gamma^{\lambda}_{\mu\nu} \equiv$ Christoffel Symbol of the Second Kind

$$h_i \equiv \sqrt{g_{ii}} = \sqrt{\sum_{k=1}^{n} \left(\frac{\partial X_k}{\partial q_i} \right)^2} \tag{2.20}$$

$$g_{ij} = g_{ii}\delta_{ij} \text{ (Diagonal Metric)} \tag{2.21}$$

$$\begin{aligned} ds^2 &= g_{11} dx_1^2 + \ldots + g_{nn} dx_n^2 && \text{(2.22a)} \\ &= h_1^2 dx_1^2 + \ldots + h_n^2 dx_n^2 && \text{(2.22b)} \end{aligned}$$

$$\begin{aligned} \nabla^2 \phi &= g^{\mu\nu} \partial_\mu \partial_\nu \phi - \Gamma^\mu \partial_\nu \phi && \text{(2.23a)} \\ &= g^{\mu\nu} \frac{\partial}{\partial_\mu} \left(\frac{\partial \phi}{\partial_\nu} \right) - \Gamma^\mu \frac{\partial \phi}{\partial_\nu} && \text{(2.23b)} \end{aligned}$$

$$\begin{aligned} \Gamma^\lambda{}_{\mu\nu} &\equiv \frac{\partial^2 \zeta^\alpha}{\partial x^\mu \partial x^\nu} \frac{\partial x^\lambda}{\partial \zeta^\alpha} && \text{(2.24a)} \\ &= \frac{1}{2} g^{\lambda\alpha} \left[\partial_\nu g_{\alpha\mu} + \partial_\mu g_{\nu\alpha} - \partial_\alpha g_{\mu\nu} \right] && \text{(2.24b)} \\ &= g^{\alpha\lambda} \left[\mu\nu, \lambda \right] && \text{(2.24c)} \end{aligned}$$

$$\begin{aligned} \Gamma_{\lambda\mu\nu} &= 0 \text{ for } \lambda \neq \mu \neq \nu && \text{(2.25a)} \\ \Gamma_{\lambda\lambda\nu} &= -\frac{1}{2} \frac{\partial g_{\lambda\lambda}}{\partial x^\nu} \text{ for } \lambda \neq \nu && \text{(2.25b)} \\ \Gamma_{\lambda\mu\lambda} &= \Gamma_{\mu\lambda\lambda} = \frac{1}{2} \frac{\partial g_{\lambda\lambda}}{\partial x^\mu} && \text{(2.25c)} \\ \Gamma^\nu{}_{\lambda\mu} &= 0 \text{ for } \lambda \neq \mu \neq \nu && \text{(2.25d)} \\ \Gamma^\nu{}_{\lambda\lambda} &= -\frac{1}{2 g_{\nu\nu}} \frac{\partial g_{\lambda\lambda}}{\partial x^\nu} \text{ for } \lambda \neq \nu && \text{(2.25e)} \\ \Gamma^\lambda{}_{\lambda\mu} &= \Gamma^\lambda{}_{\mu\lambda} = \frac{1}{2 g_{\lambda\lambda}} \frac{\partial g_{\lambda\lambda}}{\partial x^\mu} = \frac{1}{2} \frac{\partial \ln g_{\lambda\lambda}}{\partial x^\mu} && \text{(2.25f)} \end{aligned}$$

$$d\tau^2 \equiv -g_{\mu\nu} dx^\mu dx^\nu \tag{2.26}$$

$$\nabla_\mu V^\nu \equiv \partial_\mu V^\nu + \Gamma^\nu{}_{\mu\lambda} V^\lambda \tag{2.27}$$

Let $\mathrm{x}^{\mu} = \mathrm{x}^{\mu}(\lambda)$ then:

$$\frac{d^2 x^{\mu}}{d\lambda^2} + \Gamma^{\mu}{}_{\rho\sigma} \frac{dx^{\rho}}{d\lambda} \frac{dx^{\sigma}}{d\lambda} = 0 \tag{2.28}$$

2.3 Mean, Variance, Covariance, and Correlation

We will use $\langle \; \rangle_{\alpha}$ to denote the *arithmetic mean* (or *expectation value* or *average*) with respect to the variable α (e.g., time or space). These angle brackets act like an operator and can be defined by:

$$\langle f(x) \rangle_{\alpha} = \frac{\int d\alpha \; f(x)}{\int d\alpha} \tag{2.29}$$

for a continuous function[2] or

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2.30}$$

for discrete variates, x_i. We will use μ_2, σ^2, or $var()$ to denote the *variance* and $cov(x, y)$ to denote the *covariance*. Finally, we will denote the *correlation* as $cor(x, y)$.

The μ_n notation denotes the n-th moment of some probability distribution, P(x), for some function, f(x). In general, the moments are called *raw moments* (μ_n), as they are not centered on any significant value of x. Herein, we will use *central moments* ($\bar{\mu}_n$), which are centered on the *mean*. The general form of the n-th central moment is:

$$\bar{\mu}_n \equiv \langle [f(x) - \langle f(x) \rangle]^n \rangle \tag{2.31a}$$
$$= \int dx \; [f(x) - \langle f(x) \rangle]^n \; P(x) \tag{2.31b}$$

where the integral is changed to a summation for discrete f(x). The various moments of a distribution are defined as:

$$\text{normalization} \equiv \mu_0 \tag{2.32a}$$
$$\equiv \text{ density for particle velocity distributions}$$
$$\text{mean} \equiv \mu_1 = \langle f(x) \rangle \tag{2.32b}$$
$$\equiv \text{ bulk flow velocity for particle velocity distributions}$$
$$\bar{\mu}_1 = \langle [f(x) - \langle f(x) \rangle] \rangle = \langle f(x) \rangle - \langle f(x) \rangle = 0 \tag{2.32c}$$
$$\text{variance} \equiv \bar{\mu}_2 = \left\langle [f(x) - \langle f(x) \rangle]^2 \right\rangle \tag{2.32d}$$
$$\equiv \text{ pressure tensor for particle velocity distributions}$$
$$\text{skewness} \equiv \frac{\bar{\mu}_3}{\bar{\mu}_2{}^{3/2}} = \text{ measure of asymmetry of a distribution} \tag{2.32e}$$
$$\equiv \text{ heat flux tensor for particle velocity distributions}$$
$$\text{kurtosis} \equiv \frac{\bar{\mu}_4}{\bar{\mu}_2{}^2} = \text{ degree of peakedness of a distribution} \tag{2.32f}$$

[2] in general, the denominator is not present and the $\langle \; \rangle_{\alpha}$ is an unnormalized average

2.3.1 Mean

The *mean* satisfies the following relations:

$$\langle ax + by \rangle = a\langle x \rangle + b\langle y \rangle \tag{2.33a}$$
$$\langle f(x) + g(x) \rangle = \langle f(x) \rangle + \langle g(x) \rangle \tag{2.33b}$$
$$\langle f(x) \cdot g(x) \rangle = \langle f(x) \rangle \cdot \langle g(x) \rangle \tag{2.33c}$$
$$\langle ax + b \rangle = a\langle x \rangle + b \tag{2.33d}$$
$$\langle \bar{x} \rangle = \left\langle \frac{1}{N} \sum_{i=1}^{N} x_i \right\rangle \tag{2.33e}$$
$$= \frac{1}{N} \left\langle \sum_{i=1}^{N} x_i \right\rangle \tag{2.33f}$$
$$= \frac{1}{N} \sum_{i=1}^{N} \langle x_i \rangle \tag{2.33g}$$
$$= \frac{1}{N} \sum_{i=1}^{N} \mu \tag{2.33h}$$
$$= \frac{1}{N} (N\mu) \tag{2.33i}$$
$$\equiv \mu \tag{2.33j}$$

where μ is the *population mean* or the first moment of the central moments defined by:

$$\bar{\mu}_n = \langle (x - \langle x \rangle)^n \rangle \ . \tag{2.34}$$

The *mean* of a bivariate function satisfies the following relationship:

$$\langle (x - \mu_x)(y - \mu_y) \rangle = \langle xy + \mu_x \mu_y - x\mu_y - y\mu_x \rangle \tag{2.35a}$$
$$= \langle xy \rangle + \langle \mu_x \mu_y \rangle - \langle x\mu_y \rangle - \langle y\mu_x \rangle \tag{2.35b}$$
$$= \langle xy \rangle + \mu_x \mu_y - \mu_y \langle x \rangle - \mu_x \langle y \rangle \tag{2.35c}$$
$$= \langle xy \rangle - \mu_x \mu_y \tag{2.35d}$$
$$= \langle xy \rangle - \langle x \rangle \langle y \rangle \ . \tag{2.35e}$$

2.3.2 Variance

The *variance* is given by:

$$var(x) = \left\langle (x - \langle x \rangle)^2 \right\rangle \tag{2.36a}$$
$$= \left\langle \left(x^2 + \langle x \rangle^2 - 2x \langle x \rangle \right) \right\rangle \tag{2.36b}$$
$$= \left\langle x^2 \right\rangle + \left\langle \langle x \rangle^2 \right\rangle - 2 \langle x \langle x \rangle \rangle \tag{2.36c}$$
$$= \left\langle x^2 \right\rangle - \langle x \rangle^2 \ . \tag{2.36d}$$

The *variance* satisfies the following relationship:

$$var\left(\bar{x}\right) = var\left\{\frac{1}{N}\sum_{i=1}^{N} x_i\right\} \tag{2.37a}$$

$$= \left\langle \left[\frac{1}{N}\sum_{i=1}^{N} x_i - \left\langle \frac{1}{N}\sum_{i=1}^{N} x_i \right\rangle\right]^2 \right\rangle \tag{2.37b}$$

$$= \frac{1}{N^2}\left\langle \left[\sum_{i=1}^{N} x_i - \left\langle \sum_{i=1}^{N} x_i \right\rangle\right]^2 \right\rangle \tag{2.37c}$$

$$= \frac{1}{N^2} var\left\{\sum_{i=1}^{N} x_i\right\} \tag{2.37d}$$

$$= \frac{1}{N^2}\sum_{i=1}^{N} var\left\{x_i\right\} \tag{2.37e}$$

$$= \frac{1}{N^2}\sum_{i=1}^{N} \left\langle \left[x_i - \left\langle x_i \right\rangle\right]^2 \right\rangle \tag{2.37f}$$

$$= \frac{1}{N^2}\sum_{i=1}^{N} \sigma^2 = \frac{\sigma^2}{N} \ . \tag{2.37g}$$

The *variance* also satisfies the following relationship:

$$var\left(ax + b\right) = \left\langle \left[\left(ax + b\right) - \left\langle ax + b \right\rangle\right]^2 \right\rangle \tag{2.38a}$$

$$= \left\langle \left[\left(ax\right) - a\left\langle x \right\rangle\right]^2 \right\rangle \tag{2.38b}$$

$$= \left\langle a^2\left[x - \left\langle x \right\rangle\right]^2 \right\rangle \tag{2.38c}$$

$$= a^2 var\left(x\right) \ . \tag{2.38d}$$

The *variance* of two random variates is given by:

$$var\left(x + y\right) = var\left(x\right) + var\left(y\right) + 2cov\left(x, y\right) \tag{2.39a}$$

$$= {\sigma_x}^2 + {\sigma_y}^2 + 2cov\left(x, y\right) \tag{2.39b}$$

$$var\left(y - bx\right) = var\left(y\right) + var\left(-bx\right) + 2cov\left(y, -bx\right) \tag{2.39c}$$

$$= var\left(y\right) + b^2 var\left(x\right) - 2bcov\left(y, x\right) \tag{2.39d}$$

$$= {\sigma_y}^2 + b^2{\sigma_x}^2 - 2b\sigma_x\sigma_y cor\left(y, x\right) \tag{2.39e}$$

where we have used $\sigma_x \equiv \sqrt{var\left(x\right)}$ (also known as the *standard deviation*), $cov\left(y, x\right) \equiv$ the *covariance* (see Section 2.3.3 for more), and $cor\left(y, x\right) \equiv$ the *correlation* (see Section 2.3.4 for more).

2.3.3 Covariance

The *covariance* is given by:

$$cov\left(x,y\right) \equiv \left\langle \left(x - \mu_x\right)\left(y - \mu_y\right) \right\rangle = \left\langle xy \right\rangle - \left\langle x \right\rangle \left\langle y \right\rangle \tag{2.40}$$

where we note that:

$$cov\left(x,x\right) = \left\langle x^2 \right\rangle - \left\langle x \right\rangle^2 = var\left(x\right) \ . \tag{2.41}$$

The *covariance* can also be related to the *variance* through:

$$var\left(\sum_{i=1}^{N} x_i\right) = cov\left(\sum_{i=1}^{N} x_i, \sum_{j=1}^{N} x_j\right) \tag{2.42a}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} cov\left(x_i, x_j\right) \tag{2.42b}$$

$$= \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j=i}}^{N} cov\left(x_i, x_j\right) + \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} cov\left(x_i, x_j\right) \tag{2.42c}$$

$$= \sum_{i=1}^{N} cov\left(x_i, x_i\right) + \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} cov\left(x_i, x_j\right) \tag{2.42d}$$

$$= \sum_{i=1}^{N} var\left(x_i\right) + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} cov\left(x_i, x_j\right) \ . \tag{2.42e}$$

The *covariance* has the following property for linear sums:

$$var\left(\sum_{i=1}^{N} a_i x_i\right) = cov\left(\sum_{i=1}^{N} a_i x_i, \sum_{j=1}^{N} a_j x_j\right) \tag{2.43a}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j cov\left(x_i, x_j\right) \tag{2.43b}$$

$$= \sum_{i=1}^{N} {a_i}^2 var\left(x_i\right) + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} a_i a_j cov\left(x_i, x_j\right) \ . \tag{2.43c}$$

We use the results for the *variance* of two random variates is given by Equations 2.39a – 2.39e, which allows us to calculate:

$$\begin{aligned}
cov\left(x+z,y\right) &= \langle [(x+z) - \langle x+z \rangle] \cdot [(y) - \langle y \rangle] \rangle && (2.44a)\\
&= \langle (x+z)\, y \rangle - \langle x+z \rangle \langle y \rangle && (2.44b)\\
&= \langle xy + zy \rangle - [\langle x \rangle + \langle z \rangle] \langle y \rangle && (2.44c)\\
&= \langle xy \rangle + \langle zy \rangle - \langle x \rangle \langle y \rangle - \langle z \rangle \langle y \rangle && (2.44d)\\
&= [\langle xy \rangle - \langle x \rangle \langle y \rangle] + [\langle zy \rangle - \langle z \rangle \langle y \rangle] && (2.44e)\\
&= cov\left(x,y\right) + cov\left(z,y\right) \ . && (2.44f)
\end{aligned}$$

Finally, the *covariance* satisfies the following:

$$\begin{aligned}
cov\left(x,a\right) &= 0 && (2.45a)\\
cov\left(ax,by\right) &= ab\ cov\left(x,y\right) && (2.45b)\\
cov\left(x+a,y+b\right) &= cov\left(x,y\right) && (2.45c)\\
cov\left(ax+by,cw+dz\right) &= ac\ cov\left(x,w\right) + ad\ cov\left(x,z\right) + bc\ cov\left(y,w\right) + bd\ cov\left(y,z\right) && (2.45d)
\end{aligned}$$

where the proofs are just a few lines of algebra following the rules for the *variance* and the *mean*. The *covariance matrix* is given by:

$$V_{ij} = cor\left(x_i, x_j\right) \equiv \langle (x_i - \mu_i) \cdot (x_j - \mu_j) \rangle \tag{2.46}$$

where an individual matrix element, V_{ij}, is called the covariance of x_i and x_j.

The *covariance*, in the form of Equation 2.40, is similar, physically, to the *pressure* in kinetic theory. More generally, the *covariance matrix* is analogous to the *pressure tensor* in kinetic theory[3]. Recall that the *pressure tensor* is symmetric, which is the result of the following:

$$\begin{aligned}
cov\left\{\sum_{i=1}^{N} x_i, y\right\} &= \sum_{i=1}^{N} cov\left\{x_i, y\right\} && (2.47a)\\
cov\left\{\sum_{i=1}^{N} x_i, \sum_{j=1}^{M} y_j\right\} &= \sum_{i=1}^{N} cov\left\{x_i, \sum_{j=1}^{M} y_j\right\} && (2.47b)\\
&= \sum_{i=1}^{N} cov\left\{\sum_{j=1}^{M} y_j, x_i\right\} && (2.47c)\\
&= \sum_{i=1}^{N} \sum_{j=1}^{M} cov\left\{y_j, x_i\right\} && (2.47d)\\
&= \sum_{i=1}^{N} \sum_{j=1}^{M} cov\left\{x_i, y_j\right\} && (2.47e)
\end{aligned}$$

[3]just replace x_i with a component of the particle momentum(velocity) and μ_i with the first moment, or bulk flow momentum(velocity)

which shows that $V_{ij} = V_{ji} \Rightarrow$ the *pressure tensor* is symmetric.

2.3.4 Correlation

The *correlation* is given by:

$$cor\left(x,y\right) \equiv \frac{cov\left(x,y\right)}{\sigma_x \sigma_y} \tag{2.48}$$

where σ_x is defined as $\sqrt{var\left(x\right)}$. We can use a rule for the *variance*, given by:

$$var\left\{\frac{x}{\sigma_x} \pm \frac{y}{\sigma_y}\right\} = var\left\{\frac{x}{\sigma_x}\right\} + var\left\{\frac{\pm y}{\sigma_y}\right\} + 2var\left\{\frac{x}{\sigma_x}, \frac{\pm y}{\sigma_y}\right\} \tag{2.49a}$$

$$= \frac{1}{{\sigma_x}^2} var\left(x\right) + \frac{1}{{\sigma_y}^2} var\left(y\right) \pm \frac{2}{\sigma_x \sigma_y} cov\left(x,y\right) \tag{2.49b}$$

combined with the knowledge that:

$$var\left\{\frac{x}{\sigma_x} \pm \frac{y}{\sigma_y}\right\} \geq 0 \tag{2.50}$$

to prove the following:

$$-1 \leq cor\left(x,y\right) \leq 1 \; . \tag{2.51}$$

2.4 Linear Algebra

Let $\langle \mathbf{x} \rangle$ be defined as the *sample mean*, which mathematically means:

$$\langle \mathbf{x} \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i \tag{2.52}$$

where N is the number of samples in your data set. Let us define the following:

$$\hat{\mathbf{x}}_k \equiv \mathbf{x}_k - \langle \mathbf{x} \rangle \tag{2.53}$$

which leads us to a matrix whose columns have a zero sample mean, defined as:

$$\mathbf{B} = \left[\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_2 \ldots \hat{\mathbf{x}}_n \right] . \tag{2.54}$$

The *sample covariance matrix* is thus defined by:

$$\mathbf{S} \equiv \frac{\mathbf{B}\mathbf{B}^T}{N-1} . \tag{2.55}$$

If we now define a vector, $\mathbf{X}$, which varies over the set of observed vectors and denote the coordinates by x_j, then the diagonal entry, s_{jj} in $\mathbf{S}$ is called the variance of x_j. Thus, s_{jj} measures the *spread* of the values of x_j. The *total variance* is defined as:

$$\left\{ TotalVariance \right\} \equiv Tr\left[\mathbf{S} \right] \tag{2.56}$$

The *covariance*, s_{ij} for i $\neq$ j, is equal to zero when x_i and x_j are uncorrelated.

2.5 Principle Component Analysis

The main goal here is to find an orthogonal $n \times n$ matrix, $\mathbf{P} = [\mathbf{u}_1 \ \ldots \ \mathbf{u}_n]$, such that $\mathbf{X} = \mathbf{P}\ \mathbf{Y}$, with the property that the components of $\mathbf{Y}$, y_j, are uncorrelated and arranged in order of decreasing variance. This implies that each individual observed vector, $\mathbf{X}_k$, goes to a new *name*, $\mathbf{Y}_k$. This results in the following relationship:

$$\mathbf{Y}_k = \mathbf{P}^{-1}\mathbf{X}_k = \mathbf{P}^T\mathbf{X}_k \text{ for k = 1,\ldots,N} . \tag{2.57}$$

A direct result of this *re-naming* is that the covariance matrix for $\mathbf{Y}_k$ is:

$$\mathbf{S}_2 = \mathbf{P}^T\mathbf{S}\mathbf{P} \tag{2.58}$$

which forces $\mathbf{S}_2$ to be diagonal (since $\mathbf{P}$ is an orthogonal matrix). Now if we allow $\mathbf{D}$ to be a diagonal matrix with eigenvalues of $\mathbf{S}$, λ_k on the diagonal arranged so that $\lambda_1 \geq \lambda_2 \geq \ldots \ \lambda_n \geq 0$, then if $\mathbf{P}$ is an orthogonal matrix of corresponding eigenvectors we have:

$$\mathbf{S} = \mathbf{P}\mathbf{D}\mathbf{P}^T \tag{2.59a}$$

$$\mathbf{D} = \mathbf{P}^T\mathbf{S}\mathbf{P} . \tag{2.59b}$$

The eigenvectors, $\mathbf{u}_i$, of the covariance matrix, $\mathbf{S}$, are called the *principal components* of the data. The *first principal component*, $\mathbf{u}_1$, is the eigenvector corresponding to the largest eigenvalue of $\mathbf{S}$ and the *second principal component*, $\mathbf{u}_2$, corresponds to the second largest eigenvalue and so on. If we allow c_i to be entries of $\mathbf{u}_1$, then $\mathbf{Y} = \mathbf{P}^T\ \mathbf{X}$ gives:

$$y_1 = \mathbf{u}_1^T\mathbf{X} = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \tag{2.60}$$

which means y_1 is a linear combination of the original variables $\mathrm{x}_1 \ \ldots \ \mathrm{x}_n$. One thing to note, the orthogonal change of variables, $\mathbf{X} = \mathbf{P}\ \mathbf{Y}$, does NOT change the total variance of the data, or in other words:

$$\Big\{Total\,Variance\,of\,x_i\Big\} = \Big\{Total\,Variance\,of\,y_i\Big\} \tag{2.61a}$$

$$\Big\{Total\,Variance\,of\,y_i\Big\} = Tr\Big[\mathbf{D}\Big] \tag{2.61b}$$

$$= \lambda_1 + \ldots + \lambda_n \tag{2.61c}$$

$\Rightarrow$ where the variance of y_i is λ_i, and $\lambda_i/\mathrm{Tr}\big[\mathbf{S}\big]$ measures the fraction of the total variace that is *explained* or *captured* by y_i. Thus if $\mathbf{u}$ satisfies, $y = \mathbf{u}^T\ \mathbf{X}$, then the variance of the values of y as $\mathbf{X}$ varies over the original data, $\mathbf{X}_i$, is $\mathbf{u}^T\mathbf{S}\mathbf{u}$.

