

# General Math and Useful Stuff

Lynn B. Wilson

June 17, 2009

## 1 Conversion Factors

$$1 \frac{km}{s} T = 10^{-3} \frac{mV}{m} = 1 \frac{\mu V}{m} \quad (1)$$

$$f_{pe} \cong 8.98 \sqrt{n_e(cm^{-3})} \text{ kHz} \quad (2)$$

$$f_{ce} \cong 2.80 \times 10^6 B(G) \text{ Hz} = 28.0 B(nT) \text{ Hz} \quad (3)$$

$$r_{ge} \cong 2.38 \frac{\sqrt{T_e(eV)}}{B(G)} \text{ cm} = 2.38 \times 10^5 \frac{\sqrt{T_e(eV)}}{B(nT)} \text{ cm} \quad (4)$$

$$f_{pi} \cong 210 Z \sqrt{\frac{n_i(cm^{-3})}{\mu}} \text{ Hz} [\mu = m_i/m_p] \quad (5)$$

$$f_{ci} \cong 1.52 \times 10^3 Z \frac{B(G)}{\mu} \text{ Hz} = 0.0152 Z \frac{B(nT)}{\mu} \text{ Hz} \quad (6)$$

$$r_{gi} \cong 1.02 \times 10^2 \frac{\sqrt{\mu T_i(eV)}}{ZB(G)} \text{ cm} = 1.02 \times 10^7 \frac{\sqrt{\mu T_i(eV)}}{ZB(nT)} \text{ cm} \quad (7)$$

$$\lambda_{De} \cong 7.43 \times 10^2 \sqrt{\frac{T_e(eV)}{n_e(cm^{-3})}} \text{ cm} \quad (8)$$

$$V_{Te} \cong 4.19 \times 10^7 \sqrt{T_e(eV)} \text{ cm/s} \quad (9)$$

$$V_{Ti} \cong 9.79 \times 10^5 \sqrt{\frac{T_i(eV)}{\mu}} \text{ cm/s} \quad (10)$$

$$\frac{\omega_{pe}}{\Omega_{ce}} \cong 3.21 \times 10^{-3} \frac{\sqrt{n_e(cm^{-3})}}{B(G)} = 3.21 \times 10^2 \frac{\sqrt{n_e(cm^{-3})}}{B(nT)} \quad (11)$$

$$\beta \cong 4.03 \times 10^{-11} \frac{n(cm^{-3})T(eV)}{B(G)^2} = 4.03 \times 10^{-1} \frac{n(cm^{-3})T(eV)}{B(nT)^2} \quad (12)$$

## 2 General Mathematical Rules

### 2.1 The Dirac Delta Function

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

1.  $\delta(x - a) = 0$  for  $x \neq a$
2.  $\int \delta(x - a) dx = 1$  (if region includes  $x = a$  which we'll assume from here on, otherwise it is zero)
3.  $\int dx f(x) \delta(x - a) = f(a)$
4.  $\int dx f(x) \delta'(x - a) = -f'(a)$
5. The delta function transforms according to the rule seen in Equation 20, 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 21
7. The delta function has the inverse units of whatever the delta function happens to be a function of  $\Rightarrow$  the delta function in Equation 21 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 (23a - 23d)
9. Typically one assumes that  $\nabla^2(1/r) = 0$ , assuming  $r \neq 0$  and its volume integral is equal to  $-4\pi$ . 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 22.

$$\delta(x' - x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x' - x)} \quad (13)$$

$$\frac{1}{k} \delta(k - k') = \int_0^{\infty} d\rho \rho J_\nu(k\rho) J_\nu(k'\rho) \quad (14)$$

where  $J_\nu$  are *Bessel Functions* and  $\text{Re}\{v\} > -1$ .