1. The maximum value of $\mathbf{u}^T\mathbf{S}\mathbf{u}$ occurs for λ_1 and $\mathbf{u}_1$
2. y_2 has a maximum variance among all variables $y = \mathbf{u}^T\mathbf{X}$ that are uncorrelated with y_1

3. Likewise, y_3 has a maximum variance among all variables that are uncorrelated with BOTH y_1 and y_2.

2.6 Minimum Variance Analysis

Minimum variance analysis, or MVA, is the utilization of a property of plane polarized linear electromagnetic waves which allows one to assume that fluctuations in the electric ($\delta\mathbf{E}$) and magnetic ($\delta\mathbf{B}$) fields are are in a plane orthogonal to the direction of propagation ($\hat{\mathbf{k}}$) *Khrabrov and Sonnerup* [1998]. If the wave is truly a plane polarized wave, then $\hat{\mathbf{k}} \cdot \delta\mathbf{B} = 0$, which is a linear approximation of the Maxwell equation, $\nabla \cdot \mathbf{B} = 0$. The analysis is performed by minimizing the variance matrix of the magnetic field given by:

$$\mathbf{S}_{pq} = \left\langle \left(B_p - \langle B_p \rangle \right) \left(B_q - \langle B_q \rangle \right) \right\rangle \tag{2.62}$$

where $\langle B_p \rangle$ is the average of the p^{th} component of the magnetic field. We assume $\mathbf{S}_{pq}$ to be a non-degenerate matrix with three distinct eigenvalues, $\lambda_3 < \lambda_2 < \lambda_1$, and three corresponding eigenvectors, $\mathbf{e}_3$, $\mathbf{e}_2$, $\mathbf{e}_1$. Thus the minimum variance eigenvalue and eigenvector are λ_3 and $\mathbf{e}_3$. The propagation direction is said to be along $\hat{\mathbf{e}}_3$ if one assumes small isotropic noise and the condition $\lambda_2/\lambda_3 \geq 10$ is satisfied. Then the uncertainty in this direction is given by *Kawano and Higuchi* [1995]:

$$\delta\hat{\mathbf{k}} = \pm \left(\hat{\mathbf{e}}_1 \sqrt{\frac{\delta\lambda_3}{\lambda_1 - \lambda_3}} + \hat{\mathbf{e}}_2 \sqrt{\frac{\delta\lambda_3}{\lambda_2 - \lambda_3}} \right) \tag{2.63}$$

where K is the number vectors used and $\delta\lambda_3$, the uncertainty in the λ_3 eigenvalue, is given by:

$$\delta\lambda_3 = \pm \lambda_3 \sqrt{\frac{2}{(K-1)}} \; . \tag{2.64}$$

In general, the uncertainty of $\delta\lambda_i$ is given by:

$$\delta\lambda_i = \pm \sqrt{\frac{2\lambda_3(\lambda_i - \lambda_3)}{(K-1)}} \; . \tag{2.65}$$

Another useful quantity to know is the angle between the local ambient magnetic field and the propagation direction, θ_{kB}. This can be calculated in the typical manner, $\theta_{kB} \equiv \cos^{-1}\left(\hat{\mathbf{k}} \cdot \hat{\mathbf{b}}\right)$, with associated uncertainties of:

$$\delta\theta_{kB} = \pm \sqrt{\frac{\lambda_3 \lambda_2}{(K-1)(\lambda_2 - \lambda_3)^2}} \; . \tag{2.66}$$

Khrabrov and Sonnerup [1998] found analytical estimates to the error analysis of statistical noise in a vector field (i.e., B-field) with the application of minimum/maximum variance analysis. They consider two special cases of signal-to-noise ratios: 1) large and 2) small, for arbitrary noise distributions.

1. **The Ideal Case** $\equiv$ small errors and isotropic Gaussian noise

2. For the ideal case, one can determine *uncertainty cones* with elliptic cross sections for all three eigenvectors: $\mathbf{x}_1$,$\mathbf{x}_2$, $\mathbf{x}_3$, and *uncertainty intervals* for all three eigenvalues: λ_1, λ_2, λ_3
3. Note: $\lambda_3 < \lambda_2 < \lambda_1$ by definition
4. **Anisotropic Noise, No Signal:** 1) $\lambda_3 \approx \lambda_2 \equiv$ **Linearly Polarized IF** $\lambda_3 \ll \lambda_1$ **AND** the non-fluctuating part of the signal is negligible (i.e., only measuring noise due to wave packets which are broadband or spatially unresolved), 2) $\lambda_1 \approx \lambda_2 \equiv$ **Circularly Polarized** IF $\lambda_3 \ll \lambda_2$
5. **Small Anisotropic Noise:** If amplitude of noise $\ll$ amplitude of signal, then λ_3 can be said to be entirely due to noise
6. **Isotropic Gaussian Noise:** Equations 2.71 and 2.72 implicitly assume isotropic Gaussian noise
7. In Equation 2.85, $\Delta\phi_{i,j} \equiv$ the angular standard deviation (radians) of the i^{th} vector's ($\vec{x}_i$) direction towards/away from the j^{th} vector's ($\vec{x}_j$) direction
8. The *Variance* of any quantity is defined as in Equation 2.69. The use of Minimum Variance Analysis (MVA) on magnetic fields derives from the Maxwell Equation $\nabla \cdot \mathbf{B} = 0$. From this equation, one can convert the divergence into a dot product between a vector, $\mathbf{n}$, and the B-field. If this vector $\mathbf{n}$ exists, the field does not vary along it. Thus we say, $B_n = \mathbf{n} \cdot \mathbf{B} =$ constant! So we vary the B-field in each of it's component directions and the variance is described by Equation 2.73, where K $\equiv$ number of measurements/vectors.
9. *Rule of Thumb:* for K < 50, REQUIRE $\lambda_2/\lambda_3 \geq 10$, UNLESS one knows *a priori* that the noise is truly random, which then implies that 1/K is a relevant, small parameter
10. The Variance Matrix: see Equation 2.74
11. **Ensemble Average** $\equiv \langle\langle\ \rangle\rangle \equiv$ average over the ensemble of all realizations of data
12. **Average of Data** $\equiv \langle\ \rangle \equiv$ average of data in a given realization

If we have a set of functions given by:

$$\Big\{f_j\Big\} = \Big\{f\left(x_j\right)\Big\} \tag{2.67}$$

and we let x_j go to $\langle x \rangle + e_j$, then we have:

$$\langle f \rangle = \frac{1}{N}\sum_j f\left(x_j\right) \tag{2.68a}$$

$$= \frac{1}{N}\sum_j f\left(\langle x \rangle + e_j\right) \tag{2.68b}$$

$$= f\left(\langle x \rangle\right) + \frac{1}{N} f'\left(\langle x \rangle\right)\sum_j e_j + \frac{1}{2N} f''\left(\langle x \rangle\right)\sum_j \left(e_j\right)^2 + \dots \tag{2.68c}$$

$$= f\left(\langle x \rangle\right) + \frac{\sigma^2}{2} f''\left(\langle x \rangle\right) \tag{2.68d}$$

where σ^2 is defined by:

$$\sigma^2 \equiv \frac{1}{N} \sum_{i=1}^{N} \left(\mathbf{x}_i - \langle \mathbf{x} \rangle \right)^2 \tag{2.69}$$

thus, it can be shown that the fluctuations of two eigenvalues, treated as distinct and uncorrelated, have an the standard deviation of their difference, (λ_i - λ_j), as:

$$\sigma_{ij} = \sqrt{\left\langle \left\langle \left(\Delta\lambda_i \right)^2 \right\rangle \right\rangle + \left\langle \left\langle \left(\Delta\lambda_j \right)^2 \right\rangle \right\rangle} \tag{2.70}$$

where, we have

$$\left\langle \left\langle \left(\Delta\lambda_i \right)^2 \right\rangle \right\rangle = \frac{2\lambda_3 \left(2\lambda_i - \lambda_3 \right)}{\left(K - 1 \right)} \; . \tag{2.71}$$

The uncertainty in the vector, $\mathbf{x}_1$ (The maximum variance direction.), is then given by:

$$\Delta\phi_{1j} = \sqrt{\frac{\lambda_j \lambda_1}{\left(K - 1 \right) \left(\lambda_1 - \lambda_j \right)}} \tag{2.72}$$

if $\lambda_2 \ll \lambda_1$ **AND** $\lambda_3 \ll \lambda_1$.

$$Var(\mathbf{B} \cdot \mathbf{x}) \equiv \frac{1}{K} \sum_{k=1}^{K} \left[\left(\mathbf{B}^{(k)} - \langle \mathbf{B} \rangle \right) \cdot \mathbf{x} \right]^2 \equiv \left\langle \left[\left(\mathbf{B}^{(k)} - \langle \mathbf{B} \rangle \right) \cdot \mathbf{x} \right]^2 \right\rangle \tag{2.73}$$

$$M_{ij} = \left\langle \left(B_i^{(k)} - \langle B_i^{(k)} \rangle \right) \left(B_j^{(k)} - \langle B_j^{(k)} \rangle \right) \right\rangle \equiv \left\langle \delta B_i^{(k)} \delta B_j^{(k)} \right\rangle \tag{2.74}$$

now replace $\mathbf{B}^{(k)}$ by $\mathbf{B}^{*(k)} + \delta\mathbf{b}^{(k)}$, where $\mathbf{B}^{*(k)} \equiv$ signal and $\delta\mathbf{b}^{(k)} \equiv$ noise. One should note that $\mathbf{B}^{*(k)}$, k = 1, 2, . . . , K are the same in all realizations, while the K-offset noise components, $\delta\mathbf{b}^{(k)}$, contain $\langle \mathbf{b} \rangle$, therefore are functions of all K noise vectors, $\mathbf{b}^{(k)}$, k = 1, 2, . . . ,K, in the realization. By definition, the latter has the property:

$$\left\langle \left\langle \mathbf{b}^{(k)} \right\rangle \right\rangle \equiv 0 \Rightarrow \left\langle \left\langle \delta\mathbf{b}^{(k)} \right\rangle \right\rangle \equiv 0 \tag{2.75}$$

which allows us to define the following:

$$\delta\mathbf{B}^{(k)} = \delta\mathbf{B}^{*(k)} + \delta\mathbf{b}^{(k)} \tag{2.76}$$

where

$$\delta \mathbf{B}^{*(k)} \equiv \mathbf{B}^{*(k)} - \left\langle \mathbf{B}^{*} \right\rangle \tag{2.77a}$$
$$\delta \mathbf{b}^{(k)} \equiv \mathbf{b}^{(k)} - \left\langle \mathbf{b} \right\rangle \tag{2.77b}$$

so that Equation 2.74 goes to:

$$\left\langle \delta B_i^{(k)} \delta B_j^{(k)} \right\rangle = \left\langle \left(\delta B_i^{*(k)} + \delta b_i^{(k)} \right) \left(\delta B_j^{*(k)} + \delta b_j^{(k)} \right) \right\rangle \tag{2.78a}$$
$$= \left\langle \delta B_i^{*(k)} \left(\delta B_j^{*(k)} + \delta b_j^{(k)} \right) \right\rangle + \left\langle \delta b_i^{(k)} \left(\delta B_j^{*(k)} + \delta b_j^{(k)} \right) \right\rangle \tag{2.78b}$$
$$= \left\langle \delta B_i^{*(k)} \delta B_j^{*(k)} \right\rangle + \left\langle \delta B_i^{*(k)} \delta b_j^{(k)} \right\rangle + \left\langle \delta b_i^{(k)} \delta B_j^{*(k)} \right\rangle + \left\langle \delta b_i^{(k)} \delta b_j^{(k)} \right\rangle = M_{ij} \ . \tag{2.78c}$$

For the next step we have to realize that the following rule is valid:

$$\left\langle \left[\langle \langle A \rangle \rangle \right] \right\rangle = \left\langle \left\langle \left[\langle A \rangle \right] \right\rangle \right\rangle \ . \tag{2.79}$$

If we take the *ensemble average* of our variance matrix, we get:

$$\left\langle \left\langle \left(\Delta M_{ij} \right)^2 \right\rangle \right\rangle = \left\langle \left\langle \left[M_{ij} + \langle \langle M_{ij} \rangle \rangle \right]^2 \right\rangle \right\rangle \tag{2.80a}$$
$$= \left\langle \left\langle \left\{ \frac{1}{K} \sum_k \left(\delta B_i^{*(k)} + \delta b_i^{(k)} \right) \left(\delta B_j^{*(k)} + \delta b_j^{(k)} \right) - \frac{1}{K} \sum_m \left(\delta B_i^{*(m)} \delta B_j^{*(m)} + \langle \langle \delta b_i \delta b_j \rangle \rangle \right) \right\}^2 \right\rangle \right\rangle \tag{2.80b}$$

where the second term on the R.H.S. of Equation 2.80a is:

$$\left\langle \left\langle M_{ij} \right\rangle \right\rangle = \left\langle \left\langle \left[\left\langle \delta B_i^{(k)} \delta B_j^{(k)} \right\rangle \right] \right\rangle \right\rangle \tag{2.81a}$$
$$= \left\langle \left[\left\langle \left\langle \delta B_i^{(k)} \delta B_j^{(k)} \right\rangle \right\rangle \right] \right\rangle \tag{2.81b}$$
$$= \left\langle \left[\left\langle \left\langle \delta B_i^{*(k)} \delta B_j^{*(k)} \right\rangle \right\rangle \right] \right\rangle + \left\langle \left[\left\langle \left\langle \delta B_i^{*(k)} \delta b_j^{(k)} \right\rangle \right\rangle \right] \right\rangle + \left\langle \left[\left\langle \left\langle \delta b_i^{(k)} \delta B_j^{*(k)} \right\rangle \right\rangle \right] \right\rangle + \left\langle \left[\left\langle \left\langle \delta b_i^{(k)} \delta b_j^{(k)} \right\rangle \right\rangle \right] \right\rangle \tag{2.81c}$$

which is highly simplified by realizing that the middle two terms can be canceled when the ensemble average is taken due to the properties assumed in Equation 2.75. The first term on the R.H.S. is just defined as:

$$M^*_{ij} \equiv \left\langle \left\{ \left\langle \left\langle \delta B_i^{*(k)} \delta B_j^{*(k)} \right\rangle \right\rangle \right\} \right\rangle \tag{2.82}$$

which is the variance matrix of the nonfluctuating part of the field. The final result is written as:

$$\left\langle \left\langle M_{ij} \right\rangle \right\rangle = M^*_{ij} + \left\langle \left\langle \left\{ \left\langle \delta b_i^{(k)} \delta b_j^{(k)} \right\rangle \right\} \right\rangle \right\rangle \ . \tag{2.83}$$

The second term on the R.H.S. of Equation 2.83 can be dealt with in the following manner:

$$\left\langle \left\langle \left\{ \left\langle \delta b_i^{(k)} \delta b_j^{(k)} \right\rangle \right\} \right\rangle \right\rangle = \left\langle \left\langle \left\{ \left\langle \left(b_i^{(k)} - \langle b_i^{(k)} \rangle \right) \left(b_j^{(k)} - \langle b_j^{(k)} \rangle \right) \right\rangle \right\} \right\rangle \right\rangle \tag{2.84a}$$

$$= \left\langle \left\langle \left\{ \left\langle b_i^{(k)} b_j^{(k)} \right\rangle - \left\langle b_i^{(k)} \langle b_j^{(k)} \rangle \right\rangle - \left\langle \langle b_i^{(k)} \rangle b_j^{(k)} \right\rangle + \left\langle \langle b_i^{(k)} \rangle \langle b_j^{(k)} \rangle \right\rangle \right\} \right\rangle \right\rangle \tag{2.84b}$$

$$= \left\langle \left\langle \left\{ \left\langle b_i^{(k)} b_j^{(k)} \right\rangle - \langle b_i^{(k)} \rangle \langle b_j^{(k)} \rangle \right\} \right\rangle \right\rangle \tag{2.84c}$$

$$= \left\langle \left\{ \left\langle \left\langle b_i^{(k)} b_j^{(k)} \right\rangle \right\rangle \right\} \right\rangle - \left\langle \left\langle \left\{ \langle b_i^{(k)} \rangle \langle b_j^{(k)} \rangle \right\} \right\rangle \right\rangle \tag{2.84d}$$

The uncertainty in the direction between any two eigenvectors is given by:

$$\Delta \phi_{i,j} = \pm \sqrt{\left(\frac{\lambda_3 (\lambda_i + \lambda_j - \lambda_3)}{(K-1)(\lambda_i - \lambda_j)^2} \right)} \tag{2.85}$$

with an uncertainty in eigenvalues given by:

$$\Delta \lambda_i = \pm \sqrt{\left(\frac{2\lambda_3 (2\lambda_i - \lambda_3)}{(K-1)} \right)} \tag{2.86}$$

2.7 Trigonometric Identities

The following are trigonometric identities for complex functions of x:

$$sinh\ (\pm ix) = \pm i\ sin\ x \tag{2.87a}$$
$$cosh\ (\pm ix) = cos\ x \tag{2.87b}$$
$$tanh\ (\pm ix) = \pm i\ tan\ x \tag{2.87c}$$
$$sinh\ (y \pm ix) = \pm i\ cosh\,(y)\ sin\,(x) + cos\,(x)\ sinh\,(y) \tag{2.87d}$$
$$cosh\ (y \pm ix) = cos\,(x)\ cosh\,(y) \pm i\ sin\,(x)\ sinh\,(y) \tag{2.87e}$$
$$tanh\ (y \pm ix) = \frac{\pm i\ cosh\,(y)\ sin\,(x) + cos\,(x)\ sinh\,(y)}{cos\,(x)\ cosh\,(y) \pm i\ sin\,(x)\ sinh\,(y)} \tag{2.87f}$$

where we can note that letting $ix \rightarrow z$ gives:

$$sinh\ (y \pm z) = cosh\,(z)\ sinh\,(y) \pm cosh\,(y)\ sinh\,(z) \tag{2.88a}$$
$$cosh\ (y \pm z) = cosh\,(y)\ cosh\,(z) \pm sinh\,(y)\ sinh\,(z) \tag{2.88b}$$
$$tanh\ (y \pm z) = \frac{cosh\,(z)\ sinh\,(y) \pm cosh\,(y)\ sinh\,(z)}{cosh\,(y)\ cosh\,(z) \pm sinh\,(y)\ sinh\,(z)}\ . \tag{2.88c}$$

2.8 Taylor Series

The following are Taylor series expansions for general functions of x:

$$\sqrt{\frac{1}{1+x^2}} \approx 1 - \frac{x^2}{2} + \frac{3x^4}{8} + \mathcal{O}\left(x^6\right) \tag{2.89a}$$

$$\sqrt{\frac{1}{1+(x/a)^2}} \approx 1 - \frac{x^2}{2a^2} + \frac{3x^4}{8a^4} + \mathcal{O}\left(x^6\right) \tag{2.89b}$$

$$\sqrt{1+x^2} \approx 1 + \frac{x^2}{2} - \frac{x^4}{8} + \mathcal{O}\left(x^6\right) \tag{2.89c}$$

$$\sqrt{1+(x/a)^2} \approx 1 - \frac{x^2}{2a^2} + \frac{x^4}{8a^4} + \mathcal{O}\left(x^6\right) \tag{2.89d}$$

$$\sqrt{\frac{x^2}{1+x^2}} \approx x - \frac{x^3}{2} + \frac{3x^5}{8} + \mathcal{O}\left(x^7\right) \tag{2.89e}$$

$$\sqrt{\frac{1+x^2}{x^2}} \approx \frac{1}{x} + \frac{x}{2} - \frac{x^3}{8} + \mathcal{O}\left(x^5\right) \tag{2.89f}$$

$$\frac{1}{1+x^2} \approx 1 - x^2 + x^4 + \mathcal{O}\left(x^6\right) \tag{2.89g}$$

$$\frac{1}{1+(x/a)^2} \approx 1 - \frac{x^2}{a^2} + \frac{x^4}{a^4} + \mathcal{O}\left(x^6\right) \tag{2.89h}$$

The following are Taylor series expansions for exponential functions of x:

$$e^{\pm x} \approx 1 \pm x + \frac{x^2}{2!} \pm \frac{x^3}{3!} + \frac{x^4}{4!} \pm \frac{x^5}{5!} + \mathcal{O}\left(x^6\right) \tag{2.90a}$$

$$e^{\pm ix} \approx 1 \pm (ix) - \frac{x^2}{2!} \mp \frac{ix^3}{3!} + \frac{x^4}{4!} \pm \frac{ix^5}{5!} + \mathcal{O}\left(x^6\right) \tag{2.90b}$$

$$e^{\pm x^2} \approx 1 \pm x^2 + \frac{x^4}{2!} \pm \frac{x^6}{3!} + \frac{x^8}{4!} + \mathcal{O}\left(x^{10}\right) \tag{2.90c}$$

where we know that:

$$e^{\pm ix} = \cos x \pm i \sin x \; . \tag{2.91}$$

The following are Taylor series expansions for trigonometric functions of x:

$$\sin x \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \mathcal{O}\left(x^9\right) \tag{2.92a}$$

$$\cos x \approx 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \mathcal{O}\left(x^8\right) \tag{2.92b}$$

$$\tan x \approx x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \mathcal{O}\left(x^9\right) \tag{2.92c}$$

$$\sec x \approx 1 + \frac{x^2}{2!} + \frac{5x^4}{4!} + \frac{61x^6}{6!} + \mathcal{O}\left(x^8\right) \tag{2.92d}$$

$$\csc x \approx \frac{1}{x} + \frac{x}{6} + \frac{7x^3}{360} + \frac{31x^5}{15120} + \mathcal{O}\left(x^7\right) \tag{2.92e}$$

$$\cot x \approx \frac{1}{x} - \frac{x}{3} - \frac{x^3}{45} - \frac{2x^5}{945} + \mathcal{O}\left(x^7\right) \tag{2.92f}$$

The following are Taylor series expansions for hyperbolic functions of x:

$$sinh\ x \approx x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \mathcal{O}\left(x^9\right) \tag{2.93a}$$

$$cosh\ x \approx 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \mathcal{O}\left(x^8\right) \tag{2.93b}$$

$$tanh\ x \approx x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} + \mathcal{O}\left(x^9\right) \tag{2.93c}$$

$$sech\ x \approx 1 - \frac{x^2}{2!} + \frac{5x^4}{4!} - \frac{61x^6}{6!} + \mathcal{O}\left(x^8\right) \tag{2.93d}$$

$$csch\ x \approx \frac{1}{x} - \frac{x}{6} + \frac{7x^3}{360} - \frac{31x^5}{15120} + \mathcal{O}\left(x^7\right) \tag{2.93e}$$

$$coth\ x \approx \frac{1}{x} + \frac{x}{3} - \frac{x^3}{45} + \frac{2x^5}{945} + \mathcal{O}\left(x^7\right) \tag{2.93f}$$

2.9 Variational Principle

Recall that for an arbitrary function, $\mathcal{F} = \mathcal{F}\left(t, x_1, x_2, ..., x_{n-1}, x_n\right)$, the *exact derivative* or *total derivative* is given by:

$$\frac{d\mathcal{F}}{dt} = \frac{\partial \mathcal{F}}{\partial t} + \sum_{i=1}^{n} \frac{\partial \mathcal{F}}{\partial x_i} \frac{dx_i}{dt} \ . \tag{2.94}$$

If we have $\mathcal{W} = \mathcal{W}\left(\boldsymbol{\kappa}, \omega, \mathbf{x}, t\right)$, then the variation is given by:

$$\delta\mathcal{W}\left(\boldsymbol{\kappa}, \omega, \mathbf{x}, t\right) = \frac{\partial \mathcal{W}}{\partial t}\delta t + \frac{\partial \mathcal{W}}{\partial \mathbf{x}} \cdot \delta\mathbf{x} + \frac{\partial \mathcal{W}}{\partial \omega}\delta\omega + \frac{\partial \mathcal{W}}{\partial \boldsymbol{\kappa}} \cdot \delta\boldsymbol{\kappa} \tag{2.95}$$

Now let us consider the dispersion relation, $\omega = \mathcal{W}\left(\boldsymbol{\kappa}, \mathbf{x}, t\right)$, then variation can be shown to be:

$$\delta\mathcal{W}\left(\boldsymbol{\kappa}, \mathbf{x}, t\right) = \frac{\partial \mathcal{W}}{\partial t}\delta t + \frac{\partial \mathcal{W}}{\partial \mathbf{x}} \cdot \delta\mathbf{x} + \frac{\partial \mathcal{W}}{\partial \boldsymbol{\kappa}} \cdot \delta\boldsymbol{\kappa} \tag{2.96}$$

2.10 Bessel Functions

If we define J_n and Y_n as Bessel functions of the first and second kind, respectively, and we let I_n and K_n be the modified Bessel functions of the first and second kind, respectively, then:

$$J_n(x) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j(n+j)} \left(\frac{x}{2}\right)^{2j+n} \tag{2.97a}$$

$$Y_n(x) = \frac{J_n(x)\cos(n\pi) - J_{-n}(x)}{\sin(n\pi)} \tag{2.97b}$$

$$I_n(x) = (i)^{-n} J_n(ix) \tag{2.97c}$$

$$= e^{-in\pi/2} J_n\left(xe^{i\pi/2}\right) \tag{2.97d}$$

$$= \sum_{m=0}^{\infty} \frac{1}{m!\,(m+\mid n\mid)!} \left(\frac{x}{2}\right)^{2m+|n|} \tag{2.97e}$$

$$K_n(x) = \frac{\pi}{2} \frac{I_{-n}(x) - I_n(x)}{\sin(n\pi)} \tag{2.97f}$$

Some Bessel function relationships are:

$$\int_0^{\infty} du\ J_n{}^2(au)\ u\ e^{-\beta u^2} = \frac{1}{2\beta} e^{-a^2/2\beta} I_n\left(\frac{a^2}{2\beta}\right) \tag{2.98a}$$

$$e^{i[\alpha_s \phi + \beta_s \sin\phi]} = \sum_{m=0}^{\infty} J_m(\beta_s)\, e^{i[\alpha_s + m]\phi} \tag{2.98b}$$

$$J_{n+1}(x) + J_{n-1}(x) = \frac{2n}{x} J_n(x) \tag{2.98c}$$

$$J_{n+1}(x) - J_{n-1}(x) = -2\frac{dJ_n(x)}{dx} \tag{2.98d}$$

$$\int_{\phi} d\phi'\ e^{-i[\alpha_s \phi' + \beta_s \sin\phi']} = \sum_{n=0}^{\infty} J_n(\beta_s) \int_{\phi} d\phi'\ e^{-i[\alpha_s + n]\phi'} \tag{2.98e}$$

$$= i \sum_{n=0}^{\infty} \frac{J_n(\beta_s)}{\alpha_s + n} e^{-i[\alpha_s + n]\phi'} \tag{2.98f}$$

$$e^{ix\sin\theta} = \sum_{n=-\infty}^{\infty} J_n(x)\, e^{in\theta} \tag{2.98g}$$

and we can also note that:

$$\int_0^{2\pi} d\phi e^{i(m-n)\phi} = 2\pi\delta_{m,n} \tag{2.99a}$$

$$\lim_{x\to 0} nJ_n{}^2(x) = 0 \tag{2.99b}$$

Bessel function identities:

$$\sum_{n=-\infty}^{\infty} J_n{}^2(x) = 1 \tag{2.100a}$$

$$\sin\phi \sum_n J_n(k_\perp r)\ e^{in\phi} = -i \sum_n J'_n(k_\perp r)\ e^{in\phi} \tag{2.100b}$$

$$\cos\phi \sum_n J_n(k_\perp r)\ e^{in\phi} = \sum_n \frac{n\Omega}{k_\perp V_\perp} J_n(k_\perp r)\ e^{in\phi} \tag{2.100c}$$

$$\sum_n J_n(x)\ J'_n(x) = 0 \tag{2.100d}$$

$$\sum_n n^2\ J_n{}^2(x) = \frac{x^2}{2} \tag{2.100e}$$

2.11 The Plasma Dispersion Function

In this section, we will discuss the plasma dispersion function [e.g., *Gurnett and Bhattacharjee*, 2005].
If we let $\zeta_s = \sqrt{m_s/(2k_B T_s)}$ (i p/k), then we define:

$$Z(\zeta) = \frac{1}{\sqrt{\pi}} \int_C dz \, \frac{e^{-z^2}}{z - \zeta} \tag{2.101}$$

where the contour C is understood to be along the real z-axis, passing under the pole at z = ζ. To alter this, we need to consider the Plemelj relation given by:

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} dx \, \frac{f(x)}{x - (x_o \pm i\varepsilon)} = P \int_{-\infty}^{\infty} dx \, \frac{f(x)}{x - x_o} \pm i\pi f(x_o) \tag{2.102}$$

where $\epsilon > 0$ and the P refers to the principal value integral defined by:

$$P \int_{-\infty}^{\infty} ...dx = \lim_{\delta \to 0} [\int_{-\infty}^{x_o - \delta} ...dx + \int_{x_o + \delta}^{\infty} ...dx]. \tag{2.103}$$

We can define the derivative of the plasma dispersion function as:

$$\frac{dZ}{d\zeta} \equiv Z'(\zeta) = -2\left[1 + \zeta Z(\zeta)\right] \tag{2.104}$$

We can expand $Z(\zeta)$ and $Z'(\zeta)$ in the limits $|\zeta| \gg 1$ and $|\zeta| \ll 1$, which are given by:

$$Z(\zeta) = i\sqrt{\pi}\frac{k}{\mid k \mid} e^{-\zeta^2} - \left[\frac{1}{\zeta} + \frac{1}{2\zeta^3} + \frac{3}{4\zeta^5} + \cdots\right] \text{ (for } |\zeta| \gg 1) \tag{2.105a}$$

$$Z(\zeta) = i\sqrt{\pi}\frac{k}{\mid k \mid} e^{-\zeta^2} - \left[2\zeta - \frac{4}{3}\zeta^3 + \frac{8}{15}\zeta^5 + \cdots\right] \text{ (for } |\zeta| \ll 1) \tag{2.105b}$$

$$Z'(\zeta) = -2i\sqrt{\pi}\frac{k}{\mid k \mid} \zeta e^{-\zeta^2} + \left[\frac{1}{\zeta^2} + \frac{3}{2\zeta^4} + \frac{15}{4\zeta^6} + \cdots\right] \text{ (for } |\zeta| \gg 1) \tag{2.105c}$$

$$Z'(\zeta) = -2i\sqrt{\pi}\frac{k}{\mid k \mid} \zeta e^{-\zeta^2} - \left[2 - 4\zeta^2 + \frac{8}{3}\zeta^4 + \cdots\right] \text{ (for } |\zeta| \ll 1) \tag{2.105d}$$

2.12 Deriving the Quadratic Equation

2.12.1 Basic Algebra

So before we get too ahead of ourselves, let us review a few important properties of algebraic manipulation. The first is multiplication by one, which can be seen as:

$$a = a\left(\frac{b}{b}\right) \tag{2.106a}$$
$$\frac{1}{(x-y)} = \frac{1}{\frac{x}{x}(x-y)} \tag{2.106b}$$
$$= \frac{1}{x(1-\frac{y}{x})} \tag{2.106c}$$
$$\frac{1}{(x-y)^n} = \frac{1}{(\frac{x}{x})^n(x-y)^n} \tag{2.106d}$$
$$= \frac{1}{x^n(1-\frac{y}{x})^n} \tag{2.106e}$$

and so on and so forth. The point is, there are a multitude of ways to multiply any factor by the number one. Note that in the above examples, I am simply doing everything with arbitrary variables and for the anal retentive mathematicians, we'll assume that **ALL** of those variables are definite real numbers not equal to zero.

Now let's review a few properties of quotients. The two most important ones to remember, at least in my mind, are the following:

$$\left(\frac{a}{\frac{b}{c}}\right) = \left(\frac{ac}{b}\right) \tag{2.107a}$$
$$\left(\frac{\frac{a}{b}}{c}\right) = \left(\frac{a}{bc}\right) \tag{2.107b}$$

2.12.2 Quadratic Equation

If we let the following variables be defined as constants, a, b, c, where $a \neq 0$ and b and c are $\in \Re$[4]. Thus we start with a general second-order[5] polynomial equation of the form:

$$ax^2 + bx + c = 0 \tag{2.108}$$

where our undetermined variable, x, is the unknown we seek to solve for. Now when one is faced with a general second-order polynomial that cannot be factored, it is typically useful to do something called

[4]Note that $\in$ is one of the fancy mathematician ways of saying *element of...* while $\Re$ pertains to the set of numbers known as *Reals*. Also, I should be careful to point out that the ONLY real requirement on any of the constants is that $a \neq 0$, but we'll throw in the real number thing to avoid imaginaries, which tend to obfuscate things.

[5]the highest power of our undetermined variable, x, is 2

completing the square. To do so, we first divide both sides of Equation 2.108 by a[6]:

$$ax^2 + bx + c = 0 \tag{2.109a}$$

$$x^2 + \left(\frac{b}{a}\right) x + \left(\frac{c}{a}\right) = 0 \tag{2.109b}$$

Now the next step in completing the square requires that we move **ALL** terms with **ONLY** constants in them to the opposite side of the equation from that of our unknown variable, x. This changes Equation 2.109b to:

$$x^2 + \left(\frac{b}{a}\right) x = -\left(\frac{c}{a}\right) \tag{2.110}$$

and now are trying to change our original equation to something of the form of:

$$(x + B)^2 = C \tag{2.111}$$

where B and C are **NOT** the same as their lower case counterparts. To see this, we expand the left hand side of Equation 2.111 to find:

$$(x + B)^2 = x^2 + 2Bx + B^2 \tag{2.112}$$

where we see that there is a factor of 2 which must be taken into account. So let's expand the following:

$$\left(x + \left(\frac{b}{a}\right)\right)^2 = x^2 + 2\left(\frac{b}{a}\right) x + \left(\frac{b}{a}\right)^2 \tag{2.113}$$

which still leaves that pesky factor of 2 in our equation, so let's try a different approach. Instead of Equation 2.113, let's try the following:

$$\left(x + \left(\frac{b}{2a}\right)\right)^2 = \left[x^2 + \left(\frac{b}{a}\right) x\right] + \left(\frac{b}{2a}\right)^2 \tag{2.114}$$

where we can see the the terms in [ ] are the same as those on the left hand side of Equation 2.110.

[6]Were a allowed even the slightest possibility to be $= 0$, mathematicians would have their panties all in a bunch over this step...

Thus, we rearrange Equation 2.114 in the following manner to find:

$$\left\{x + \left(\frac{b}{2a}\right)\right\}^2 - \left(\frac{b}{2a}\right)^2 = \left[x^2 + \left(\frac{b}{a}\right)x\right] \tag{2.115a}$$

$$= -\left(\frac{c}{a}\right) \tag{2.115b}$$

Thus we have the following equation of the following form:

$$\left\{x + \left(\frac{b}{2a}\right)\right\}^2 - \left[\left(\frac{b}{2a}\right)^2 - \left(\frac{c}{a}\right)\right] = 0 \tag{2.116}$$

which allows us to see that B and C in Equation 2.111 are:

$$B \equiv \left(\frac{b}{2a}\right) \tag{2.117a}$$

$$C \equiv \left[\left(\frac{b}{2a}\right)^2 - \left(\frac{c}{a}\right)\right] \tag{2.117b}$$

Now we return to Equation 2.111 and solve for x by first taking the square-root of both sides finding:

$$(x + B) = \pm\sqrt{C} \tag{2.118a}$$

$$x = -B \pm \sqrt{C} \tag{2.118b}$$

and now substitute in our definitions of B and C from Equations 2.117a and 2.117b to find:

$$x = -\left(\frac{b}{2a}\right) \pm \sqrt{\left[\left(\frac{b}{2a}\right)^2 - \left(\frac{c}{a}\right)\right]} \tag{2.119a}$$

$$= -\left(\frac{b}{2a}\right) \pm \sqrt{\left(\frac{2a}{2a}\right)^2 \left[\left(\frac{b}{2a}\right)^2 - \left(\frac{c}{a}\right)\right]} \tag{2.119b}$$

$$= -\left(\frac{b}{2a}\right) \pm \left(\frac{1}{2a}\right)\sqrt{b^2 - \left(\frac{4a^2c}{a}\right)} \tag{2.119c}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{2.119d}$$

where Equation 2.119d is the commonly seen form of the general solution to any second-order polynomial of the form of Equation 2.108.

3 Heat flux from 10 Moment Diffusive Term

Let us define the following quantities:

1. $\mathbb{Q} \equiv$ heat flux $= \mathbb{Q}_0 + \mathbb{Q}_1$
2. $\mathbb{T}(\mathbb{P}) \equiv$ temperature(pressure) tensor
3. $\mathfrak{T} \equiv \mathbb{T}^{-1}$ (inverse of the temperature tensor)
4. $\mathbb{I} \equiv$ identity tensor
5. $\tilde{\mathbb{A}} \equiv$ arbitrary 3-rank tensor
6. $\mathbb{A}\mathbf{:}\mathbb{B} \equiv$ Frobenius inner product $= \sum_{i,j} \mathbb{A}_{ij} \, \mathbb{B}_{ij}$
7. Tr[] $\equiv$ trace
8. Sym[] $\equiv$ tensor symmetrization operator

where for example, we define:

$$Sym\left[\nabla \mathbf{u}\right] = \frac{1}{2}\left[\nabla \mathbf{u} + \left(\nabla \mathbf{u}\right)^T\right] \tag{3.1a}$$

$$Sym\left[\mathbb{P} \cdot \nabla \mathbf{u}\right] = Sym\left[\nabla \cdot \left(\mathbb{P}\mathbf{u}\right) - \mathbf{u}\nabla \cdot \mathbb{P}\right] \tag{3.1b}$$

which we have chosen because in general, diffusion terms have the form $\nabla \cdot (\tilde{\mathbb{A}}\mathbf{:}\nabla \mathbf{h})$. Thus we can write:

$$\mathbb{Q} = 3A_1 Sym\left[\nabla \mathfrak{T}\right] + 3A_0 Sym\left[\mathbb{I} \cdot Tr\left[\nabla \mathfrak{T}\right]\right] \tag{3.2a}$$

$$\left[\mathbb{Q}_1\right]_{ijk} = A_1\left(\partial_i \mathfrak{T}_{jk} + \partial_k \mathfrak{T}_{ij} + \partial_j \mathfrak{T}_{ik}\right) \tag{3.2b}$$

$$\left[\mathbb{Q}_0\right]_{ijk} = A_0\left[\delta_{ij} \cdot \left(2\partial_n \mathfrak{T}_{nk} + \partial_k \mathfrak{T}_{nn}\right) + \delta_{ik} \cdot \left(2\partial_n \mathfrak{T}_{nj} + \partial_j \mathfrak{T}_{nn}\right) + \delta_{jk} \cdot \left(2\partial_n \mathfrak{T}_{ni} + \partial_i \mathfrak{T}_{nn}\right)\right] \tag{3.2c}$$

where $A_{0,1}$ are positive parameters determined by a collisional integral.

4 Rotations and Transformations

4.1 Constructing Rotation Matrices

Let's assume we have two arbitrary vectors, **A** and **B**. Let their unit vectors be denoted by: $\hat{a}$ and $\hat{b}$. If we want to find the parts of vector **A** which are parallel and perpendicular to **B**, we can do a couple of things:

1. We can find $\mathbf{A}_{\perp}$ and $\mathbf{A}_{\parallel}$ with dot and cross products, but leave the the resultant vectors in the original coordinate basis

2. We can find $\mathbf{A}_{\perp}$ and $\mathbf{A}_{\parallel}$ by rotating both vectors to a new coordinate basis where **B**' is now the Z'-Axis and **A**' is in the X'Z'-Plane.