$$\frac{d^n \delta(x' - x)}{dx^n} = \delta(x' - x) \frac{d^n}{dx^n} \quad (15)$$

$$\delta(xa) = \frac{\delta(x)}{|a|} \quad (16)$$

$$\delta'(x' - x) = \frac{d}{dx} \delta(x' - x) = -\frac{d}{dx'} \delta(x' - x) \quad (17)$$

$$\delta(x' - x) = \frac{d}{dx} \Theta(x' - x) \quad (18)$$

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

$$\Theta(x' - x) = \begin{cases} 0 & \text{if } (x' - x) < 0, \\ 1 & \text{if } (x' - x) > 0. \end{cases} \quad (19)$$

$$\delta(f(x)) = \sum_i \frac{1}{\left| \frac{df}{dx_i} \right|} \delta(x - x_i) \quad [x_i \text{ are the zeros of } f(x)] \quad (20)$$

$$\delta(\mathbf{x} - \mathbf{x}') = \delta(x_1 - x'_1) \delta(x_2 - x'_2) \delta(x_3 - x'_3) \quad (21)$$

$$\nabla^2 \left( \frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) = -4\pi \delta(\mathbf{x} - \mathbf{x}') \quad (22)$$

$$\vec{r}_{i+\delta} = \vec{r}_i + \vec{r}_{\delta i} \quad (23a)$$

$$\frac{|\vec{r}_{\delta i}|}{|\vec{r}_{i+\delta}|} \ll 1 \quad (23b)$$

$$\delta(\vec{r} - \vec{r}_{i+\delta}) \rightarrow \delta(\vec{r} - \vec{r}_i - \vec{r}_{\delta i}) \quad (23c)$$

$$\approx \delta(\vec{r} - \vec{r}_i) - \vec{r}_{\delta i} \cdot \nabla_{\vec{r}} \left( \delta(\vec{r} - \vec{r}_i) \right) \quad (23d)$$

## 2.2 Vector and Tensor Calculus

1. Assume that the vector  $\mathbf{A}$  and the scalars,  $\psi$  and  $\phi$ , are well behaved vector functions
2.  $V \equiv$  3D volume with volume element  $d^3x$
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  $da$

$$\int_V d^3x \nabla \cdot \mathbf{A} = \int_S da \mathbf{n} \cdot \mathbf{A} \quad (24a)$$

$$\int_V d^3x \nabla \psi = \int_S da \mathbf{n} \psi \quad (24b)$$

$$\int_V d^3x \nabla \times \mathbf{A} = \int_S da \mathbf{n} \times \mathbf{A} \quad (24c)$$

$$\int_V d^3x \left[ \mathbf{A} \cdot (\nabla \times (\nabla \times \mathbf{B})) - \mathbf{B} \cdot (\nabla \times (\nabla \times \mathbf{A})) \right] = \int_S da \mathbf{n} \cdot \left[ \mathbf{B} \times (\nabla \times \mathbf{A}) - \mathbf{A} \times (\nabla \times \mathbf{B}) \right] \quad (24d)$$

$$\int_V d^3x \left( \phi \nabla^2 \psi + \nabla \phi \cdot \nabla \psi \right) = \int_S da \phi (\mathbf{n} \cdot \nabla \psi) \quad (\text{Green's 1}^{st} \text{ Identity}) \quad (24e)$$

$$\int_V d^3x \left( \phi \nabla^2 \psi + \psi \nabla^2 \phi \right) = \int_S da \phi (\phi \nabla \psi - \psi \nabla \phi) \quad (\text{Green's Theorem}) \quad (24f)$$

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

$$\int_S da (\nabla \times \mathbf{A}) \cdot \mathbf{n} = \oint_C \mathbf{A} \cdot d\mathbf{l} \quad (\text{Stokes's Theorem}) \quad (25a)$$

$$\int_S da (\mathbf{n} \times \nabla) \psi = \oint_C \psi d\mathbf{l} \quad (25b)$$

$$\int_S da (\mathbf{n} \times \nabla) \times \mathbf{A} = \oint_C d\mathbf{l} \times \mathbf{A} \quad (25c)$$

$$\int_S da \mathbf{n} \cdot (\nabla f \times \nabla g) = \oint_C dg f = - \oint_C df g \quad (25d)$$