The method to deal with the first method is the following: 1) First find the unit vectors in the typical manner:

$$\mathbf{a} \equiv \frac{\mathbf{A}}{|\mathbf{A}|} \tag{4.1a}$$

$$\mathbf{b} \equiv \frac{\mathbf{B}}{|\mathbf{B}|} \ , \tag{4.1b}$$

2) then we find the parallel vector by the following method:

$$\mathbf{a}_{\parallel} = \Big(\mathbf{a} \cdot \mathbf{b}\Big)\mathbf{b} = |\mathbf{a}||\mathbf{a}| \cos\theta_{ab}\mathbf{b} \tag{4.2a}$$

$$\mathbf{a}_{\perp} \equiv \Big(\mathbf{b} \times \mathbf{a}\Big) \times \mathbf{a} = \mathbf{a} - \Big(\mathbf{b} \cdot \mathbf{a}\Big)\mathbf{b} \ , \tag{4.2b}$$

which only need to be multiplied by the magnitude of the vector, **A**, to be turned back into vectors. It should be noted that these two vectors, $\mathbf{A}_{\parallel}$ and $\mathbf{A}_{\perp}$, satisfy the following condition:

$$|\mathbf{A}| = \sqrt{\Big(\mathbf{A}_{\parallel}\Big)^2 + \Big(\mathbf{A}_{\perp}\Big)^2} \equiv \sqrt{\left(\sum_{i}^{3} A_i^2\right)} \ . \tag{4.3}$$

The second method to find these vectors is by constructing a matrix which can rotate both vectors into a new coordinate system where **b**' is parallel to the new Z'-Axis and **a**' is in the X'Z'-Plane. To do this, we start with the unit vectors again. The first thing we do is define the following two vectors:

$$\mathbf{c} \equiv \mathbf{b} \times \mathbf{a} \tag{4.4a}$$

$$\mathbf{d} \equiv \mathbf{c} \times \mathbf{b} \tag{4.4b}$$

which we use to construct the following matrix:

$$\mathbf{R} = \begin{bmatrix} d_1 & d_2 & d_3 \\ c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \end{bmatrix} \tag{4.5}$$

The original vectors can now be rotated into a new coordinate system. Let's consider an example for illustrative purposes. Let the following be true:

$$\mathbf{A} = \{0.2, 0.3, 0.4\} \tag{4.6a}$$
$$\mathbf{B} = \{0.1, 0.5, 0.7\} \tag{4.6b}$$
$$|\mathbf{A}| = 0.5385165215 \tag{4.6c}$$
$$|\mathbf{B}| = 0.8660253882 \tag{4.6d}$$
$$\mathbf{a} = \{0.37139, 0.55709, 0.74278\} \tag{4.6e}$$
$$\mathbf{b} = \{0.11547, 0.57735, 0.80829\} \tag{4.6f}$$
$$\mathbf{c} = \{-0.02144, 0.21442, -0.15010\} \tag{4.6g}$$
$$\mathbf{d} = \{0.25997, 1.30385 \times 10^{-8}, -0.03714\} \tag{4.6h}$$

where the Y-component of **d** is a consequence of rounding errors, which I'll show turn out to actually matter. Thus our matrix is:

$$\mathbf{R} = \begin{bmatrix} 0.25997 & 1.3038 \times 10^{-8} & -0.03714 \\ -0.02144 & 0.21442 & -0.15010 \\ 0.11547 & 0.57735 & 0.80829 \end{bmatrix} \tag{4.7}$$

which produces the following new vectors:

$$\mathbf{a}' = \{0.06897, -1.84871 \times 10^{-9}, 0.964901\} \tag{4.8a}$$
$$\mathbf{b}' = \{-1.87482 \times 10^{-9}, -7.16156 \times 10^{-9}, 1.00000\} \; . \tag{4.8b}$$

One can see that **a**' and **b**' are not normalized, nor are they what we *expected* them to be. Meaning, I claimed that **b**' should be PURELY in the Z'-direction, but this has small, finite values in the X'Y'-Plane. Before we complain too much about this atrocity, let's normalize the unit vectors, which makes them now:

$$\mathbf{a}' = \{0.07129, -1.91108 \times 10^{-9}, 0.99746\} \tag{4.9a}$$
$$\mathbf{b}' = \{-1.87482 \times 10^{-9}, -7.16156 \times 10^{-9}, 1.00000\} \; . \tag{4.9b}$$

Recall that I claimed these *small* rounding errors made a difference in your final answer, so let's go back to our first set of rotated unit vectors in Equations 4.8a and 4.8b and intentionally force those *small* rounding errors to zero before we renormalize the unit vectors. Let's define these new ones as **w**' and **u**' to avoid confusion with our vectors in Equations 4.9a and 4.9b, and they become (after renormalizing):

$$\mathbf{w}' = \{0.07129, 0.00000, 0.99746\} \tag{4.10a}$$
$$\mathbf{u}' = \{0.00000, 0.00000, 1.00000\} \; . \tag{4.10b}$$

We now take the magnitudes of our original vectors and multiply that by these unit vectors to get the new vectors:

$$|\mathbf{A}| * \mathbf{a}' \equiv \mathbf{A}' = \left\{0.0383921, -1.02915 \times 10^{-9}, 0.537146\right\} \tag{4.11a}$$
$$|\mathbf{B}| * \mathbf{b}' \equiv \mathbf{B}' = \left\{-1.62364 \times 10^{-9}, -6.20209 \times 10^{-9}, 0.866025\right\} \tag{4.11b}$$
$$|\mathbf{A}| * \mathbf{w}' \equiv \mathbf{W}' = \left\{0.0383921, 0.00000, 0.537146\right\} \tag{4.11c}$$
$$|\mathbf{B}| * \mathbf{u}' \equiv \mathbf{U}' = \left\{0.00000, 0.00000, 0.866025\right\} . \tag{4.11d}$$

If we use double precision instead of single, our rotation matrix is now:

$$\mathbf{R}_d = \begin{bmatrix} 0.25997347 & -6.5919492 \times 10^{-17} & -0.037139068 \\ -0.021442251 & 0.21442251 & -0.15009575 \\ 0.11547005 & 0.57735027 & 0.80829038 \end{bmatrix} \tag{4.12}$$

which produces the following new vectors (after normalization):

$$\mathbf{a}'_d = \left\{0.071292300580, 9.63019443558 \times 10^{-18}, 0.997455466614\right\} \tag{4.13a}$$
$$\mathbf{b}'_d = \left\{-3.88116288919 \times 10^{-18}, 1.40180631841 \times 10^{-18}, 1.000000000000\right\} . \tag{4.13b}$$

Again we step back and intentionally remove the rounding errors before renormalizing to get (keep the same names this time):

$$\mathbf{a}'_d = \left\{0.071292300580, 0.000000000000, 0.997455466614\right\} \tag{4.14a}$$
$$\mathbf{b}'_d = \left\{0.000000000000, 0.000000000000, 1.000000000000\right\} . \tag{4.14b}$$

4.2 Normal Incidence Frame and Coordinate Basis

4.2.1 The Normal Incidence Frame

In this section, we will define our reference frame transformation into the Normal Incidence Frame (NIF) and coordinate basis rotations into the Normal incidence frame Coordinate Basis (NCB). We will present the transformations/rotations in a generalized manner, but for the purposes of this manuscript the measurements are in the SpaceCraft Frame (SCF) and GSE coordinate basis. We define the generalized basis as the Input Coordinate Basis (ICB). In the following, we will use the notation ${\mathbf{V}_{Coord}}^{Ref}$ to represent a 3-vector in the coordinate basis, *Coord*, and reference frame, *Ref*.

We can define the velocity transformation from any arbitrary frame of reference (e.g., SCF) to the shock frame of reference (SHF) as:

$$
{\mathbf{V}_{ICB}}^{SHF} = {\mathbf{V}_{ICB}}^{arb.} - \left({\mathbf{V}_{sh,ICB}}^{arb.} \cdot \hat{\mathbf{n}} \right) \hat{\mathbf{n}} \tag{4.15}
$$

where $\hat{\mathbf{n}}$ is the vector normal to the assumed planar shock front (see Appendix E). For an experimentalist's purposes with spacecraft observations, ${\mathbf{V}_{ICB}}^{arb.} \rightarrow {\mathbf{V}_{bulk,ICB}}^{SCF} \equiv$ the bulk flow solar wind velocity in the SCF and ICB. Let us define $\left({\mathbf{V}_{sh,ICB}}^{arb.} \cdot \hat{\mathbf{n}} \right) = {V_{sh,n}}^{SCF}$ as the shock speed along the unit normal vector, $\hat{\mathbf{n}}$, in the SCF in the upstream region and ${U_{j,n}}^{SHF}$ as the shock normal speed in the SHF, determined from the numerical Rankine-Hugoniot solution techniques [e.g., *Vinas and Scudder*, 1986; *Koval and Szabo*, 2008], in the j^{th} region. Let us also define $\langle Q \rangle_{region}$ as the spatial ensemble average of any parameter, Q, over a given space (i.e., upstream or downstream)[7].

Therefore, we can define the average upstream incident bulk flow velocity in the SHF, which is given by:

$$
\langle {\mathbf{V}_{ICB}}^{SHF} \rangle_{up} = \langle {\mathbf{V}_{bulk,ICB}}^{SCF} \rangle_{up} - \left({V_{sh,n}}^{SCF} \hat{\mathbf{n}} \right) \ . \tag{4.16}
$$

From the relationship for $\langle {\mathbf{V}_{ICB}}^{SHF} \rangle_j$, we can show that:

$$
{U_{j,n}}^{SHF} = \langle {\mathbf{V}_{ICB}}^{SHF} \rangle_j \cdot \hat{\mathbf{n}} \ . \tag{4.17}
$$

There are two physically significant frames of reference: the Normal Incidence Frame (NIF) and the de Hoffmann-Teller frame (dHT). The NIF is useful because the upstream flow velocity is entirely along the shock normal vector[8]. The dHT frame is useful because the upstream flow velocity is entirely along the upstream averaged quasi-static magnetic field ($\mathbf{B}_u$)[9]. The transformation velocity from the

[7]Note that $\hat{\mathbf{n}}$, ${V_{sh,n}}^{SCF}$, and ${U_{j,n}}^{SHF}$ are, by definition, assumed to be averages over the upstream or downstream regions. I did not include $\langle\rangle$'s out of laziness. Note I have also omitted the fact that $\hat{\mathbf{n}}$ is generally defined in the ICB in the SCF.

[8]assuming a locally planar discontinuity, as in Appendix E

[9]which results in the convective electric field $\rightarrow 0$

SHF to the NIF or dHT are given by:

$$\mathbf{V}_{ICB}{}^{NIF} = \hat{\mathbf{n}} \times \left(\langle \mathbf{V}_{ICB}{}^{SHF} \rangle_{up} \times \hat{\mathbf{n}} \right) \tag{4.18a}$$

$$\mathbf{V}_{ICB}{}^{dHT} = \frac{\hat{\mathbf{n}} \times \left(\langle \mathbf{V}_{ICB}{}^{SHF} \rangle_{up} \times \langle \mathbf{B}_{ICB}{}^{SHF} \rangle_{up} \right)}{\hat{\mathbf{n}} \cdot \langle \mathbf{B}_{ICB}{}^{SHF} \rangle_{up}} \tag{4.18b}$$

so that the upstream flow velocity in each reference frame is given by:

$$\langle \mathbf{V}_{ICB}{}^{NIF} \rangle_{up} = \langle \mathbf{V}_{ICB}{}^{SHF} \rangle_{up} - \mathbf{V}_{ICB}{}^{NIF} \tag{4.19a}$$

$$\langle \mathbf{V}_{ICB}{}^{dHT} \rangle_{up} = \langle \mathbf{V}_{ICB}{}^{SHF} \rangle_{up} - \mathbf{V}_{ICB}{}^{dHT} \; . \tag{4.19b}$$

Note that $\mathbf{V}_{ICB}{}^{NIF} = \hat{\mathbf{n}} \times \left(\langle \mathbf{V}_{ICB}{}^{SHF} \rangle_{up} \times \hat{\mathbf{n}} \right) = \hat{\mathbf{n}} \times \left(\langle \mathbf{V}_{bulk,ICB}{}^{SCF} \rangle_{up} \times \hat{\mathbf{n}} \right)$ because $\hat{\mathbf{n}} \times \hat{\mathbf{n}} = 0$. Since the change in velocity between any shock rest frame the local SC frame satisfies $|\boldsymbol{\beta}| \equiv |\Delta \mathbf{V}\ |/c \ll 1$ for any shock within the heliosphere, the Lorentz transformations of the electric and magnetic fields [page 558 of *Jackson*, 1998] can be given by:

$$\mathbf{E}' = \gamma \left(\mathbf{E} + \boldsymbol{\beta} \times \mathbf{B} \right) - \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta} \left(\boldsymbol{\beta} \cdot \mathbf{E} \right) \tag{4.20a}$$

$$\lim_{\gamma \to 1} \mathbf{E}' \approx \left(\mathbf{E} + \boldsymbol{\beta} \times \mathbf{B} \right) \tag{4.20b}$$

$$\mathbf{B}' = \gamma \left(\mathbf{B} - \boldsymbol{\beta} \times \mathbf{E} \right) - \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta} \left(\boldsymbol{\beta} \cdot \mathbf{B} \right) \tag{4.20c}$$

$$\lim_{\gamma \to 1} \mathbf{B}' \approx \mathbf{B} \; . \tag{4.20d}$$

The difference in flow velocity, $\Delta \mathbf{V}_{ICB}{}^{SCF2NIF(dHT)}$, between the SCF and relevant shock rest frames, i.e., NIF and dHT, is given by:

$$\Delta \mathbf{V}_{ICB}{}^{SCF2NIF(dHT)} = \langle \mathbf{V}_{bulk,ICB}{}^{SCF} \rangle_{up} - \langle \mathbf{V}_{ICB}{}^{NIF(dHT)} \rangle_{up} \tag{4.21a}$$

$$= \langle \mathbf{V}_{bulk,ICB}{}^{SCF} \rangle_{up} - \left[\langle \mathbf{V}_{bulk,ICB}{}^{SCF} \rangle_{up} - \left(V_{sh,n}{}^{SCF} \hat{\mathbf{n}} \right) \right] + \mathbf{V}_{ICB}{}^{NIF(dHT)} \tag{4.21b}$$

$$= \left(V_{sh,n}{}^{SCF} \hat{\mathbf{n}} \right) + \mathbf{V}_{ICB}{}^{NIF(dHT)} \tag{4.21c}$$

which allows us to show that the electric field in a relevant shock rest frame, $\mathbf{E}_{ICB}{}^{NIF,dHT}$, can be determined from the electric field observed in the SCF, $\mathbf{E}_{ICB}{}^{SCF}$, through the following:

$$\mathbf{E}_{ICB}{}^{NIF(dHT)} = \mathbf{E}_{ICB}{}^{SCF} + \left(\Delta \mathbf{V}_{ICB}{}^{SCF2NIF(dHT)} \times \mathbf{B}_{ICB}{}^{SCF} \right) \tag{4.22a}$$

$$= \mathbf{E}_{ICB}{}^{SCF} + \left[\left(V_{sh,n}{}^{SCF} \hat{\mathbf{n}} + \mathbf{V}_{ICB}{}^{NIF(dHT)} \right) \times \mathbf{B}_{ICB}{}^{SCF} \right] \; . \tag{4.22b}$$

4.2.2 NIF Coordinate Basis

We can rotate into the Normal incidence frame Coordinate Basis (NCB) from the Input Coordinate Basis (ICB) by defining a rotation matrix, $\mathbb{A}$ [*Scudder et al.*, 1986], given by:

$$\mathbb{A} = \begin{bmatrix} n_x & n_y & n_z \\ \beta_x & \beta_y & \beta_z \\ \zeta_x & \zeta_y & \zeta_z \end{bmatrix} \tag{4.23}$$

where $\hat{\mathbf{n}}$ is the shock normal vector and $\boldsymbol{\beta}$ and $\boldsymbol{\zeta}$ are given by:

$$\hat{\mathbf{y}} = \boldsymbol{\beta} = \frac{\langle \mathbf{B}_{ICB}{}^{SCF} \rangle_{dn} \times \langle \mathbf{B}_{ICB}{}^{SCF} \rangle_{up}}{\mid \langle \mathbf{B}_{ICB}{}^{SCF} \rangle_{up} \times \langle \mathbf{B}_{ICB}{}^{SCF} \rangle_{dn} \mid} \tag{4.24a}$$

$$\hat{\mathbf{z}} = \boldsymbol{\zeta} = \frac{\hat{\mathbf{n}} \times \boldsymbol{\beta}}{\mid \hat{\mathbf{n}} \times \boldsymbol{\beta} \mid} \tag{4.24b}$$

where $\langle \mathbf{B}_{ICB}{}^{SCF} \rangle_{up(dn)}$ is the average upstream(downstream) magnetic field vector. If the vectors $\hat{\mathbf{n}}$, $\boldsymbol{\beta}$, and $\boldsymbol{\zeta}$ start in the ICB (e.g., GSE), then one would expect that $\mathbb{A}$ acting on $\hat{\mathbf{n}}$, $\boldsymbol{\beta}$, or $\boldsymbol{\zeta}$ should give the corresponding NCB axis unit vector. Meaning, we expect the following to be true:

$$\mathbb{A} \cdot \hat{\mathbf{n}} = \langle 1, 0, 0 \rangle \tag{4.25a}$$

$$\mathbb{A} \cdot \boldsymbol{\beta} = \langle 0, 1, 0 \rangle \tag{4.25b}$$

$$\mathbb{A} \cdot \boldsymbol{\zeta} = \langle 0, 0, 1 \rangle \ . \tag{4.25c}$$

Thus, $\mathbb{A}$ should rotate any ICB vector into the NCB.

If the coordinate vectors used to create $\mathbb{A}$ are not orthogonal, then the correct rotation tensor is given by $\mathbb{R} = (\mathbb{A}^T)^{-1}$, or the inverse transpose of $\mathbb{A}$. The need to perform the inverse transpose of $\mathbb{A}$ arises from the non-orthogonal nature of the NIF basis. If the NIF were created from an orthogonal basis, then $\mathbb{A}$ would be an orthogonal matrix, which means $\mathbb{A}^T = \mathbb{A}^{-1}$. For any invertible matrix, the following is true: $(\mathbb{A}^T)^{-1} = (\mathbb{A}^{-1})^T$. Thus, an orthogonal NIF basis would imply $\mathbb{R} = (\mathbb{A}^T)^{-1} = (\mathbb{A}^T)^T = \mathbb{A}$. In general, however, the NIF basis vectors are not orthogonal and thus $\mathbb{R} \neq \mathbb{A}$.

5 Notes from the Shaggy Steed of Physics Book

5.1 The Action Principle

Action [*Oliver*, 2004] is mathematically defined by the integral:

$$S = \int_{t_1}^{t_2} L dt \tag{5.1}$$

where, L is defined as the Lagrangian[10]. The action principle can be stated as follows: *of all the possible paths the particles may take between any two given points in space and time, they take those paths for which the action, S, has the least possible value.* The Lagrangian is really nothing more than the difference between kinetic and potential energy, in Galilean space-time, but in its evolution, nature seeks to minimize any deviation between kinetic and potential energy, regardless of the continual interchange between the two. In relativistic space-time, the action becomes the dominating factor in what path, for any given particle, the universe will conspire to create. If we define the total energy as:

$$H = \sum_{\alpha} \mathbf{p}_{\alpha} \cdot \dot{\mathbf{x}}_{\alpha} - L\left(\mathbf{x}_{\alpha}, \dot{\mathbf{x}}_{\alpha}\right) \tag{5.2}$$

where $\mathbf{p}_{\alpha}$ is the momentum of a particle defined by:

$$\mathbf{p}_{\alpha} \equiv \frac{\partial L\left(\mathbf{x}_{\alpha}, \dot{\mathbf{x}}_{\alpha}\right)}{\partial \dot{\mathbf{x}}_{\alpha}} \ . \tag{5.3}$$

The kinetic energy can be defined as:

$$T \equiv \frac{1}{2} \sum_{\alpha} \frac{\partial L\left(\mathbf{x}_{\alpha}, \dot{\mathbf{x}}_{\alpha}\right)}{\partial \dot{\mathbf{x}}_{\alpha}} \cdot \dot{\mathbf{x}}_{\alpha} \ , \tag{5.4}$$

and the angular momentum is:

$$\mathbf{J} \equiv \mathbf{x}_{\alpha} \times \frac{\partial L\left(\mathbf{x}_{\alpha}, \dot{\mathbf{x}}_{\alpha}\right)}{\partial \dot{\mathbf{x}}_{\alpha}} \ . \tag{5.5}$$

We are also met with another way of describing what is meant when one claims *the action must be minimized:* think of the bottom of a valley as minimum in potential energy. It can also be said that if the valley is a smoothly varying "bowl," if you will, then one might claim that the bottom of the valley has an approximately zero slope. What does it mean to have no *slope*? For any given function, f(q_1,q_2,q_3,...,q_n), the following statement defines what it means to have *no slope* at some point, q_o, in space-time:

$$\frac{\partial f}{\partial q_i}\Big|_{q_i = q_o} = 0 \ , \tag{5.6}$$

[10]Note: Oliver refers to the Lagrangian as the *the gene of motion*.

or one could also say that the *variation* of a function must vanish at some point, q_o, in space-time:

$$\delta f = \frac{\partial f}{\partial q_i} \delta q_i = 0 \ , \tag{5.7}$$

where δq_i are the *arguments* of the function, f. The variation of the function, f, illustrate how small changes in its arguments, q_i, cause changes in the function itself. That means, if the partial derivative of one of the arguments vanishes, the variation in f suffers no change (e.g. $\partial f/\partial q_i = 0$, thus $\delta f = 0$ regardless of δq_i). Since the variation of the arguments, δq_i, are arbitrary, the vanishing variation of f requires that every partial derivative, $\partial f/\partial q_i$, vanishes at the arbitrary point, q_o. Thus, *least action* is define as the conditions for which the functions, q(t), satisfy the requirements to force $\delta S = 0$. The action is defined as a *functional* ≡ *a quantity that has a single value corresponding to an entire function.* Thus, the variation of the action is defined as:

$$\delta S \equiv \int_{t_1}^{t_2} \delta L dt = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) dt \tag{5.8}$$

where the second term in the equation is defined as:

$$\frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i = \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \Big(\delta q_i \Big) \tag{5.9a}$$

$$= \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) - \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i \ . \tag{5.9b}$$

An important thing to note is that the path variations, δq, vanish at the end points, $(\delta q,t)_1$ and $(\delta q,t)_2$. So we look at the first term in Equation 5.9b and notice the following:

$$\int_{t_1}^{t_2} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) dt = \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) \Bigg|_{t_1}^{t_2} = 0 \ . \tag{5.10}$$

So we then have the following from Equation 5.9b:

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i dt = - \int_{t_1}^{t_2} \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i dt \tag{5.11}$$

which means we've now transformed the second term in Equation 5.8 into a variation of δ q_i, instead of $\delta \dot{q}_i$. This implies that we can do the following:

$$\delta S = - \int_{t_1}^{t_2} \left(\frac{d}{dt} \Big(\frac{\partial L}{\partial \dot{q}_i} \Big) - \frac{\partial L}{\partial q_i} \right) \delta q_i dt \ . \tag{5.12}$$

5.2 Other important definitions

The Total Time Derivative

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \left(\frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i \right) \tag{5.13a}$$

$$= \frac{\partial f}{\partial t} + \Big\{ f, \mathcal{H} \Big\} \tag{5.13b}$$

where $\{f,\mathcal{H}\}$ is called a *Poisson bracket* and $\mathcal{H}$ is the Hamiltonian, defined by:

$$\Big\{ f, g \Big\} \equiv \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i} \right) . \tag{5.14}$$

It is important to note that the position-momentum pair is an *antisymmetric manifestation of a symplectic structure*. Here are some general rules of Poisson brackets:

1. If both f and g are scalars, $\{f,g\}$ is a scalar
2. If f is a vector and g is a scalar, $\{f,g\}$ is a vector
3. If both f and g are vectors, $\{f,g\}$ is a second rank tensor.

The Hamilton equations of motion allow/show an example of where this notation can be of constructive use with the following two examples/definitions:

$$\dot{q}_i = \Big\{ q_i, \mathcal{H} \Big\} \tag{5.15a}$$

$$\dot{p}_i = \Big\{ p_i, \mathcal{H} \Big\} \tag{5.15b}$$

$$\Big\{ q_i, q_j \Big\} = 0 \tag{5.15c}$$

$$\Big\{ p_i, p_j \Big\} = 0 \tag{5.15d}$$

$$\Big\{ q_i, p_j \Big\} = \delta_{ij} \tag{5.15e}$$

$$\begin{aligned} \Big\{ J_i, J_j \Big\} &= \epsilon_{ijk} J_k \\ \mathbf{J} \cdot \Big\{ \mathbf{J}, J_j \Big\} &= J_i \Big\{ J_i, J_j \Big\} \\ &= \epsilon_{ijk} J_i J_k \\ &\equiv 0 \end{aligned} \tag{5.16a}$$

$$\begin{aligned}\left\{J^2, J_j\right\} &= 0 \\ &= \left\{J_i J_i, J_j\right\} \\ &= 2J_i\left\{J_i, J_j\right\}\end{aligned} \tag{5.17a}$$

$$\left\{x_i, J_j\right\} = \epsilon_{ijk} x_k \tag{5.18a}$$

$$\left\{p_i, J_j\right\} = \epsilon_{ijk} p_k \tag{5.18b}$$

$$\left\{F_i, J_j\right\} = \epsilon_{ijk} F_k \tag{5.18c}$$

$$\left\{f, J_j\right\} = 0 \tag{5.18d}$$

(where $\mathbf{F}$ is any arbitrary vector function and f is any arbitrary scalar). Phase space is always an even-dimensional manifold that describes the space of motion. This motion fills phase space with the phase trajectories described by q(t) and p(t). The essence of configuration space is Euclidean while phase space is symplectic[11]. If we define the following as the single 2s state vector of phase space:

$$\xi = \left(q, p\right) \tag{5.19}$$

where the first s-components of ξ are position coordinates of all the particles in the phase space you're trying to describe. The *symplectic* is a 2s $\times$ 2s antisymmetric matrix defined as:

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \tag{5.20}$$

where I is the identity or unit matrix (of dimension s). The J is a perfect example of symplectic space (Oliver calls it the *signature* of symplectic phase space). It has the following properties:

$$J = -J^{\dagger} = -J^{-1} \tag{5.21}$$

where $J^{\dagger}$ is defined by:

$$J^{\dagger}_{ij} = -J^{*}_{ji} \tag{5.22}$$

or, in other words, the *transpose conjugate*. The symplectic also has a square:

$$J^2 = -\begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} = -J \tag{5.23}$$

[11] *Symplectic structure* is named after the Greek word, $\pi\lambda\epsilon\kappa\acute{o}\varsigma$, meaning *twined* or *braided*. It is an antisymmetric pairing of coordinates induced by the action principle.

and it has the *defining property* of inducing a vanishing scalar product on any phase space vector:

$$\xi_i J_{ij} \xi_j = 0 \text{ (often seen as } \xi J \xi = 0). \tag{5.24}$$

The symplectic allows us to rewrite the Poisson bracket notation in a different manner:

$$\Big\{ f, g \Big\} = \frac{\partial f}{\partial \xi} J \frac{\partial g}{\partial \xi} \tag{5.25}$$

which allows one to look at the Poisson bracket of any function, f(ξ), with the phase space vector, ξ:

$$\Big\{ \xi, f \Big\} = J \frac{\partial f}{\partial \xi} \ . \tag{5.26}$$

We are finally allowed to look at the Hamiltonian equations using the symplectic and phase space vectors:

$$\dot{\xi} = \Big\{ \xi, \mathcal{H} \Big\} = J \frac{\partial \mathcal{H}}{\partial \xi} \ . \tag{5.27}$$

A useful relationship between any two quantities, F$\big(\xi\big)$ and G$\big(\xi\big)$, can be shown to be:

$$\Big\{ F, G \Big\} = \frac{\partial F}{\partial \xi} J \frac{\partial G}{\partial \xi} = \frac{\partial F}{\partial \xi} \cdot \Big\{ \xi, G \Big\} = -\frac{\partial G}{\partial \xi} \cdot \Big\{ \xi, F \Big\} \ . \tag{5.28}$$

So the Poisson bracket illustrates a number of points:

1. The Poisson bracket is a *projection of the normals of the level surfaces of one quantity upon the tangents of the level surfaces of the other.*

2. If $\big\{$F,G$\big\} \neq 0$, the *flow of F* does NOT stay on level surfaces of G, rather it cuts ACROSS them! $\Rightarrow$ G is NOT a constant on the flow of F.

3. If $\big\{$F,G$\big\} = 0$, the *flow of F* not only stays on its own level surface, it is on the level surfaces of G too. $\Rightarrow$ The two quantities become one common surface to which both flows are confined.

4. Every mechanical quantity, F$\big(\xi\big)$, has an image in phase space as *sheets* of level surfaces filled with streamlines generated by $\big\{\xi$,F$\big\} \equiv$ the flow. $\Rightarrow$ The Poisson bracket describes the intersection of the two flows produced by F$\big(\xi\big)$ and G$\big(\xi\big)$.

5. The *phase flow* is incompressible.

5.3 Hamilton-Jacobi Theory

The motion of the world is imaged as a flow in phase space. The manner in which to find these *flows* involves the integrals to Hamilton's equations. The mathematical forms of the action principle, Hamilton's equations, and Poisson brackets are independent of the coordinate system they are expressed in. *Motion with s-degrees of freedom has 2s canonical coordinates which form s-conjugate pairs.* $\Rightarrow$ Any set of canonical coordinates is related to another set by transformations which preserve the action principle. Consider the set of canonical coordinates (Q,P), where Q = (Q$_1$,Q$_2$,...,Q$_s$) and P =

(P_1,P_2,...,P_s). Though it may appear that Q and P are actual position and momentum coordinates, they need not be. Now Hamilton's equations are:

$$\dot{Q}_i = \frac{\partial \mathcal{H}}{\partial P_i} \tag{5.29a}$$

$$\dot{P}_i = -\frac{\partial \mathcal{H}}{\partial Q_i} \tag{5.29b}$$

so that now $\mathcal{H} \rightarrow \mathcal{H}'\big(Q, P\big)$ which still satisfies:

$$\delta S = \delta \int_{t_1}^{t_2} \Big(p_i dq_i - \mathcal{H} dt\Big) = \delta \int_{t_1}^{t_2} \Big(P_i dQ_i - \mathcal{H}' dt\Big) \ . \tag{5.30}$$

These two integrals may differ by any function, F, which has a vanishing variation (i.e. δF = 0). Thus the difference between the integrals in Equation 5.30 must be the total differential of F:

$$dF = p_i dq_i - P_i dQ_i - \Big(\mathcal{H} - \mathcal{H}'\Big) dt \tag{5.31}$$

which defines what is referred to as *the generating function*. The generating function, from Equation 5.31, is F = F(q,Q,t). The relationship of any coordinate can be described as follows:

$$p_i = \frac{\partial F}{\partial q_i} \tag{5.32a}$$

$$P_i = -\frac{\partial F}{\partial Q_i} \tag{5.32b}$$

$$\Big(\mathcal{H} - \mathcal{H}'\Big) = \frac{\partial F}{\partial t} \tag{5.32c}$$

but the generating function can be written in a different form as:

$$dG = d\Big(F + P_i Q_i\Big) = p_i dq_i + P_i dQ_i - \Big(\mathcal{H} - \mathcal{H}'\Big) dt \ . \tag{5.33}$$

Here, G = G(q,P,t), where the coordinates are:

$$p_i = \frac{\partial G}{\partial q_i} \tag{5.34a}$$

$$Q_i = \frac{\partial G}{\partial P_i} \tag{5.34b}$$

$$\Big(\mathcal{H} - \mathcal{H}'\Big) = -\frac{\partial G}{\partial t} \tag{5.34c}$$

1. The generating function incorporates ONE coordinate from the old pair and ONE coordinate from the new pair.

2. The canonical transformation presents one of the coordinates explicitly and one of them implicitly. Meaning, in the transformation from the state (q,p) to the state (Q,P) by the generating function, G(q,P,t), the implicit-explicit nature can be seen in Equations 5.35a and 5.35b.

3. *Explicit Dependence* $\equiv$ A direct relationship between two quantities, e.g. f(t) explicitly depends on t if the variable t exists directly in the function f(t), NOT if a variable in f(t) has a dependence on time. Meaning, $\mathrm{f}\big(\mathbf{x}(\mathrm{t})\big)$ would not explicitly depend on t UNLESS t existed indpendent of $\mathbf{x}$(t) in the function.

4. *Implicit Dependence* $\equiv$ An indirect relationship between two quantities, e.g. $\mathrm{f}\big(\mathbf{x}(\mathrm{t})\big)$ implicitly depends on the variable t, but $\mathbf{x}$ explicitly depends on t

$$\frac{\partial}{\partial q_i} G\Big(q, P, t\Big) = p_i \tag{5.35a}$$

$$Q_i = \frac{\partial}{\partial P_i} G\Big(q, P, t\Big) \tag{5.35b}$$

where the new coordinates, $\mathrm{Q}\big(\mathrm{q,p}\big)$, are given **explicitly** by Equation 5.35b and the coordinates, $\mathrm{P}\big(\mathrm{q,p}\big)$, are given **implicitly** by Equation 5.35a. The two following examples illustrate some trivial transformations:

$$G\Big(q, P, t\Big) = q_i P_i \tag{5.36a}$$

$$F\Big(q, Q, t\Big) = q_i Q_i \tag{5.36b}$$

yield the following identity and inverse-identity transformations: 1) for $\mathrm{G}\big(\mathrm{q,P,t}\big)$ we have the identity transformation given by:

$$Q_i = q_i \tag{5.37a}$$

$$P_i = p_i \tag{5.37b}$$

$$\mathcal{H}' = \mathcal{H} \tag{5.37c}$$

and 2) for $\mathrm{F}\big(\mathrm{q,Q,t}\big)$ we have the inverse-identity transformation given by:

$$Q_i = \quad p_i \tag{5.38a}$$

$$P_i = -q_i \tag{5.38b}$$

$$\mathcal{H}' = \quad \mathcal{H} \; . \tag{5.38c}$$

A nineteenth-century astronomer, C.E. Delaunay, used the fact that the transformations must ONLY be canonical in nature by attempting to simplify the problem of motion. In his attempts, he found coordinates that are now referred to as *elementary flow* coordinates, which occur when one of the coordinates, P_i or Q_i, are selected as constants (i.e. assume we chose $\mathrm{P}_i \equiv \mathrm{I}_i$, then $\dot{I}_i = 0$ and $\mathrm{Q}_i \equiv \alpha_i$). Now Hamilton's equations simplify dramatically to:

$$\dot{\alpha}_i = \quad \frac{\partial \mathcal{H}'}{\partial I_i} \tag{5.39a}$$

$$\dot{I}_i = -\frac{\partial \mathcal{H}'}{\partial \alpha_i} = 0 \tag{5.39b}$$

which simplifies our job of solving Hamilton's equations even more than just having Equation 5.39b be null. The reason for the underlying simplicity is due to the coordinate's symplectic union. This equation also shows us that the Hamiltonian is now only a function of one canonical coordinate, namely $\mathcal{H}$' = $\mathcal{H}$'$\left(I\right)$. We may now rewrite the R.H.S. of Equation 5.40 as a function of only I also:

$$\omega_i\left(I\right) \equiv \frac{\partial \mathcal{H}'}{\partial I_i} \ . \tag{5.40}$$

This reformulation of the coordinates allows for a remarkably trivial integral form, which might otherwise be seemingly impossible:

$$\alpha_i = \int dt \Big(\frac{\partial \mathcal{H}'}{\partial I_i}\Big) = \omega_i t + \beta_i \tag{5.41}$$

where β_i (i = 1, 2, . . . , s) are the integration constants.

1. The elementary flow ONLY "flows" along the α-coordinates $\Rightarrow$ its phase velocity has no components in the invariant coordinate, I, since $\dot{I} = 0$.

2. The integrals of elementary flow depend upon the two constants of integration, I and β, which make up the two sets of s quantities.

3. We can define the *Phase Vector* (Equation 5.42a), the phase vector's *Phase Velocity* (Equation 5.42b), and the *Integration Constants* (Equation 5.42c).

$$\Xi = \left(\alpha, I\right) \tag{5.42a}$$
$$\Omega = \left(\omega, 0\right) \tag{5.42b}$$
$$\mathcal{I} = \left(\beta, I\right) \tag{5.42c}$$

Thus the elementary flow can be described by:

$$\Xi\left(t\right) = \Omega t + \mathcal{I} \ . \tag{5.43}$$

The elementary flow can be seen to depend upon the constants of flow, $\mathcal{I}$, which are also invariant in coordinate phase space because: *Quantities invariant on the flow in one set of coordinates are invariant on the IMAGE of this flow in all other canonical coordinates* $\Rightarrow$ they are the 2s invariants of motion! This is important because the elementary phase space coordinates, $\Xi = \left(\alpha, I\right)$, are NOT connected to their image phase space coordinates, $\xi = \left(q, p\right)$, in any simple way (typically a VERY "ugly" transformation connects them). However, the invariants are the "same" in both phase spaces, linking the two together, satisfying an important conservation law:

$$\Delta\left(\mathcal{I}\right) = 0 \ . \tag{5.44}$$

These invariants also satisfy the condition that its rate of change along the "flow" (i.e. its total time derivative) vanishes by:

$$\frac{d\mathcal{I}}{dt} = \frac{\partial \mathcal{I}}{\partial t} + \Big\{\mathcal{I}, \mathcal{H}\Big\} = 0 \ . \tag{5.45}$$

It is often the case where $\mathcal{I} \neq \mathcal{I}(t)$, **explicitly** (i.e. $\partial \mathcal{I}/\partial t = 0$), but only a function of the canonical phase space coordinates, $\mathcal{I} = \mathcal{I}(q, p)$. Thus any quantity which does **NOT EXPLICITLY** depend upon time satisfies the following:

$$\left\{\mathcal{I}, \mathcal{H}\right\} = 0 \ , \tag{5.46}$$

namely, it's Poisson bracket with the Hamiltonian vanishes.

5.4 Action Again

Since we can say that the Lagrangian is really just the total time derivative of the action, we can also say:

$$dS = -\mathcal{H}dt + p_i dq_i = \frac{\partial S}{\partial t}dt + \frac{\partial S}{\partial q_i}dq_i \ . \tag{5.47}$$

As one might expect from our previous treatments of such things, we can say:

$$\mathcal{H} = -\frac{\partial S}{\partial t} \tag{5.48a}$$

$$p_i = \quad \frac{\partial S}{\partial q_i}dq_i \tag{5.48b}$$

and since $\mathcal{H} = \mathcal{H}(q, p)$, we can rewrite Equation 5.48a as:

$$\frac{\partial S}{\partial t} + \mathcal{H}\left(q, \frac{\partial S}{\partial q}\right) = 0 \ . \tag{5.49}$$

We now know that $\mathrm{S} = \mathrm{S}(q, t)$ is a solution to the 1^{st}-Order partial differential equation in the s-position coordinates, q, and the time, t. The solution, in general, depends upon $s + 1$ constants of integration, where one of these constants is purely additive; meaning, if $\mathrm{S}(q, t)$ is a solution of the Hamilton-Jacobi equation, then $\mathrm{S}(q, t) + A$ is too (assuming A is an additive constant)[12]. We also assume the total energy of our system is constant, meaning, $\mathcal{H} = \mathcal{E} = \partial S/\partial t$, which leads to an action of the form:

$$S\left(q, I, t\right) = -\mathcal{E}t + S_o\left(q, I\right) \tag{5.50}$$

where $\mathrm{S}_o\left(q, I\right)$ is the time-independent part of the action. One should also note that the constant, $\mathcal{E}$, is one of the invariants, $I = \left(I_1, I_2, \ldots, I_s\right)$. So now we have the new momentum-like coordinates, $\mathrm{P} \equiv I$ in $\mathrm{G}(q, P, t) \equiv \mathrm{S}_o\left(q, I\right)$. This leaves the remaining coordinates as:

$$p_i = \frac{\partial}{\partial q_i}S_o\left(q, I\right) \tag{5.51a}$$

$$\alpha_i = \frac{\partial}{\partial I_i}S_o\left(q, I\right) \tag{5.51b}$$

$$\tag{5.51c}$$

[12]So, it's not entirely clear why, but the remaining constants must be invariants of motion.

or they may also be expressed as:

$$p_i = \frac{\partial}{\partial q_i} S\Big(q, I, t\Big) \tag{5.52a}$$

$$\beta_i = \frac{\partial}{\partial I_i} S\Big(q, I, t\Big) \tag{5.52b}$$

$$\tag{5.52c}$$

which leads us to the conclusion that the Hamiltonians in the two sets of coordinates are the same, just of different form, given by:

$$\mathcal{H}'\Big(I\Big) = \mathcal{H}\Big(q, p\Big) \ . \tag{5.53}$$

We already know that we can write the action as an integral of the canonical coordinates:

$$S = \int \Big(p_i dq_i - \mathcal{H} dt\Big) \tag{5.54}$$

but this form is *clearly* not an invariant[13], so we reconsider this case as a contour integral over a closed contour, γ, in phase space as:

$$S = \oint_{\gamma} \Big(p_i dq_i - \mathcal{H} dt\Big) \ . \tag{5.55}$$

One should note that the contour, γ, itself is not invariant because it is deformed as it is swept through phase space by the flow, however, the integral over this moving contour can be invariant[14]. This closed integral is the Poincar*é* invariant. Though this invariant exists for all motion, its form is completely opaque unless the canonical coordinates are known functions and one can actually solve the integrals.

5.4.1 Hooke Motion

As a way to illustrate how these transformations work, we'll consider a few examples. The first of which is an idea proposed by Robert Hooke, one of Newton's most prominent contemporaries[15] involved a force corresponding to the potential:

$$V\big(q\big) = \frac{\kappa q^2}{2} \ , \tag{5.56}$$

where κ is a constant. It gives rise to a force, denoted by:

$$-\frac{\partial}{\partial q} V\big(q\big) = -\kappa q \equiv \mathcal{F} \ , \tag{5.57}$$

[13]Due the indefinite nature of the integral and the lack of rotational invaraince in the $\mathrm{p}_i \mathrm{dq}_i$ terms, the integral can't be said to be an invariant of motion.

[14]The invariance arises from the integration within a closed boundary. When one considers the integral at hand, one can see we are really integrating the Lagrangian within a closed boundary. That means, for this integral to NOT be an invariant of motion, requires a violation of the conservation of energy

[15]Hooke happened to be a rather short man, while Newton was very tall. Both men did not get along very well, and as a way to mock Hooke, Newton said the famous quote: *If I have seen further it is by standing on ye shoulders of Giants.*

which really doesn't correspond to any fundamental force in nature, it's only an approximation to the force between two bodies bound by an elastic material (e.g. a spring)[16]. The Hamiltonian can be written as:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{\kappa q^2}{2} \ , \tag{5.58}$$

corresponding to the Hamiltonian for the simple harmonic oscillator with natural frequency, $\omega_o = \sqrt{\kappa/m}$. The canonical invariant of motion and its elementary flow (like all elementary flows) turns out to be:

$$I = \mathcal{H}/\omega_o \tag{5.59a}$$
$$\alpha_i = \omega_i + \beta_i \tag{5.59b}$$

which gives us a new Hamiltonian, $\mathcal{H}'\big(I\big) = \omega_o I$, and the coordinates, $\big(\alpha, I\big)$. The generating function for this case is the action found in Equation 5.50. This can be found from solving Equation 5.49 for one degree of freedom, which reduces to:

$$\frac{dS_o}{dq} = p\big(q\big) \tag{5.60}$$

After some algebraic manipulation, the integral for S$_o$ can be found[17] to be:

$$S_o\big(q, I\big) = \int dq \sqrt{2m\omega_o \Big(I - \frac{m\omega_o q^2}{2}\Big)} \ . \tag{5.61}$$

[16]Oliver goes into a paragraph-long explanation of how the inverse-square law Coulomb force actually averages out to a linear force when dealing with the macroscopic scale of a spring. This results in enormous cancelations of forces.