1. In the following equations, we define  $\mathbf{x} \equiv$  coordinate of some point with respect to some origin

2.  $r \equiv$  the magnitude of  $\mathbf{x}$  ( $= |\mathbf{x}|$ )
3.  $\mathbf{k} \equiv \mathbf{x}/r =$  unit radial vector
4.  $f(r) \equiv$  a well-behaved function of  $r$
5.  $\mathbf{a} \equiv$  an arbitrary vector
6.  $\mathbf{L} \equiv$  the angular momentum operator defined in Equation 26g

$$\nabla \cdot \mathbf{x} = 3 \quad (26a)$$

$$\nabla \times \mathbf{x} = 0 \quad (26b)$$

$$\nabla \cdot [\mathbf{n}f(r)] = \frac{2}{r}f(r) + \frac{\partial f}{\partial r} \quad (26c)$$

$$\nabla \times [\mathbf{n}f(r)] = 0 \quad (26d)$$

$$(\mathbf{a} \cdot \nabla)\mathbf{n}f(r) = \frac{f(r)}{r}[\mathbf{a} - \mathbf{n}(\mathbf{a} \cdot \mathbf{n})] + \mathbf{n}(\mathbf{a} \cdot \mathbf{n})\frac{\partial f}{\partial r} \quad (26e)$$

$$\nabla(\mathbf{x} \cdot \mathbf{a}) = \mathbf{a} + \mathbf{x}(\nabla \cdot \mathbf{a}) + i(\mathbf{L} \times \mathbf{a}) \quad (26f)$$

$$\mathbf{L} = -i(\mathbf{x} \times \nabla) \quad (26g)$$

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) \quad (27a)$$

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \quad (27b)$$

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) \quad (27c)$$

$$\nabla \times \nabla \psi = 0 \quad (27d)$$

$$\nabla \cdot (\nabla \times \mathbf{a}) = 0 \quad (27e)$$

$$\nabla \times (\nabla \times \mathbf{a}) = \nabla(\nabla \cdot \mathbf{a}) - \nabla^2 \mathbf{a} \quad (27f)$$

$$\nabla \cdot (\psi \mathbf{a}) = \mathbf{a} \cdot \nabla \psi + \psi \nabla \cdot \mathbf{a} \quad (27g)$$

$$\nabla \times (\psi \mathbf{a}) = \nabla \psi \times \mathbf{a} + \psi \nabla \times \mathbf{a} \quad (27h)$$

$$\nabla(\mathbf{a} \cdot \mathbf{b}) = (\mathbf{a} \cdot \nabla)\mathbf{b} + (\mathbf{b} \cdot \nabla)\mathbf{a} + \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) \quad (27i)$$

$$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b}) \quad (27j)$$

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a}) + (\mathbf{b} \cdot \nabla)\mathbf{a} - (\mathbf{a} \cdot \nabla)\mathbf{b} \quad (27k)$$

$$(\nabla \mathbf{b}) \cdot \mathbf{a} = \mathbf{a} \times (\nabla \times \mathbf{b}) + (\mathbf{a} \cdot \nabla)\mathbf{b} \quad (27l)$$

$$\nabla^2 \mathbf{a} = \nabla(\nabla \cdot \mathbf{a}) - \nabla \times (\nabla \times \mathbf{a}) \quad (27m)$$

$$(27n)$$

$$\nabla^2 \equiv \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial u_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial}{\partial u_1} \right) + \frac{\partial}{\partial u_2} \left( \frac{h_1 h_3}{h_2} \frac{\partial}{\partial u_2} \right) + \frac{\partial}{\partial u_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial}{\partial u_3} \right) \right] \quad (28)$$

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

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	$\phi$	z	1	r	1
Spherical	z	$\theta$	$\phi$	1	r	$r \sin \phi$
Oblate Sph.	$\xi$	$\eta$	$\phi$	$a\sqrt{\sinh^2 \xi + \sin^2 \eta}$	$a\sqrt{\sinh^2 \xi + \sin^2 \eta}$	$\operatorname{acosh} \xi \cos \eta$
Elliptic Cyl.	u	v	z	$a\sqrt{\sinh^2 u + \sin^2 v}$	$a\sqrt{\sinh^2 u + \sin^2 v}$	1

4.  $h_i \equiv$  scale factors in the coordinate system metric

5.  $g_{\mu\nu} \equiv$  coordinate system metric

6.  $\Gamma_{\mu\nu}^\lambda \equiv$  Christoffel Symbol of the Second Kind

$$h_i \equiv \sqrt{g_{ii}} = \sqrt{\sum_{k=1}^n \left( \frac{\partial X_k}{\partial q_i} \right)^2} \quad (29)$$