[17]Solve for p(q) in Equation 5.58 by replacing κ with $m\omega_o^2$, and $\mathcal{H}$ with $\mathcal{H}' = \mathrm{I}\omega_o$.

A Definitions

A.1 Symbols and Parameters

1. **Fundamental Constants**
   (a) $\varepsilon_o \equiv$ permittivity of free space [F m^{-1} or s^2 C^2 kg^{-1} m^{-3}]
   (b) $\mu_o \equiv$ permeability of free space [H m^{-1} or kg m C^{-2}]
   (c) $c = 1/\sqrt{\mu_o \varepsilon_o} \equiv$ speed of light in vacuum [m s^{-1}]
   (d) $e \equiv$ fundamental charge [C]
   (e) $\mathrm{k}_B \equiv$ Boltzmann constant [J K^{-1} or kg m^2 s^{-2} K^{-1}]
2. **Particle/Plasma-Related Parameters**
   (a) $\mathrm{m}_s \equiv$ mass of particle species s [kg]
   (b) $\mathrm{q}_s \equiv$ charge of particle species s [C]
   (c) $\mathrm{n}_s \equiv$ number density of particle species s [m^{-3}]
   (d) $\rho_s = \mathrm{m}_s\ \mathrm{n}_s \equiv$ mass density of particle species s [kg m^{-3}]
   (e) $\mathrm{B}_o \equiv$ quasi-static magnetic field magnitude [T]
   (f) $\mathbf{j} \equiv$ electrical current density [e.g., from Ampere's law]
   (g) $\omega_{ps} = \sqrt{n_s {q_s}^2/(m_s \varepsilon_o)} \equiv$ plasma frequency of particle species s [rad s^{-1}]
   (h) $\Omega_{cs} = \mathrm{q}_s \mathrm{B}_o/(\gamma\ \mathrm{m}_s) \equiv$ cyclotron frequency of particle species s [rad s^{-1}][18]
   (i) $\mathbf{V}_s \equiv$ bulk flow velocity[19] of species s [m s^{-1}]
   (j) $\mathrm{T}_s \equiv$ average temperature[20] of particle species s [eV or K]
   (k) $\mathrm{V}_{Ts} = \sqrt{(2 k_B T_s)/m_s} \equiv$ average thermal speed[21] of particle species s [m s^{-1}]
   (l) $\lambda_s = \mathrm{c}/\omega_{ps} \equiv$ inertial length (or skin depth) of particle species s [m]
   (m) $\lambda_{Ds} = \sqrt{(\varepsilon_o k_B T_s)/(n_s {q_s}^2)} \equiv$ Debye length of particle species s [m]
   (n) $\rho_{cs} = \mathrm{V}_{Ts}/\omega_{ps} \equiv$ thermal gyroradius of particle species s [m]
   (o) $\mathrm{V}_A = \sqrt{{B_o}^2/(\mu_o m_i n_i)} \equiv$ Alfvén speed [m s^{-1}][22]
   (p) $\mathrm{C}_s = \sqrt{k_B \left(Z_i \gamma_e T_e + \gamma_i T_i\right)/m_i} \equiv$ ion sound speed in an electron-ion plasma with differing ratios of specific heats[23], γ_e and γ_i, for each species and ion charge state, Z_i.
   (q) $\mathrm{P}_s = \mathrm{n}_s\ \mathrm{k}_B\ \mathrm{T}_s \equiv$ thermal pressure of particle species s [Pa or N m^{-2} or J m^{-3} or kg m^{-1} s^{-2}]
3. **Wave/Fluctuation-Related Parameters**
   (a) $\omega \equiv$ angular frequency (typically used for waves)
   (b) $\mathrm{k} \equiv$ wavenumber [$= 2\pi/\lambda$]
   (c) $\lambda \equiv$ wavelength (typically when no subscript is present)
4. **Useful Relationships**
   (a) $\omega_{pi} = (c\ \Omega_{ci})/\mathrm{V}_A$
   (b) $\omega_{pe} = (c\ \Omega_{ce})/\mathrm{V}_{Ae}$

[18] $\gamma \equiv$ Lorentz frame transformation factor $= \left[1 - (v/c)^2\right]^{-1/2}$
[19] refers to the 1st velocity moment
[20] refers to the 2nd velocity moment in the bulk flow rest frame per unit mass
[21] here it is the *most probable speed*, whereas $\mathrm{V}_{Ts}/\sqrt{2}$ is the *root mean square speed*
[22] we also refer to an electron Alfvén speed, V_{Ae}, on occasion, but it does not have the same physical significance as V_A
[23] typical values are $\gamma_e = 3$ (or 1) and $\gamma_i = 1$

A.2 Terminology and Jargon

1. **Phase Front:** plane of constant phase
2. $\mathbf{V}_{ph} \equiv$ phase velocity
   (a) the velocity associated with a fixed value of phase $\Rightarrow$ representing an advance of position, $\mathbf{r}$, with t [see page 233 of *French*, 1971]
3. **Dispersive:** a medium where the phase speed of a wave depends upon the frequency of the wave [see page 398 of *Griffiths*, 1999]
4. $\mathbf{V}_{gr} \equiv$ group velocity
   (a) the velocity associated with a modulated envelope, which encloses a group of phase fronts (or short waves)
   (b) so long as the wave source is slowly varying, constructive interference will maximize where the phase is stationary (i.e., where $d\mathbf{r}/dt = \partial\omega/\partial\mathbf{k}$) $\Rightarrow$ the locus of points satisfying this condition define the group velocity [see page 76 of *Stix*, 1992]
   (c) is perpendicular to a contour of constant ω in $\mathbf{k}$-space [pages 82-83 of *Gurnett and Bhattacharjee*, 2005]
5. **Index of Refraction:** a dimensionless vector that has the direction of $\mathbf{k}$ and the magnitude of $c/\mathbf{V}_{ph}$ (defined as $\mathbf{n}$)
6. **Wave Normal Surface:** the locus of $\mathbf{V}_{ph}$ azimuthally revolved around $\mathbf{B}_o$ with a 2D cross-section shown as a polar plot of ω/k vs. θ (= angle between $\mathbf{k}$ and $\mathbf{B}_o$)
   (a) formed by plotting u $(= \omega/kc = 1/n)$ vs. θ
   (b) formed by the locus of the tip of the vector $\mathbf{n}^{-1} \equiv \mathbf{n}/n^2$

B Maxwell Equations

We start with the Maxwell equations:

$$\nabla \cdot \mathbf{E} = \frac{\rho_c}{\varepsilon_o} \tag{B.1a}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{B.1b}$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{B.1c}$$

$$\nabla \times \mathbf{B} = \mu_o \mathbf{j} + \mu_o \varepsilon_o \frac{\partial \mathbf{E}}{\partial t} \tag{B.1d}$$

Since we know from vector calculus that the divergence of the curl of a vector is zero, then from Equation B.1b we can define the vector potential as:

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{B.2}$$

which we then substitute into Equation B.1c to get:

$$\nabla \times \left[\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right] = 0 \ . \tag{B.3}$$

From vector calculus, we know that the curl of the gradient of a scalar function is zero, thus we can say:

$$\mathbf{E} = -\nabla \Phi - \frac{\partial \mathbf{A}}{\partial t} \tag{B.4}$$

where the negative sign is chosen due to the physical relationship between potentials and forces. The arbitrariness in the choice of $\mathbf{A}$ and Φ suggest that the following two operations would not affect the electric or magnetic fields:

$$\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla \Lambda \tag{B.5a}$$

$$\Phi \rightarrow \Phi' = \Phi - \frac{\partial \Lambda}{\partial t} \tag{B.5b}$$

which means we can choose any ($\mathbf{A}$, Φ) such that they satisfy:

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \Phi}{\partial t} = 0 \ . \tag{B.6}$$

This allows us to rewrite Equations B.1a and B.1d as:

$$\nabla^2 \Phi - \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} = -\frac{\rho}{\varepsilon_o} \tag{B.7a}$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_o \mathbf{j} \ . \tag{B.7b}$$

For more information, see Chapter 6 of *Jackson* [1998].

B.1 Poynting's Theorem

We can define the, if there exists a continuous distribution of charge and current, the total rate of work done by electromagnetic fields in a volume by:

$$\frac{dW_{EM}}{dt} = \int_V d^3x \, \mathbf{j} \cdot \mathbf{E} \ . \tag{B.8}$$

This power represents the rate at which electromagnetic field energy is converted into mechanical or thermal energy. To conserve energy, the increase in mechanical or thermal energy must be balanced by a decrease in electromagnetic field energy. We can change the form of this equation and represent it in terms of fields only by using Equation B.1d to replace $\mathbf{j}$, which results in:

$$\frac{dW_{EM}}{dt} = \int_V d^3x \ \left[\frac{1}{\mu_o} \nabla \times \mathbf{B} - \varepsilon_o \frac{\partial \mathbf{E}}{\partial t} \right] \cdot \mathbf{E} \tag{B.9a}$$

$$= \int_V d^3x \ \left[\frac{1}{\mu_o} \mathbf{E} \cdot (\nabla \times \mathbf{B}) - \varepsilon_o \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} \right] \tag{B.9b}$$

$$= \int_V d^3x \ \left\{ \frac{1}{\mu_o} \left[\mathbf{B} \cdot (\nabla \times \mathbf{E}) - \nabla \cdot (\mathbf{E} \times \mathbf{B}) \right] - \varepsilon_o \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} \right\} \tag{B.9c}$$

$$= -\frac{1}{\mu_o} \int_V d^3x \ \left\{ \left[\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} + \nabla \cdot (\mathbf{E} \times \mathbf{B}) \right] + \mu_o \varepsilon_o \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} \right\} \tag{B.9d}$$

$$= -\int_V d^3x \ \left\{ \nabla \cdot \left(\frac{\mathbf{E} \times \mathbf{B}}{\mu_o} \right) + \frac{\partial}{\partial t} \left[\frac{\mathbf{B} \cdot \mathbf{B}}{2\mu_o} + \frac{\varepsilon_o (\mathbf{E} \cdot \mathbf{E})}{2} \right] \right\} \tag{B.9e}$$

$$-\mathbf{j} \cdot \mathbf{E} = \nabla \cdot \mathbf{S} + \frac{\partial}{\partial t} \left(\mathcal{W}_B + \mathcal{W}_E \right) \tag{B.9f}$$

where we have used $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A})$ - $\mathbf{A} \cdot (\nabla \times \mathbf{B})$, Equation B.1c, and $\mathbf{S}$ is the Poynting flux. $\mathbf{S}$ is the flow of electromagnetic field energy per unit area per unit time. Since $(\mathbf{j} \cdot \mathbf{E})$ is a rate of work to convert electromagnetic field energy into a mechanical energy, E_{mech}, we can relate this to Newton's 2nd law through the following:

$$\frac{d\mathbf{P}_{mech}}{dt} = \int_V d^3x \ \left[\rho_c \mathbf{E} + \mathbf{j} \times \mathbf{B} \right] \tag{B.10}$$

where ρ_c is the charge density and $\mathbf{P}_{mech}$ is the total momenta of all particles in the volume. Using

Equations B.1a and B.1d, we can change the integrand to the following:

$$\left[\rho_c \mathbf{E} + \mathbf{j} \times \mathbf{B}\right] = \varepsilon_o \mathbf{E} \left(\nabla \cdot \mathbf{E}\right) + \frac{1}{\mu_o} \left(\nabla \times \mathbf{B}\right) \times \mathbf{B} - \varepsilon_o \frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B} \tag{B.11a}$$

$$= \varepsilon_o \left[\mathbf{E} \left(\nabla \cdot \mathbf{E}\right) + \mathbf{B} \times \frac{\partial \mathbf{E}}{\partial t} - c^2 \mathbf{B} \times \left(\nabla \times \mathbf{B}\right)\right] \ . \tag{B.11b}$$

We can manipulate this further by using the following:

$$\mathbf{B} \times \frac{\partial \mathbf{E}}{\partial t} = -\frac{\partial}{\partial t} \left(\mathbf{E} \times \mathbf{B}\right) + \mathbf{E} \times \frac{\partial \mathbf{B}}{\partial t} \tag{B.12a}$$

$$= -\frac{\partial}{\partial t} \left(\mathbf{E} \times \mathbf{B}\right) - \mathbf{E} \times \left(\nabla \times \mathbf{E}\right) \ . \tag{B.12b}$$

Now we can take this result and use $c^2 \mathbf{B}(\nabla \cdot \mathbf{B}) = 0$ to change the terms in brackets to:

$$\begin{aligned} \left[\rho_c \mathbf{E} + \mathbf{j} \times \mathbf{B}\right] = & \\ & \varepsilon_o \left\{ \left[\mathbf{E} \left(\nabla \cdot \mathbf{E}\right) + c^2 \mathbf{B} \left(\nabla \cdot \mathbf{B}\right)\right] - \left[\mathbf{E} \times \left(\nabla \times \mathbf{E}\right) + c^2 \mathbf{B} \times \left(\nabla \times \mathbf{B}\right)\right] \right\} \\ & - \varepsilon_o \frac{\partial}{\partial t} \left(\mathbf{E} \times \mathbf{B}\right) \ . \end{aligned} \tag{B.13}$$

We can simplify this slightly by recognizing the following rule about the divergence of a 2nd rank tensor:

$$\left[\mathbf{A} \left(\nabla \cdot \mathbf{A}\right) - \mathbf{A} \times \left(\nabla \times \mathbf{A}\right)\right]_{\alpha} = \sum_{\beta} \frac{\partial}{\partial x^{\beta}} \left[A_{\alpha} A_{\beta} - \frac{\mathbf{A} \cdot \mathbf{A}}{2} \delta_{\alpha\beta}\right] \tag{B.14}$$

where we have used tensor notation[24], therefore, we can rewrite Equation B.10 as:

$$\begin{aligned} \frac{dP_{mech,\alpha}}{dt} + \frac{d}{dt} \int_{V} d^3x \ \varepsilon_o \left(\mathbf{E} \times \mathbf{B}\right)_{\alpha} = & \\ & \varepsilon_o \int_{V} d^3x \ \left\{ \left[\sum_{\beta} \frac{\partial}{\partial x^{\beta}} \left(E_{\alpha} E_{\beta} - \frac{\mathbf{E} \cdot \mathbf{E}}{2} \delta_{\alpha\beta} \right) \right] + \left[c^2 \sum_{\beta} \frac{\partial}{\partial x^{\beta}} \left(B_{\alpha} B_{\beta} - \frac{\mathbf{B} \cdot \mathbf{B}}{2} \delta_{\alpha\beta} \right) \right] \right\} \ . \end{aligned} \tag{B.15}$$

At this point, we can define the *Maxwell stress tensor*, $T_{\alpha\beta}$, as:

$$T_{\alpha\beta} = \varepsilon_o \left[\left(E_{\alpha} E_{\beta} + B_{\alpha} B_{\beta}\right) - \frac{1}{2} \left(\mathbf{E} \cdot \mathbf{E} + c^2 \mathbf{B} \cdot \mathbf{B}\right) \delta_{\alpha\beta} \right] \tag{B.16}$$

[24]Meaning, $\partial/\partial x^{\beta} \equiv \partial_{\beta} = \left\{\partial/\partial x^0, \nabla\right\}$, where $x^{\beta} \equiv$ *contravariant vector* (or rank one tensor) and $x_{\beta} \equiv$ *covariant vector* [e.g., see Chapter 11 of *Jackson*, 1998].

which shows us that Equation B.15 can be rewritten, in component form, as:

$$\frac{d}{dt}\left(\mathbf{P}_{mech} + \mathbf{P}_{EM}\right)_{\alpha} = \sum_{\beta} \int_{V} d^{3}x \ \frac{\partial}{\partial x^{\beta}} T_{\alpha\beta} \tag{B.17a}$$

$$= \oint_{S} dA \ \sum_{\beta} T_{\alpha\beta} \ n_{\beta} \ . \tag{B.17b}$$

We should note that $T_{\alpha\beta}$ is not the rank two antisymmetric field-strength tensor $F^{\alpha\beta} = \partial^{\alpha} A^{\beta}$ - $\partial^{\beta} A^{\alpha}$ (where A^{α} is the 4-vector electromagnetic potentials) and $\partial^{\beta} A^{\lambda}$ = -$F^{\lambda\beta}$ + $\partial^{\lambda} A^{\beta}$, but the two are related through:

$$T^{\alpha\beta} = \frac{1}{4\pi} \left\{ \left[g^{\alpha\mu} F_{\mu\lambda} F^{\lambda\beta} + \frac{1}{4} g^{\alpha\beta} F_{\mu\nu} F^{\mu\nu} \right] - g^{\alpha\mu} F_{\mu\lambda} \partial^{\lambda} A^{\beta} \right\} \ . \tag{B.18}$$

For more information, see Chapters 6, 11, and 12 of *Jackson* [1998].

B.2 Lorentz Transformation

The general Lorentz transformation for electromagnetic fields is derived from the properties of the second-rank antisymmetric field-strength tensor given by:

$$F^{\alpha\beta} = \partial^{\alpha} A^{\beta} - \partial^{\beta} A^{\alpha} \tag{B.19}$$

where A^{α} [= (Φ,**A**)] is the scalar and vector potentials. The covariant form of the inhomogeneous Maxwell's equations are given by:

$$\partial_{\alpha} F^{\alpha\beta} = \frac{4\pi}{c} J^{\beta} \tag{B.20}$$

where J^{α} [= (cρ,**J**)] is the 4-vector current density. The covariant form of the homogeneous Maxwell's equations are given by:

$$\partial_{\alpha} \mathcal{F}^{\alpha\beta} = 0 \tag{B.21a}$$
$$\mathcal{F}^{\alpha\beta} = \frac{1}{2} \epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta} \tag{B.21b}$$

where $\epsilon^{\alpha\beta\gamma\delta}$ is defined on page 556 of *Jackson* [1998]. Note that the Lorentz force can be written in its covariant form as:

$$\frac{dp^{\alpha}}{d\tau} = m \frac{dU^{\alpha}}{d\tau} = \frac{q}{c} F^{\alpha\beta} U_{\beta} \tag{B.22}$$

which means that the equations of motion are given by:

$$\frac{1}{c} \frac{d\mathcal{E}}{d\tau} = \frac{q}{c} \mathbf{U} \cdot \mathbf{E} \tag{B.23a}$$
$$\frac{d\mathbf{p}}{d\tau} = \frac{q}{c} \left(\frac{\mathcal{E}}{mc} \mathbf{E} + \mathbf{U} \times \mathbf{B} \right) \tag{B.23b}$$

where $\mathcal{E}$ is the scalar energy, (U_{o},**U**) is the 4-vector velocity [$U_{o} = \mathcal{E}$/mc], (p_{o},**p**) is the 4-vector momentum [= m (U_{o},**U**)], and **E** and **B** are the electric and magnetic fields. Because **E** and **B** are elements of the second-rank tensor $F^{\alpha\beta}$, their values can be expressed in different reference frames in terms of their values in other reference frames [see page 558 of *Jackson*, 1998]. Mathematically, this is expressed as:

$$F'^{\alpha\beta} = \frac{\partial x'^{\alpha}}{\partial x^{\gamma}} \frac{\partial x'^{\beta}}{\partial x^{\delta}} F^{\gamma\delta} \ . \tag{B.24}$$

For a general Lorentz transformation, **E** and **B** are transformed as:

$$\mathbf{E}' = \gamma \left(\mathbf{E} + \boldsymbol{\beta} \times \mathbf{B} \right) - \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta} \left(\boldsymbol{\beta} \cdot \mathbf{E} \right) \tag{B.25a}$$

$$\mathbf{B}' = \gamma \left(\mathbf{B} - \boldsymbol{\beta} \times \mathbf{E} \right) - \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta} \left(\boldsymbol{\beta} \cdot \mathbf{B} \right) \tag{B.25b}$$

where $\boldsymbol{\beta} = \mathbf{v}/c$ [see page 558 of *Jackson*, 1998]. As an aside, if we have **J**' = σ **E**' (Ohm's law), where primes denote the rest frame, then the covariant generalization of Ohm's law is written as:

$$J^{\alpha} - \frac{1}{c^2} \left(U_{\beta} J^{\beta} \right) U^{\alpha} = \frac{\sigma}{c} F^{\alpha\beta} U_{\beta} \tag{B.26}$$

where we have allowed for the possibility of convection currents and conduction currents [see problem 11.16 of *Jackson*, 1998].