$$g_{ij} = g_{ii} \delta_{ij} \text{ (Diagonal Metric)} \quad (30)$$

$$ds^2 = g_{11} dx_1^2 + \dots + g_{nn} dx_n^2 \quad (31a)$$

$$= h_1^2 dx_1^2 + \dots + h_n^2 dx_n^2 \quad (31b)$$

$$\nabla^2 \phi = g^{\mu\nu} \partial_\mu \partial_\nu \phi - \Gamma^\mu \partial_\nu \phi \quad (32a)$$

$$= g^{\mu\nu} \frac{\partial}{\partial x_\mu} \left( \frac{\partial \phi}{\partial x_\nu} \right) - \Gamma^\mu \frac{\partial \phi}{\partial x_\nu} \quad (32b)$$

$$\Gamma_{\mu\nu}^\lambda \equiv \frac{\partial^2 \zeta^\alpha}{\partial x^\mu \partial x^\nu} \frac{\partial x^\lambda}{\partial \zeta^\alpha} \quad (33a)$$

$$= \frac{1}{2} g^{\lambda\alpha} \left[ \partial_\nu g_{\alpha\mu} + \partial_\mu g_{\nu\alpha} - \partial_\alpha g_{\mu\nu} \right] \quad (33b)$$

$$= g^{\alpha\lambda} [\mu\nu, \lambda] \quad (33c)$$

$$\Gamma_{\lambda\mu\nu} = 0 \text{ for } \lambda \neq \mu \neq \nu \quad (34a)$$

$$\Gamma_{\lambda\lambda\nu} = -\frac{1}{2} \frac{\partial g_{\lambda\lambda}}{\partial x^\nu} \text{ for } \lambda \neq \nu \quad (34b)$$

$$\Gamma_{\lambda\mu\lambda} = \Gamma_{\mu\lambda\lambda} = \frac{1}{2} \frac{\partial g_{\lambda\lambda}}{\partial x^\mu} \quad (34c)$$

$$\Gamma_{\lambda\mu}^\nu = 0 \text{ for } \lambda \neq \mu \neq \nu \quad (34d)$$

$$\Gamma_{\lambda\lambda}^\nu = -\frac{1}{2g_{\nu\nu}} \frac{\partial g_{\lambda\lambda}}{\partial x^\nu} \text{ for } \lambda \neq \nu \quad (34e)$$

$$\Gamma_{\lambda\mu}^\lambda = \Gamma_{\mu\lambda}^\lambda = \frac{1}{2g_{\lambda\lambda}} \frac{\partial g_{\lambda\lambda}}{\partial x^\mu} = \frac{1}{2} \frac{\partial \ln g_{\lambda\lambda}}{\partial x^\mu} \quad (34f)$$

$$d\tau^2 \equiv -g_{\mu\nu} dx^\mu dx^\nu \quad (35)$$

$$\nabla_\mu V^\nu \equiv \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \quad (36)$$

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

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu_{\rho\sigma} \frac{dx^\rho}{d\lambda} \frac{dx^\sigma}{d\lambda} = 0 \quad (37)$$

### 3 Tricks and Useful Techniques

#### 3.1 Rotating Vectors

Let's assume we have two arbitrary vectors,  $\mathbf{A}$  and  $\mathbf{B}$ . Let their unit vectors be denoted by:  $\hat{a}$  and  $\hat{b}$ . If we want to find the parts of vector  $\mathbf{A}$  which are parallel and perpendicular to  $\mathbf{B}$ , 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 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  $\mathbf{B}'$  is now the  $Z'$ -Axis and  $\mathbf{A}'$  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:

$$\mathbf{a} \equiv \frac{\mathbf{A}}{|\mathbf{A}|} \quad (38a)$$

$$\mathbf{b} \equiv \frac{\mathbf{B}}{|\mathbf{B}|}, \quad (38b)$$

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

$$\mathbf{a}_\parallel = (\mathbf{a} \cdot \mathbf{b}) \mathbf{b} = |\mathbf{a}| |\mathbf{a}| \cos \theta_{ab} \mathbf{b} \quad (39a)$$

$$\mathbf{a}_\perp \equiv (\mathbf{b} \times \mathbf{a}) \times \mathbf{a} = \mathbf{a} - (\mathbf{b} \cdot \mathbf{a}) \mathbf{b}, \quad (39b)$$

which only need to be multiplied by the magnitude of the vector,  $\mathbf{A}$ , 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:

$$|\mathbf{A}| = \sqrt{(\mathbf{A}_\parallel)^2 + (\mathbf{A}_\perp)^2} \equiv \sqrt{\left( \sum_i^3 A_i^2 \right)}. \quad (40)$$