C Wave Properties

Before we begin, we should define some terms and parameters/functions that will be used later:

1. **Wave Number:** $\equiv$ effectively the number of wave crests (i.e., anti-node of local maximum) per unit length $\leftrightharpoons$ "density" of waves $\rightarrow \mathbf{k} = \mathbf{k}\left(\omega, \mathbf{x}, t\right)$ in general (sometimes denoted by $\boldsymbol{\kappa}$ as in *Whitham* [1999])
2. **Wave Frequency:** $\equiv$ effectively the number of wave crests crossing position $\mathbf{x}$ per unit time $\leftrightharpoons$ "flux" of waves $\rightarrow \omega = \omega\left(\mathbf{k}, \mathbf{x}, t\right)$ in general
3. **Wave Phase:** $\equiv$ position on a wave cycle between a crest and a trough (i.e., anti-node of local minimum) $\rightarrow \phi = \phi\left(\mathbf{x}, t\right)$ in general
4. **Dispersive Wave:** $\equiv$ a propagating fluctuation where the wave frequency, $\omega\left(\mathbf{k}, \mathbf{x}, t\right)$, is nonlinearly dependent upon the wave number, $\mathbf{k}\left(\omega, \mathbf{x}, t\right) \Rightarrow$ modes with different $\mathbf{k}$ will propagate at different speeds $\Rightarrow$ modes will spread out spatially = *disperse*.
5. **Dispersion Relation:** $\equiv$ the mathematical dependence of ω on $\mathbf{k}$ (or vice versa) $\leftrightharpoons$ mathematical relationship between ω and $\mathbf{k}$
6. **[Wave] Mode:** $\equiv$ a general solution to a dispersion relation[25]

We can define an elementary solution to periodic wave equations as:

$$\psi\left(\mathbf{x}, t\right) = \mathcal{A}\, e^{i\left(\boldsymbol{\kappa} \cdot \mathbf{x} - \omega t\right)} \tag{C.1}$$

where $\mathcal{A}$ is the wave amplitude and, in general, can be a function of $\boldsymbol{\kappa}$ and/or ω, but we will assume constant for now. Let us assume that a *dispersion relation*, $\omega = \mathcal{W}\left(\boldsymbol{\kappa}, \mathbf{x}, t\right)$, exists and may be solved for positive real roots. In general, there will be multiple solutions to the dispersion relation, where each solution is referred to as different *modes*. The term in the exponent is known as the *wave phase*, given by:

$$\phi\left(\mathbf{x}, t\right) = \boldsymbol{\kappa}\left(\omega, \mathbf{x}, t\right) \cdot \mathbf{x} - \omega\left(\boldsymbol{\kappa}, \mathbf{x}, t\right)\ t + \phi_o \tag{C.2}$$

where we have used a general form for the frequency, ω, and wave number, $\boldsymbol{\kappa}$. Recall that the Fourier transform of an arbitrary function, $f(\mathbf{x}, t)$, in four-dimensions is given by:

$$f\left(\mathbf{x}, t\right) = \frac{1}{\left(2\pi\right)^2} \int_{-\infty}^{\infty} d^3\kappa\ d\omega\ \tilde{f}\left(\boldsymbol{\kappa}, \omega\right) e^{i\left(\boldsymbol{\kappa} \cdot \mathbf{x} - \omega t\right)} \tag{C.3}$$

where $\mathbf{x}(t)$ is an arbitrary position on a plane of constant phase at time t. In other words, "$\mathbf{x}$ represents the point of maximum constructive interference at time t for a wave packet centered, in Fourier space, on $\boldsymbol{\kappa}$ and ω" [page 71 of *Stix*, 1992]. Because $\phi\left(\mathbf{x}, t\right)$ results from solutions of the wave equation, its derivatives must satisfy the dispersion relation through the following:

$$-\frac{\partial \phi\left(\mathbf{x}, t\right)}{\partial t} = \mathcal{W}\left(\frac{\partial \phi\left(\mathbf{x}, t\right)}{\partial \mathbf{x}}, \mathbf{x}, t\right) \tag{C.4}$$

[25] There can be multiple *modes* for one dispersion relation

and we can see from Equation C.2 that the following is true:

$$\boldsymbol{\kappa} = \frac{\partial \phi (\mathbf{x}, t)}{\partial \mathbf{x}} \tag{C.5a}$$
$$\omega = -\frac{\partial \phi (\mathbf{x}, t)}{\partial t} \ . \tag{C.5b}$$

We also know that $\partial^2 \phi / \partial \mathbf{x} \partial t = \partial^2 \phi / \partial t \partial \mathbf{x}$ [page 119 of *Kulsrud*, 2005], therefore:

$$\frac{\partial^2 \phi}{\partial t \partial \mathbf{x}} - \frac{\partial^2 \phi}{\partial \mathbf{x} \partial t} = 0 \tag{C.6a}$$
$$= \frac{\partial \boldsymbol{\kappa}}{\partial t} - \frac{-\partial \omega}{\partial \mathbf{x}} = 0 \tag{C.6b}$$
$$= \frac{\partial \boldsymbol{\kappa}}{\partial t} + \frac{\partial \omega}{\partial \mathbf{x}} = 0 \tag{C.6c}$$
$$= \frac{\partial \boldsymbol{\kappa}}{\partial t} + \nabla \omega = 0 \ . \tag{C.6d}$$

One can see that Equation C.6d looks similar to the *continuity equation*, so long as $\boldsymbol{\kappa} \leftrightharpoons$ "density" of the waves, and $\omega \leftrightharpoons$ "flux" of the waves.

From the above relations, we can see that on *contours* of constant $\phi (\mathbf{x}, t)$, we are sitting on local wave crests (i.e., *phase fronts*) where $\boldsymbol{\kappa}$ is orthogonal to these *contours*. These phase fronts move parallel to $\boldsymbol{\kappa}$ at a speed, $\mathbf{V}_{\phi}$, known as the *phase velocity*. The general form for this speed is given by:

$$\mathbf{V}_{\phi} \equiv \frac{\mathcal{W} (\boldsymbol{\kappa}, \mathbf{x}, t)}{\kappa} \hat{\boldsymbol{\kappa}} \ . \tag{C.7}$$

If we multiply the 2nd term in Equation C.6c by unity, we get:

$$\frac{\partial \boldsymbol{\kappa}}{\partial t} + \frac{\partial \omega}{\partial \mathbf{x}} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\kappa}} = 0 \tag{C.8a}$$
$$\frac{\partial \boldsymbol{\kappa}}{\partial t} + \frac{\partial \omega}{\partial \boldsymbol{\kappa}} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} = 0 \tag{C.8b}$$
$$\frac{\partial \boldsymbol{\kappa}}{\partial t} + (\mathbf{V}_{g} \cdot \nabla) \boldsymbol{\kappa} = 0 \tag{C.8c}$$

where $\mathbf{V}_{g}$ is called the *group velocity*, where we note that:

$$\frac{\partial \omega}{\partial \mathbf{x}} = \frac{\partial \mathcal{W} (\boldsymbol{\kappa}, \mathbf{x}, t)}{\partial \boldsymbol{\kappa}} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} + \frac{\partial \mathcal{W} (\boldsymbol{\kappa}, \mathbf{x}, t)}{\partial \mathbf{x}} \tag{C.9}$$

which shows that $\partial \mathcal{W} / \partial \boldsymbol{\kappa} = (\partial \omega / \partial \boldsymbol{\kappa})_{\mathbf{x}} \Rightarrow$ *different* $\boldsymbol{\kappa}$*'s propagate with velocity* $\mathbf{V}_{g}$ [page 376 of *Whitham*, 1999]. In other words, $\mathbf{V}_{g}$ *is the propagation velocity for* κ [page 380 of *Whitham*, 1999] and $|\mathcal{A}|^2$ *propagates with velocity* $\mathbf{V}_{g}$ [page 379 of *Whitham*, 1999].

We wish to define the term *dispersive* more appropriately, so we require the following constraints:

$$\mathcal{W}\left(\boldsymbol{\kappa}, \mathbf{x}, t\right) = \text{ real and } \neq 0 \tag{C.10a}$$

$$\frac{\partial^2 \, \mathcal{W}\left(\boldsymbol{\kappa}, \mathbf{x}, t\right)}{\partial \kappa^2} \neq 0 \tag{C.10b}$$

and finally:

$$\det \left| \frac{\partial^2 \, \mathcal{W}\left(\boldsymbol{\kappa}, \mathbf{x}, t\right)}{\partial \kappa_i \, \partial \kappa_j} \right| \neq 0 \ . \tag{C.11}$$

Thus, an observer moving with the phase fronts (crests) moves at $\mathbf{V}_{\phi}$, but they observe the local wave number and frequency to change in time $\Rightarrow$ neighboring phase fronts (crests) move away from the observer in this frame. In contrast, for an observer moving with $\mathbf{V}_{g}$, they observe constant local wave number and frequency (with respect to time), but phase fronts (crests) continuously move past the observer in this frame [page 377 of *Whitham*, 1999].

C.1 Inhomogeneous Media

Recall that for an arbitrary function, $\mathcal{F} = \mathcal{F}\left(t, x_1, x_2, ..., x_{n-1}, x_n\right)$, the *exact derivative* or *total derivative* is given by:

$$\frac{d\mathcal{F}}{dt} = \frac{\partial \mathcal{F}}{\partial t} + \sum_{i=1}^{n} \frac{\partial \mathcal{F}}{\partial x_i} \, \frac{dx_i}{dt} \ . \tag{C.12}$$

Let us start with the assumption of a lossless dispersion relation, $\mathcal{W}\left(\boldsymbol{\kappa}, \omega, \mathbf{x}, t\right) = \mathcal{W}\left(\boldsymbol{\kappa}, \omega, \mathbf{x}, t\right) = 0$, where $\mathcal{W} = 0$ for all points along a trajectory following $\mathbf{V}_{g}$. The variation of the dispersion relation, $\delta\mathcal{W}$, is given by:

$$\delta\mathcal{W}\left(\boldsymbol{\kappa}, \omega, \mathbf{x}, t\right) = \frac{\partial \mathcal{W}}{\partial t}\delta t + \frac{\partial \mathcal{W}}{\partial \mathbf{x}} \cdot \delta \mathbf{x} + \frac{\partial \mathcal{W}}{\partial \omega}\delta \omega + \frac{\partial \mathcal{W}}{\partial \boldsymbol{\kappa}} \cdot \delta \boldsymbol{\kappa} = 0 \tag{C.13}$$

which we require to $= 0$ as well[26]. Now if we let τ be a measure of distance along this trajectory and we vary the parameters with respect to τ (e.g., $\delta\boldsymbol{\kappa} \rightarrow d\boldsymbol{\kappa}/d\tau \ \delta\tau$), then we find the following relations

[26] this is actually the constraint that makes $\mathcal{W} \rightarrow 0$ for all points along the $\mathbf{V}_{g}$-trajectory

that cause the terms in Equation C.13 to cancel accordingly:

$$\frac{d\mathbf{x}}{d\tau} = \frac{\partial \mathcal{W}}{\partial \boldsymbol{\kappa}} \tag{C.14a}$$
$$\frac{d\boldsymbol{\kappa}}{d\tau} = -\frac{\partial \mathcal{W}}{\partial \mathbf{x}} \tag{C.14b}$$
$$\frac{dt}{d\tau} = -\frac{\partial \mathcal{W}}{\partial \omega} \tag{C.14c}$$
$$\frac{d\omega}{d\tau} = \frac{\partial \mathcal{W}}{\partial t} \tag{C.14d}$$

therefore,

$$\delta\mathbf{x} \rightarrow \frac{\partial \mathcal{W}}{\partial \boldsymbol{\kappa}}\delta\tau \tag{C.14e}$$
$$\delta\boldsymbol{\kappa} \rightarrow \frac{\partial \mathcal{W}}{\partial \mathbf{x}}\delta\tau \tag{C.14f}$$
$$\delta t \rightarrow \frac{\partial \mathcal{W}}{\partial \omega}\delta\tau \tag{C.14g}$$
$$\delta\omega \rightarrow \frac{\partial \mathcal{W}}{\partial t}\delta\tau \tag{C.14h}$$

which shows us that the 2nd and 4th terms and 1st and 3rd terms in Equation C.13 cancel, respectively [Chapter 4.7 of *Stix*, 1992]. Combining Equations C.14a and C.14c, we find:

$$\frac{d\mathbf{x}}{dt} = -\frac{\partial \mathcal{W}/\partial \boldsymbol{\kappa}}{\partial \mathcal{W}/\partial \omega} = \frac{\partial \omega}{\partial \boldsymbol{\kappa}} \equiv \mathbf{V}_g \ . \tag{C.15}$$

Now let us consider the dispersion relation, $\omega = \mathcal{W}\left(\boldsymbol{\kappa}, \mathbf{x}, t\right)$, then variation can be shown to be:

$$\delta\mathcal{W}\left(\boldsymbol{\kappa}, \mathbf{x}, t\right) = \frac{\partial \mathcal{W}}{\partial t}\delta t + \frac{\partial \mathcal{W}}{\partial \mathbf{x}} \cdot \delta\mathbf{x} + \frac{\partial \mathcal{W}}{\partial \boldsymbol{\kappa}} \cdot \delta\boldsymbol{\kappa} = 0 \tag{C.16}$$

which gives us the following relationships:

$$\frac{d\mathbf{x}}{dt} = \frac{\partial \mathcal{W}}{\partial \boldsymbol{\kappa}} = \frac{\partial \omega}{\partial \boldsymbol{\kappa}} \tag{C.17a}$$
$$\frac{d\boldsymbol{\kappa}}{dt} = -\frac{\partial \mathcal{W}}{\partial \mathbf{x}} = -\frac{\partial \omega}{\partial \mathbf{x}} \tag{C.17b}$$
$$\frac{d\omega}{dt} = \frac{\partial \mathcal{W}}{\partial t} = \frac{\partial \omega}{\partial t} \ . \tag{C.17c}$$

We can show this by considering that the total derivative of $\boldsymbol{\kappa}$ is given by:

$$
\begin{aligned}
\boldsymbol{\kappa} &= \frac{\partial \boldsymbol{\kappa}}{\partial t} dt + \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} \cdot d\mathbf{x} \\
\frac{d\boldsymbol{\kappa}}{dt} &= \frac{\partial \boldsymbol{\kappa}}{\partial t} + \mathbf{V}_{g} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} && \text{(C.18a)} \\
&= -\left(\frac{\partial \omega}{\partial \mathbf{x}} \right)_{t} + \mathbf{V}_{g} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} && \text{(C.18b)} \\
&= -\left\{ \frac{\partial \mathcal{W}}{\partial \boldsymbol{\kappa}} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} + \frac{\partial \mathcal{W}}{\partial \mathbf{x}} \right\} + \mathbf{V}_{g} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} && \text{(C.18c)} \\
&= -\left\{ \mathbf{V}_{g} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} + \frac{\partial \mathcal{W}}{\partial \mathbf{x}} \right\} + \mathbf{V}_{g} \cdot \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{x}} && \text{(C.18d)}
\end{aligned}
$$

therefore,

$$
\frac{d\boldsymbol{\kappa}}{dt} = -\frac{\partial \mathcal{W}}{\partial \mathbf{x}} \equiv -\left(\frac{\partial \omega}{\partial \mathbf{x}} \right)_{\boldsymbol{\kappa}} \quad \text{(C.18e)}
$$

where the notation $()_{\alpha}$ considers the expression within the parentheses a constant with respect to α. Equation C.18e is known as the *wave normal equation*, or sometimes as the *eikonal equation* [Chapter 5.6 of *Kulsrud*, 2005]. This equation defines the total rate of change of $\boldsymbol{\kappa}$ for a wave packet while following the wave packet[27].

[27] this is also the wave packet's *refraction*

C.2 Anisotropic Media

In anisotropic media, the angle between $\mathbf{V}_{gr}$ and $\mathbf{V}_{ph}$, α, is given by:

$$\tan\alpha = \frac{\frac{1}{k}\left(\frac{\partial\omega}{\partial\theta}\right)_k}{\left(\frac{\partial\omega}{\partial k}\right)_\theta} \tag{C.19a}$$

$$= \frac{\frac{1}{n}\left(\frac{\partial\omega}{\partial\theta}\right)_n}{\left(\frac{\partial\omega}{\partial n}\right)_\theta} \tag{C.19b}$$

$$= -\frac{1}{n}\left(\frac{\partial n}{\partial\theta}\right)_\omega \ . \tag{C.19c}$$

Often times it is useful to write $\mathbf{V}_{gr}$ in terms of $\mathbf{V}_{ph}$. This can be done in the following way:

$$\frac{\partial\omega}{\partial k} = \frac{c}{n} - \frac{kc}{n^2}\left(\frac{\partial n}{\partial k}\right) \tag{C.20a}$$

$$= \frac{c}{n}\left[1 - \frac{k}{n}\left(\frac{\partial\omega}{\partial k}\cdot\frac{\partial n}{\partial\omega}\right)\right] \tag{C.20b}$$

$$\frac{c}{n} = \left[1 + \frac{\omega}{n}\left(\frac{\partial n}{\partial\omega}\right)\right]\frac{\partial\omega}{\partial k} \tag{C.20c}$$

$$\frac{\partial\omega}{\partial k} = \frac{c}{n + \omega\frac{\partial n}{\partial\omega}} \tag{C.20d}$$

$$\frac{\partial\omega}{\partial k} = V_{gr} = \frac{V_{ph}}{\frac{c}{V_{ph}} + c\omega\frac{\partial {V_{ph}}^{-1}}{\partial\omega}} \tag{C.20e}$$

$$= \frac{c}{\frac{c}{V_{ph}}\left(1 - \frac{\omega}{V_{ph}}\frac{\partial V_{ph}}{\partial\omega}\right)} \tag{C.20f}$$

which gives us a simple relation between the group and phase speeds, given by:

$$\mathbf{V}_{gr} = \frac{V_{ph}}{\left(1 - \frac{\omega}{V_{ph}}\frac{\partial V_{ph}}{\partial\omega}\right)} \ . \tag{C.21}$$

C.3 Nonlinear Optics

To be more general, let us assume that a nonlinear phase, $\vartheta(\mathbf{x}, t)$, contains separable terms where one part represents the phase of the wave if the medium was uniform, stationary, and linear ($\mathbf{k}_o$, ω_o) and a second part that allows for nonlinear terms ($\varkappa$, ϖ). Under these conditions, we can see that $\omega_o = \omega_o(\mathbf{k}_o, 0)$. Therefore, we can say:

$$\vartheta(\mathbf{x}, t) = [\mathbf{k}_o \cdot \mathbf{x} - \omega_o\, t] + \varphi(\mathbf{x}, t) \tag{C.22}$$

where we now use $\varphi(\mathbf{x}, t)$ to describe the nonlinear part of the wave phase [Chapter 15 of *Sagdeev et al.*, 1988]. From this, we can see that the total wave number, $\boldsymbol{\kappa}$, and frequency, ω, are given by:

$$\boldsymbol{\kappa} = \nabla\vartheta(\mathbf{x}, t) \tag{C.23a}$$
$$= \nabla(\mathbf{k}_o \cdot \mathbf{x}) - \nabla(\omega_o\, t) + \nabla\varphi(\mathbf{x}, t) \tag{C.23b}$$
$$= \nabla(\mathbf{k}_o \cdot \mathbf{x}) + \nabla\varphi(\mathbf{x}, t) \tag{C.23c}$$

and

$$\omega = -\partial_t\, \vartheta(\mathbf{x}, t) \tag{C.23d}$$
$$= -\partial_t(\mathbf{k}_o \cdot \mathbf{x}) - \partial_t(-\omega_o\, t) - \partial_t\, \varphi(\mathbf{x}, t) \tag{C.23e}$$
$$= \partial_t(\omega_o\, t) - \partial_t\, \varphi(\mathbf{x}, t) \ . \tag{C.23f}$$

If we assume that ∂_t and ∇ acting on either $\mathbf{k}_o$ or $\omega_o \to 0$, then we have:

$$\boldsymbol{\kappa} = \nabla(\mathbf{k}_o \cdot \mathbf{x}) + \nabla\varphi(\mathbf{x}, t) \tag{C.24a}$$

and

$$\omega = \omega_o - \partial_t\, \varphi(\mathbf{x}, t) \ . \tag{C.24b}$$

To proceed further, we need to recall some rules for vector calculus. These rules are:

$$\nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\ \mathbf{B} + (\mathbf{B} \cdot \nabla)\ \mathbf{A} \tag{C.25a}$$
$$\nabla \cdot (\mathbf{A}\ \mathbf{B}) = (\nabla \cdot \mathbf{A})\ \mathbf{B} + (\mathbf{A} \cdot \nabla)\ \mathbf{B} \tag{C.25b}$$
$$\mathbf{A} \times (\nabla \times \mathbf{B}) = (\nabla\ \mathbf{B}) \cdot \mathbf{A} - (\mathbf{A} \cdot \nabla)\ \mathbf{B} \tag{C.25c}$$
$$\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A}\ (\nabla \cdot \mathbf{B}) - \mathbf{B}\ (\nabla \cdot \mathbf{A}) + (\mathbf{B} \cdot \nabla)\ \mathbf{A} - (\mathbf{A} \cdot \nabla)\ \mathbf{B} \tag{C.25d}$$

and

$$\nabla \times \mathbf{x} = 0 \tag{C.25e}$$
$$\nabla \cdot \mathbf{x} = 3 \tag{C.25f}$$
$$\nabla\ \mathbf{x} = \overleftrightarrow{\mathbb{I}} \tag{C.25g}$$

where $\overleftrightarrow{\mathbb{I}}$ is the unit dyad[28]. Using these relationships, we can show:

$$\nabla \left(\mathbf{k}_o \cdot \mathbf{x} \right) = \mathbf{k}_o \times \left(\nabla \times \mathbf{x} \right) + \mathbf{x} \times \left(\nabla \times \mathbf{k}_o \right) + \left(\mathbf{k}_o \cdot \nabla \right) \ \mathbf{x} + \left(\mathbf{x} \cdot \nabla \right) \ \mathbf{k}_o \tag{C.26a}$$
$$= 0 + \left\{ \left(\nabla \ \mathbf{k}_o \right) \cdot \mathbf{x} - \left(\mathbf{k}_o \cdot \nabla \right) \ \mathbf{x} \right\} + \left(\mathbf{k}_o \cdot \nabla \right) \ \mathbf{x} + \left(\mathbf{x} \cdot \nabla \right) \ \mathbf{k}_o \tag{C.26b}$$
$$= \left(\nabla \ \mathbf{k}_o \right) \cdot \mathbf{x} + \left(\mathbf{k}_o \cdot \nabla \right) \ \mathbf{x} \tag{C.26c}$$
$$= 0 + \left(\mathbf{k}_o \cdot \nabla \right) \ \mathbf{x} \ . \tag{C.26d}$$