The second method to find these vectors is by constructing a matrix which can rotate both vectors into a new coordinate system where  $\mathbf{b}'$  is parallel to the new  $Z'$ -Axis and  $\mathbf{a}'$  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:

$$\mathbf{c} \equiv \mathbf{b} \times \mathbf{a} \quad (41a)$$

$$\mathbf{d} \equiv \mathbf{c} \times \mathbf{b} \quad (41b)$$

which we use to construct the following matrix:

$$\mathbf{R} = \begin{bmatrix} d_1 & d_2 & d_3 \\ c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \end{bmatrix} \quad (42)$$

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:

$$\mathbf{A} = \{0.2, 0.3, 0.4\} \quad (43a)$$

$$\mathbf{B} = \{0.1, 0.5, 0.7\} \quad (43b)$$

$$|\mathbf{A}| = 0.5385165215 \quad (43c)$$

$$|\mathbf{B}| = 0.8660253882 \quad (43d)$$

$$\mathbf{a} = \{0.37139, 0.55709, 0.74278\} \quad (43e)$$

$$\mathbf{b} = \{0.11547, 0.57735, 0.80829\} \quad (43f)$$

$$\mathbf{c} = \{-0.02144, 0.21442, -0.15010\} \quad (43g)$$

$$\mathbf{d} = \{0.25997, 1.30385 \times 10^{-8}, -0.03714\} \quad (43h)$$

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

$$\mathbf{R} = \begin{bmatrix} 0.25997 & 1.3038 \times 10^{-8} & -0.03714 \\ -0.02144 & 0.21442 & -0.15010 \\ 0.11547 & 0.57735 & 0.80829 \end{bmatrix} \quad (44)$$

which produces the following new vectors:

$$\mathbf{a}' = \{0.06897, -1.84871 \times 10^{-9}, 0.964901\} \quad (45a)$$

$$\mathbf{b}' = \{-1.87482 \times 10^{-9}, -7.16156 \times 10^{-9}, 1.00000\} . \quad (45b)$$

One can see that  $\mathbf{a}'$  and  $\mathbf{b}'$  are not normalized, nor are they what we *expected* them to be. Meaning, I claimed that  $\mathbf{b}'$  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:

$$\mathbf{a}' = \{0.07129, -1.91108 \times 10^{-9}, 0.99746\} \quad (46a)$$

$$\mathbf{b}' = \{-1.87482 \times 10^{-9}, -7.16156 \times 10^{-9}, 1.00000\} . \quad (46b)$$

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 45a and 45b and intentionally force those *small* rounding errors to zero before we renormalize the unit vectors. Let's define these new ones as  $\mathbf{w}'$  and  $\mathbf{u}'$  to avoid confusion with our vectors in Equations 46a and 46b, and they become (after renormalizing):

$$\mathbf{w}' = \{0.07129, 0.00000, 0.99746\} \quad (47a)$$

$$\mathbf{u}' = \{0.00000, 0.00000, 1.00000\} . \quad (47b)$$

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

$$|\mathbf{A}| * \mathbf{a}' \equiv \mathbf{A}' = \{0.0383921, -1.02915 \times 10^{-9}, 0.537146\} \quad (48a)$$

$$|\mathbf{B}| * \mathbf{b}' \equiv \mathbf{B}' = \{-1.62364 \times 10^{-9}, -6.20209 \times 10^{-9}, 0.866025\} \quad (48b)$$

$$|\mathbf{A}| * \mathbf{w}' \equiv \mathbf{W}' = \{0.0383921, 0.00000, 0.537146\} \quad (48c)$$

$$|\mathbf{B}| * \mathbf{u}' \equiv \mathbf{U}' = \{0.00000, 0.00000, 0.866025\} . \quad (48d)$$

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

$$\mathbf{R}_d = \begin{bmatrix} 0.25997347 & -6.5919492 \times 10^{-17} & -0.037139068 \\ -0.021442251 & 0.21442251 & -0.15009575 \\ 0.11547005 & 0.57735027 & 0.80829038 \end{bmatrix} \quad (49)$$

which produces the following new vectors (after normalization):

$$\mathbf{a}'_d = \{0.071292300580, 9.63019443558 \times 10^{-18}, 0.997455466614\} \quad (50a)$$

$$\mathbf{b}'_d = \{-3.88116288919 \times 10^{-18}, 1.40180631841 \times 10^{-18}, 1.000000000000\} . \quad (50b)$$

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

$$\mathbf{a}'_d = \{0.071292300580, 0.000000000000, 0.997455466614\} \quad (51a)$$

$$\mathbf{b}'_d = \{0.000000000000, 0.000000000000, 1.000000000000\} . \quad (51b)$$

## 4 Linear Algebra

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

$$\langle \mathbf{x} \rangle \equiv \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad (52)$$

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

$$\hat{\mathbf{x}}_k \equiv \mathbf{x}_k - \langle \mathbf{x} \rangle \quad (53)$$

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

$$\mathbf{B} = [\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_2 \dots \hat{\mathbf{x}}_n] . \quad (54)$$

The *sample covariance matrix* is thus defined by:

$$\mathbf{S} \equiv \frac{\mathbf{B}\mathbf{B}^T}{N-1} . \quad (55)$$

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:

$$\{TotalVariance\} \equiv Tr[\mathbf{S}] \quad (56)$$

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

### 4.1 Principle Component Analysis

The main goal here is to find an orthogonal  $n \times n$  matrix,  $\mathbf{P} = [\mathbf{u}_1 \dots \mathbf{u}_n]$ , 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:

$$\mathbf{Y}_k = \mathbf{P}^{-1} \mathbf{X}_k = \mathbf{P}^T \mathbf{X}_k \text{ for } k = 1, \dots, N . \quad (57)$$

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

$$\mathbf{S}_2 = \mathbf{P}^T \mathbf{S} \mathbf{P} \quad (58)$$



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  $\mathbf{S}$ ,  $\lambda_k$  on the diagonal arranged so that  $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq 0$ , then if  $\mathbf{P}$  is an orthogonal matrix of corresponding eigenvectors we have:

$$\mathbf{S} = \mathbf{PDP}^T \quad (59a)$$

$$\mathbf{D} = \mathbf{P}^T \mathbf{S} \mathbf{P} . \quad (59b)$$

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:

$$y_1 = \mathbf{u}_1^T \mathbf{X} = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \quad (60)$$

which means  $y_1$  is a linear combination of the original variables  $x_1 \dots x_n$ . 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:

$$\left\{ \text{Total Variance of } x_i \right\} = \left\{ \text{Total Variance of } y_i \right\} \quad (61a)$$

$$\left\{ \text{Total Variance of } y_i \right\} = \text{Tr}[\mathbf{D}] \quad (61b)$$

$$= \lambda_1 + \dots + \lambda_n \quad (61c)$$

$\Rightarrow$  where the variance of  $y_i$  is  $\lambda_i$ , and  $\lambda_i/\text{Tr}[\mathbf{S}]$  measures the fraction of the total variance 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  $\lambda_1$  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  $y_2$ .

## 4.2 Minimum Variance Analysis

Minimum variance analysis 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 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:

$$\mathbf{S}_{pq} = \left\langle \left( B_p - \langle B_p \rangle \right) \left( B_q - \langle B_q \rangle \right) \right\rangle \quad (62)$$

where  $\langle B_p \rangle$  is the average of the  $p^{\text{th}}$  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  $\lambda_3$  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]:

$$\delta \hat{\mathbf{k}} = \pm \left( \hat{\mathbf{e}}_1 \sqrt{\frac{\delta \lambda_3}{\lambda_1 - \lambda_3}} + \hat{\mathbf{e}}_2 \sqrt{\frac{\delta \lambda_3}{\lambda_2 - \lambda_3}} \right) \quad (63)$$

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

$$\delta \lambda_3 = \pm \lambda_3 \sqrt{\frac{2}{(K-1)}} . \quad (64)$$

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

$$\delta\lambda_i = \pm \sqrt{\frac{2\lambda_3(\lambda_i - \lambda_3)}{(K-1)}}. \quad (65)$$

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

$$\delta\theta_{kB} = \pm \sqrt{\frac{\lambda_3\lambda_2}{(K-1)(\lambda_2 - \lambda_3)^2}}. \quad (66)$$