The final term can be reduced even further by noticing:

$$\left(\mathbf{k}_o \cdot \nabla \right) \ \mathbf{x} = \left(\mathbf{k}_{ox} \ \partial_x + \mathbf{k}_{oy} \ \partial_y + \mathbf{k}_{oz} \ \partial_z \right) \ \left\{ x \ \mathbf{i} + y \ \mathbf{j} + z \ \mathbf{k} \right\} \tag{C.27a}$$
$$= \left(\mathbf{k}_{ox} \ \partial_x \ x \right) \mathbf{i} + \left(\mathbf{k}_{oy} \ \partial_y \ y \right) \mathbf{j} + \left(\mathbf{k}_{oz} \ \partial_z \ z \right) \mathbf{k} \tag{C.27b}$$
$$= \mathbf{k}_o \tag{C.27c}$$

which means that the final forms for $\boldsymbol{\kappa}$ and ω are given by:

$$\boldsymbol{\kappa} = \mathbf{k}_o + \nabla \varphi \left(\mathbf{x}, t \right) \tag{C.28a}$$
$$\omega = \omega_o - \partial_t \ \varphi \left(\mathbf{x}, t \right) \ . \tag{C.28b}$$

[28]often this is called the unit tensor or unit matrix or identity matrix because it satisfies the multiplication identity for rank two tensors

C.4 Doppler Effect

If we assume we are at rest with respect to a fluid moving at a velocity of $\mathbf{V}_{sw}$, then the frequency of a signal convecting with the fluid would be given by:

$$\omega_{obs} = \gamma \left(\omega_o + \mathbf{k}_o \cdot \mathbf{V}_{sw} \right) \tag{C.29}$$

where ω_{obs} is the frequency we observe, ω_o is the actual frequency of the source, γ is the relativistic factor, and $\mathbf{k}_o$ is the wave vector of the source. The relationship in Equation C.29 holds because the phase of any signal, ϕ, is a Lorentz invariant [see page 529 of *Jackson*, 1998]. This means:

$$\phi = \omega_{obs} t_{obs} - \mathbf{k}_{obs} \cdot \mathbf{x}_{obs} = \omega_o t_o - \mathbf{k}_o \cdot \mathbf{x}_o \tag{C.30}$$

where t_j is the time in the j-frame and $\mathbf{x}_j$ the position in the j-frame. The wave number is given by:

$$k_j \left(\omega_j \right) = \frac{\omega_j}{c} n \left(\omega_j \right) \tag{C.31}$$

where $n(\omega_j)$ is the wave index of refraction in the j-frame. Thus, we can show for stationary phase ($\partial \, \phi / \partial \, \omega = 0$):

$$c \frac{d \, k_j}{d \, \omega_j} = n \left(\omega_j \right) + \omega_j \frac{d \, n \left(\omega_j \right)}{d \, \omega_j} \ . \tag{C.32}$$

In general the frequency and wave number are a four vector with the following Lorentz transformations:

$$\frac{\omega'}{c} = \gamma \left(\frac{\omega}{c} - \frac{\mathbf{v}}{c} \cdot \mathbf{k} \right) \tag{C.33a}$$

$$k_{\parallel}' = \gamma \left(\frac{\mathbf{v} \cdot \mathbf{k}}{\mid \mathbf{v} \mid} - \frac{v}{c} \frac{\omega}{c} \right) \tag{C.33b}$$

$$\mathbf{k}_{\perp}' = \mathbf{k}_{\perp} \ . \tag{C.33c}$$

For an electromagnetic wave with the angle between $\mathbf{k}$ and $\mathbf{v}$ defined as θ, we have:

$$\tan \theta' = \frac{\sin \theta}{\gamma \left(\cos \theta - \frac{v}{c} \right)} \tag{C.34}$$

which shows that there exists a Doppler shift even when $\mathbf{k}$ is orthogonal to $\mathbf{v}$ (i.e. $\theta \rightarrow \pi/2$) [see page 530 of *Jackson*, 1998].

D Fluid Moment Definitions

Let us assume we have a function, $f_s(\mathbf{x},\mathbf{v},t)$, which defines the number of particles of species s in the following way:

$$dN = f_s\left(\mathbf{x}, \mathbf{v}, t\right)\ d^3x\ d^3v \tag{D.1}$$

which tells us that $f_s(\mathbf{x},\mathbf{v},t)$ is the particle distribution function of species s that defines a probability density in phase space. We can define moments of the distribution function as expectation values of any dynamical function, $g(\mathbf{x},\mathbf{v})$, as:

$$\langle g\left(\mathbf{x}, \mathbf{v}\right)\rangle = \frac{1}{N} \int d^3x\ d^3v\ g\left(\mathbf{x}, \mathbf{v}\right)\ f\left(\mathbf{x}, \mathbf{v}, t\right) \tag{D.2}$$

where $\langle\ \rangle$ is the average, which can mean ensemble average, arithmetic mean, etc.

If we define a set of fluid moments with similar format to that of Equations 2.31a and 2.31b, then we have:

$$\text{number density: } n_s = \int d^3v\ f_s\left(\mathbf{x}, \mathbf{v}, t\right) \tag{D.3a}$$

$$\text{average velocity: } \mathbf{U}_s = \frac{1}{n_s} \int d^3v\ \mathbf{v}\ f_s\left(\mathbf{x}, \mathbf{v}, t\right) \tag{D.3b}$$

$$\text{kinetic energy density: } W_s = \frac{m_s}{2} \int d^3v\ v^2\ f_s\left(\mathbf{x}, \mathbf{v}, t\right) \tag{D.3c}$$

$$\text{pressure tensor: } \overleftrightarrow{\mathbb{P}}_s = m_s \int d^3v\ \left(\mathbf{v} - \mathbf{U}_s\right)\left(\mathbf{v} - \mathbf{U}_s\right)\ f_s\left(\mathbf{x}, \mathbf{v}, t\right) \tag{D.3d}$$

$$\text{heat flux tensor: } \left(\overleftrightarrow{\mathbb{Q}}_s\right)_{i,j,k} = m_s \int d^3v\ \left(\mathbf{v} - \mathbf{U}_s\right)_i \left(\mathbf{v} - \mathbf{U}_s\right)_j \left(\mathbf{v} - \mathbf{U}_s\right)_k\ f_s\left(\mathbf{x}, \mathbf{v}, t\right) \tag{D.3e}$$

where the pressure tensor can be written as:

$$\overleftrightarrow{\mathbb{P}}_s = \begin{bmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{yx} & P_{yy} & P_{yz} \\ P_{zx} & P_{zy} & P_{zz} \end{bmatrix} \tag{D.4}$$

which can be reduced to a symmetric tensor (consequence of covariance symmetry, see Section 2.3.3) with the only off-diagonal elements being $P_{xy} = P_{yx}$, $P_{xz} = P_{zx}$, and $P_{yz} = P_{zy}$. In a magnetized plasma, the magnetic field direction can often organize the collective particle motion so that the pressure tensor is reduced to a diagonal tensor. In general, one can separate the pressure tensor into a diagonal part and an off-diagonal part[29]. The general diagonal elements of the pressure tensor are:

$$\overleftrightarrow{\mathbb{P}}_s = \begin{bmatrix} P_{\perp,1} & 0 & 0 \\ 0 & P_{\perp,2} & 0 \\ 0 & 0 & P_{\parallel} \end{bmatrix} \tag{D.5}$$

where a gyrotropic assumption will result in $\mathrm{P}_{\perp,1} = \mathrm{P}_{\perp,2}$. Thus, a gyrotropic plasma will have:

$$P_{\perp,s} = n_s k_B T_{\perp,s} \tag{D.6a}$$
$$P_{\parallel,s} = n_s k_B T_{\parallel,s} \tag{D.6b}$$

and a non-gyrotropic plasma will have:

$$T_{\perp,s} = \frac{1}{2 n_s k_B} \left(P_{\perp,1,s} + P_{\perp,2,s} \right) \tag{D.7a}$$
$$T_{\parallel,s} = \frac{1}{n_s k_B} P_{\parallel,s} \ . \tag{D.7b}$$

Therefore, if we have the following relationships:

$$\text{gyrotropic: } V_{T_s} = \sqrt{\frac{1}{2} \left({V_{T_s,\perp}}^2 + {V_{T_s,\parallel}}^2 \right)} \tag{D.8a}$$
$$\text{non-gyrotropic: } V_{T_s} = \sqrt{\frac{2}{3 m_s} Tr \left[\frac{\overleftrightarrow{\mathbb{P}}_s}{n_s k_B} \right]} \tag{D.8b}$$
$$= \sqrt{\frac{2 k_B \langle T_s \rangle}{m_s}} \tag{D.8c}$$

where we have used Tr[] as the trace and defined:

$$\langle T_s \rangle = \frac{1}{3} Tr \left[\frac{\overleftrightarrow{\mathbb{P}}_s}{n_s k_B} \right] \ . \tag{D.9}$$

The average temperature of particle species s shown in Equation D.9 is the one most often used when calculating temperatures from electrostatic plasma analyzers [e.g., *Curtis et al.*, 1989]. The temperature is physically a measure of the average kinetic energy density of particle species s, and can be represented

[29] which is usually called the stress tensor

as:

$$T_{\perp,s} = \frac{1}{2}\left(T_{\perp,1,s} + T_{\perp,2,s}\right) \tag{D.10a}$$

$$\langle T_s \rangle = \frac{1}{3}\left(T_{\perp,1,s} + T_{\perp,2,s} + T_{\parallel,s}\right) \tag{D.10b}$$

therefore, if we already have $V_{T_s,\perp}$ and $V_{T_s,\parallel}$ and we assume $T_{\perp,1} \neq T_{\perp,2}$ (i.e., non-gyrotropic)[30], then we have:

$$V_{T_s} = \sqrt{\frac{1}{3}\left({V_{T_s,\perp,1}}^2 + {V_{T_s,\perp,2}}^2 + {V_{T_s,\parallel}}^2\right)} \tag{D.11a}$$

$$= \sqrt{\frac{2{V_{T_s,\perp}}^2}{3} + \frac{{V_{T_s,\parallel}}^2}{3}} \tag{D.11b}$$

$$\neq \sqrt{\frac{1}{2}\left({V_{T_s,\perp}}^2 + {V_{T_s,\parallel}}^2\right)} \tag{D.11c}$$

The heat flux tensor (or kinetic energy flux in the bulk flow reference frame), in its general form, is a 3×3×3-element array, which, without symmetries, would have 27 distinct elements. However, due to symmetries imposed by math[31], assumptions, and physical aspects of fluids, we can reduce this tensor to only its symmetric components (10 total). The 10 variations of $Q_{l,m,n}$ are: $Q_{x,x,x}$, $Q_{x,y,y}$, $Q_{x,z,z}$, $Q_{x,x,y}$, $Q_{x,x,z}$, $Q_{x,y,z}$, $Q_{y,y,z}$, $Q_{y,z,z}$, $Q_{y,y,y}$, and $Q_{z,z,z}$. The result is a simple rank-2 tensor or 3×3 matrix where the sum of the i^{th} row results in the i^{th} component of the resultant *heat flux vector*. This is how one typically defines a heat flux in practical applications e.g., the solar wind electron heat flux, given by:

$$\vec{\mathbf{q}} = \frac{m_e}{2} \int d^3v \; f_e(\vec{\mathbf{x}}, \vec{\mathbf{v}}, t) \; (\mathbf{v} - \mathbf{U}_i) \; (\mathbf{v} - \mathbf{U}_i)^2 \tag{D.12}$$

where m_e is the electron mass, $\mathbf{U}_i$ the bulk flow velocity, and $f_e(\vec{\mathbf{v}},\vec{\mathbf{x}},t)$ represents a general form of the electron velocity distribution function.

[30] In most cases, it is assumed that the electrons are gyrotropic and ions are non-gyrotropic. Physically, this is due to the relatively long sample period ($\geq$3 s) of current particle detectors compared to ${\Omega_{ce}}^{-1}$ for electrons, which causes the resulting measured distribution to appeared *smeared out* in phase space. Most non-gyrotropic features are lost due to the relatively long sample periods. For ions, however, ${\Omega_{cp}}^{-1}$ can be ∼1-10 s (for B_o ∼ 1-10 nT). Therefore, non-gyrotropic features (e.g., gyrophase bunching) can often be observed in ion distributions.

[31] similar covariance rules to those used to make the pressure tensor symmetric

E Conservation Relations

In the case of a planar shock, we can define the conservation relations called the Rankine-Hugoniot relations across the shock ramp. If we define $\Delta[X] = \langle X \rangle_{dn}$ - $\langle X \rangle_{up}$, where the subscript up(dn) corresponds to upstream(downstream). Then we have from *Vinas and Scudder* [1986]; *Koval and Szabo* [2008]:

$$\Delta \left[G_n \right] \equiv \Delta \left[\rho \left(V_n - V_{shn} \right) \right] = 0 \tag{E.1a}$$
$$\Delta \left[B_n \right] \equiv \Delta \left[\hat{\mathbf{n}} \cdot \mathbf{B} \right] = 0 \tag{E.1b}$$
$$\Delta \left[\mathbf{S}_t \right] \equiv \Delta \left[\rho \left(V_n - V_{shn} \right) \mathbf{V}_t - \frac{B_n}{\mu_o} \mathbf{B}_t \right] = 0 \tag{E.1c}$$
$$\Delta \left[\mathbf{S}_t \right] \equiv \Delta \left[\left(\hat{\mathbf{n}} \times \mathbf{V}_t \right) B_n - \left(V_n - V_{shn} \right) \left(\hat{\mathbf{n}} \times \mathbf{B}_t \right) \right] = 0 \tag{E.1d}$$
$$\Delta \left[S_n \right] \equiv \Delta \left[P + \frac{\mathbf{B}_t \cdot \mathbf{B}_t}{2 \mu_o} + \rho \left(V_n - V_{shn} \right)^2 \right] = 0 \tag{E.1e}$$
$$\begin{aligned} \Delta \left[\varepsilon \right] \equiv \Delta \Bigg[& \rho \left(V_n - V_{shn} \right) \left\{ \frac{1}{2} \left(\mathbf{V}_{sw} - V_{shn} \hat{\mathbf{n}} \right)^2 + \frac{\gamma}{\gamma - 1} \frac{P}{\rho} + \frac{\mathbf{B} \cdot \mathbf{B}}{\rho \mu_o} \right\} \\ & - \frac{B_n \left(\mathbf{V}_{sw} - V_{shn} \hat{\mathbf{n}} \right) \cdot \mathbf{B}}{\mu_o} \Bigg] = 0 \end{aligned} \tag{E.1f}$$

where we have defined:

$$Q_n = \mathbf{Q} \cdot \hat{\mathbf{n}} \tag{E.2a}$$
$$\mathbf{Q}_t = \left(\hat{\mathbf{n}} \times \mathbf{Q} \right) \times \hat{\mathbf{n}} \tag{E.2b}$$
$$= \mathbf{Q} \cdot \left(\mathbb{I} - \hat{\mathbf{n}} \hat{\mathbf{n}} \right) \tag{E.2c}$$
$$V_{shn} = \frac{\Delta \left[\rho \mathbf{V}_{sw} \right]}{\Delta \left[\rho \right]} \cdot \hat{\mathbf{n}} \tag{E.2d}$$

and ρ is the mass density, P is scalar total (ion plus electron) thermal pressure, and γ is the ratio of specific heats or polytrope index. We note that P = $\hat{\mathbf{n}} \cdot \mathbb{P} \cdot \hat{\mathbf{n}}$ = 1/3 Tr[$\mathbb{P}$] $\sim$ n_o k_B (T_e + T_i).
The more general form of the above equations are as follows:

Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \frac{\rho_c}{\varepsilon_o} \tag{E.3a}$$
$$\nabla \cdot \mathbf{B} = 0 \tag{E.3b}$$
$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{E.3c}$$
$$\nabla \times \mathbf{B} = \mu_o \mathbf{j} + \mu_o \varepsilon_o \frac{\partial \mathbf{E}}{\partial t} \tag{E.3d}$$

mass flux continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \left(\rho \mathbf{V} \right) = 0 \tag{E.3e}$$

charge flux continuity equation:

$$\frac{\partial \rho_c}{\partial t} = -\nabla \cdot \left[e \left(n_i \mathbf{V}_i - n_e \mathbf{V}_e \right) \right] = -\nabla \cdot \mathbf{j} \tag{E.3f}$$

momentum flux continuity equation:

$$\frac{\partial}{\partial t} \left[\rho \mathbf{V} + \frac{1}{c^2} \left(\frac{\mathbf{E} \times \mathbf{B}}{\mu_o} \right) \right] + \nabla \cdot \left[\rho \mathbf{V} \mathbf{V} + \mathbb{P} + \left(\frac{\varepsilon_o \mathbf{E} \cdot \mathbf{E}}{2} + \frac{\mathbf{B} \cdot \mathbf{B}}{2\mu_o} \right) \mathbb{I} - \varepsilon_o \mathbf{E} \mathbf{E} - \frac{\mathbf{B} \mathbf{B}}{\mu_o} \right] = 0 \tag{E.3g}$$

energy flux continuity equation:

$$\frac{\partial}{\partial t} \left[\frac{1}{2} \rho \mathbf{V} \cdot \mathbf{V} + \frac{3}{2} P + \left(\frac{\varepsilon_o \mathbf{E} \cdot \mathbf{E}}{2} + \frac{\mathbf{B} \cdot \mathbf{B}}{2\mu_o} \right) \right] + \nabla \cdot \left[\left(\frac{1}{2} \rho \mathbf{V} \cdot \mathbf{V} + \frac{3}{2} P \right) \mathbf{V} + \mathbb{P} \cdot \mathbf{V} + \mathbf{q} + \left(\frac{\mathbf{E} \times \mathbf{B}}{\mu_o} \right) \right] = 0 \tag{E.3h}$$

and generalized Ohm's law:

$$\frac{\partial \mathbf{j}}{\partial t} + \nabla \cdot \sum_{\alpha} \left(q_{\alpha} n_{\alpha} \mathbf{V}_{\alpha} \mathbf{V}_{\alpha} + \frac{q_{\alpha}}{m_{\alpha}} \mathbb{P}_{\alpha} \right) = \sum_{\alpha} \frac{{q_{\alpha}}^2 n_{\alpha}}{m_{\alpha}} \left(\mathbf{E} + \mathbf{V}_{\alpha} \times \mathbf{B} \right) \tag{E.3i}$$

References

Curtis, D. W., C. W. Carlson, R. P. Lin, G. Paschmann, and H. Reme (1989), On-board data analysis techniques for space plasma particle instruments, *Rev. Sci. Instr.*, *60*, 372–380, doi:10.1063/1.1140441.

French, A. P. (1971), *Vibrations and Waves*, W. W. Norton & Company, New York, NY.

Griffiths, D. J. (1999), *Introduction to Electrodynamics (3rd Edition)*, Prentice Hall, Upper Saddle River, NJ.

Gurnett, D. A., and A. Bhattacharjee (2005), *Introduction to Plasma Physics: With Space and Laboratory Applications*, ISBN 0521364833. Cambridge, UK: Cambridge University Press.

Jackson, J. D. (1998), *Classical Electrodynamics, 3rd Edition*, John Wiley & Sons, Inc., New York, NY.

Kawano, H., and T. Higuchi (1995), The bootstrap method in space physics: Error estimation for the minimum variance analysis, *Geophys. Res. Lett.*, *22*, 307–310, doi:10.1029/94GL02969.

Khrabrov, A. V., and B. U. Ö. Sonnerup (1998), Error estimates for minimum variance analysis, *J. Geophys. Res.*, *103*, 6641–6652, doi:10.1029/97JA03731.

Koval, A., and A. Szabo (2008), Modified "Rankine-Hugoniot" shock fitting technique: Simultaneous solution for shock normal and speed, *J. Geophys. Res.*, *113*, 10,110, doi:10.1029/2008JA013337.

Kulsrud, R. M. (2005), *Plasma physics for astrophysics*, Princeton, NJ: Princeton University Press; ISBN:0-691-12073-0.

Oliver, D. (2004), *The shaggy steed of physics : mathematical beauty in the physical world*.

Sagdeev, R. Z., D. A. Usikov, and G. M. Zaslavsky (1988), *Nonlinear physics. From the pendulum to turbulence and chaos*, Chur, Switzerland: Harwood Academic Publishers, ISBN:3-7186-4829-6, ISSN:0272-2488.

Scudder, J. D., T. L. Aggson, A. Mangeney, C. Lacombe, and C. C. Harvey (1986), The resolved layer of a collisionless, high beta, supercritical, quasi-perpendicular shock wave. I - Rankine-Hugoniot geometry, currents, and stationarity, *J. Geophys. Res.*, *91*, 11,019–11,052, doi:10.1029/JA091iA10p11019.

Stix, T. H. (1992), *Waves in plasmas*, Springer.

Vinas, A. F., and J. D. Scudder (1986), Fast and optimal solution to the 'Rankine-Hugoniot problem', *J. Geophys. Res.*, *91*, 39–58, doi:10.1029/JA091iA01p00039.

Whitham, G. B. (1999), *Linear and Nonlinear Waves*, New York, NY: John Wiley & Sons, Inc.; ISBN:0-471-35942-4.