*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:  $\lambda_1, \lambda_2, \lambda_3$
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  $\lambda_3$  can be said to be entirely due to noise
6. **Isotropic Gaussian Noise:** Equations 71 and 72 implicitly assume isotropic Gaussian noise
7. In Equation 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 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,  $B_n = \mathbf{n} \cdot \mathbf{B} = \text{constant!}$  So we vary the B-field in each of it's component directions and the variance is described by Equation 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 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:

$$\{f_j\} = \{f(x_j)\} \quad (67)$$

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

$$\langle f \rangle = \frac{1}{N} \sum_j f(x_j) \quad (68a)$$

$$= \frac{1}{N} \sum_j f(\langle x \rangle + e_j) \quad (68b)$$

$$= f(\langle x \rangle) + \frac{1}{N} f'(\langle x \rangle) \sum_j e_j + \frac{1}{2N} f''(\langle x \rangle) \sum_j (e_j)^2 + \dots \quad (68c)$$

$$= f(\langle x \rangle) + \frac{\sigma^2}{2} f''(\langle x \rangle) \quad (68d)$$

where  $\sigma^2$  is defined by:

$$\sigma^2 \equiv \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \langle \mathbf{x} \rangle)^2 \quad (69)$$

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

$$\sigma_{ij} = \sqrt{\langle \langle (\Delta \lambda_i)^2 \rangle \rangle + \langle \langle (\Delta \lambda_j)^2 \rangle \rangle} \quad (70)$$

where,

$$\langle \langle (\Delta \lambda_i)^2 \rangle \rangle = \frac{2\lambda_3(2\lambda_i - \lambda_3)}{(K-1)}. \quad (71)$$

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

$$\Delta \phi_{1j} = \sqrt{\frac{\lambda_j \lambda_1}{(K-1)(\lambda_1 - \lambda_j)}} \quad (72)$$

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

$$Var(\mathbf{B} \cdot \mathbf{x}) \equiv \frac{1}{K} \sum_{k=1}^K [(\mathbf{B}^{(k)} - \langle \mathbf{B} \rangle) \cdot \mathbf{x}]^2 \equiv \langle [(\mathbf{B}^{(k)} - \langle \mathbf{B} \rangle) \cdot \mathbf{x}]^2 \rangle \quad (73)$$

$$M_{ij} = \langle (B_i^{(k)} - \langle B_i^{(k)} \rangle) (B_j^{(k)} - \langle B_j^{(k)} \rangle) \rangle \equiv \langle \delta B_i^{(k)} \delta B_j^{(k)} \rangle \quad (74)$$

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  $\mathbf{B}^{*(k)}$ ,  $k = 1, 2, \dots, K$  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,  $\mathbf{b}^{(k)}$ ,  $k = 1, 2, \dots, K$ , in the realization. By definition, the latter has the property:

$$\langle \langle \mathbf{b}^{(k)} \rangle \rangle \equiv 0 \Rightarrow \langle \langle \delta \mathbf{b}^{(k)} \rangle \rangle \equiv 0 \quad (75)$$

which allows us to define the following:

$$\delta \mathbf{B}^{(k)} = \delta \mathbf{B}^{*(k)} + \delta \mathbf{b}^{(k)} \quad (76)$$

where

$$\delta \mathbf{B}^{*(k)} \equiv \mathbf{B}^{*(k)} - \langle \mathbf{B}^* \rangle \quad (77a)$$

$$\delta \mathbf{b}^{(k)} \equiv \mathbf{b}^{(k)} - \langle \mathbf{b} \rangle \quad (77b)$$

so that Equation 74 goes to:

$$\langle \delta B_i^{(k)} \delta B_j^{(k)} \rangle = \langle (\delta B_i^{*(k)} + \delta b_i^{(k)}) (\delta B_j^{*(k)} + \delta b_j^{(k)}) \rangle \quad (78a)$$

$$= \langle \delta B_i^{*(k)} (\delta B_j^{*(k)} + \delta b_j^{(k)}) \rangle + \langle \delta b_i^{(k)} (\delta B_j^{*(k)} + \delta b_j^{(k)}) \rangle \quad (78b)$$

$$= \langle \delta B_i^{*(k)} \delta B_j^{*(k)} \rangle + \langle \delta B_i^{*(k)} \delta b_j^{(k)} \rangle + \langle \delta b_i^{(k)} \delta B_j^{*(k)} \rangle + \langle \delta b_i^{(k)} \delta b_j^{(k)} \rangle = M_{ij}. \quad (78c)$$

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

$$\langle \langle \langle \langle A \rangle \rangle \rangle \rangle = \langle \langle \langle [A] \rangle \rangle \rangle. \quad (79)$$

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

$$\langle \langle \langle (\Delta M_{ij})^2 \rangle \rangle \rangle = \langle \langle \langle [M_{ij} + \langle M_{ij} \rangle]^2 \rangle \rangle \rangle \quad (80a)$$

$$= \langle \langle \langle \left\{ \frac{1}{K} \sum_k (\delta B_i^{*(k)} + \delta b_i^{(k)}) (\delta B_j^{*(k)} + \delta b_j^{(k)}) - \frac{1}{K} \sum_m (\delta B_i^{*(m)} \delta B_j^{*(m)} + \langle \delta b_i \delta b_j \rangle) \right\}^2 \rangle \rangle \rangle \quad (80b)$$

where the second term on the R.H.S. of Equation 80a is:

$$\langle \langle \langle M_{ij} \rangle \rangle \rangle = \langle \langle \langle [\langle \delta B_i^{(k)} \delta B_j^{(k)} \rangle] \rangle \rangle \rangle \quad (81a)$$

$$= \langle \langle \langle [\langle \langle \delta B_i^{(k)} \delta B_j^{(k)} \rangle \rangle] \rangle \rangle \rangle \quad (81b)$$

$$= \langle \langle \langle [\langle \langle \delta B_i^{*(k)} \delta B_j^{*(k)} \rangle \rangle] \rangle \rangle \rangle + \langle \langle \langle [\langle \langle \delta B_i^{*(k)} \delta b_j^{(k)} \rangle \rangle] \rangle \rangle \rangle + \langle \langle \langle [\langle \langle \delta b_i^{(k)} \delta B_j^{*(k)} \rangle \rangle] \rangle \rangle \rangle + \langle \langle \langle [\langle \langle \delta b_i^{(k)} \delta b_j^{(k)} \rangle \rangle] \rangle \rangle \rangle \quad (81c)$$

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 75. The first term on the R.H.S. is just defined as:

$$M_{ij}^* \equiv \langle \langle \langle \langle \delta B_i^{*(k)} \delta B_j^{*(k)} \rangle \rangle \rangle \rangle \quad (82)$$

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

$$\langle \langle \langle M_{ij} \rangle \rangle \rangle = M_{ij}^* + \langle \langle \langle \langle \delta b_i^{(k)} \delta b_j^{(k)} \rangle \rangle \rangle \rangle. \quad (83)$$

The second term on the R.H.S. of Equation 83 can be dealt with in the following manner:

$$\langle\langle\langle\langle\delta b_i^{(k)} \delta b_j^{(k)}\rangle\rangle\rangle\rangle = \langle\langle\langle\langle(b_i^{(k)} - \langle b_i^{(k)}\rangle)(b_j^{(k)} - \langle b_j^{(k)}\rangle)\rangle\rangle\rangle\rangle \quad (84a)$$

$$= \langle\langle\langle\langle\langle b_i^{(k)} b_j^{(k)}\rangle - \langle b_i^{(k)}\rangle\langle b_j^{(k)}\rangle - \langle\langle b_i^{(k)}\rangle b_j^{(k)}\rangle + \langle\langle b_i^{(k)}\rangle\langle b_j^{(k)}\rangle\rangle\rangle\rangle\rangle \quad (84b)$$

$$= \langle\langle\langle\langle\langle b_i^{(k)} b_j^{(k)}\rangle - \langle b_i^{(k)}\rangle\langle b_j^{(k)}\rangle\rangle\rangle\rangle \quad (84c)$$

$$= \langle\langle\langle\langle\langle b_i^{(k)} b_j^{(k)}\rangle\rangle\rangle\rangle - \langle\langle\langle\langle\langle b_i^{(k)}\rangle\langle b_j^{(k)}\rangle\rangle\rangle\rangle \quad (84d)$$

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

$$\Delta\phi_{i,j} = \pm \sqrt{\left(\frac{\lambda_3(\lambda_i + \lambda_j - \lambda_3)}{(K-1)(\lambda_i - \lambda_j)^2}\right)} \quad (85)$$

with an uncertainty in eigenvalues given by:

$$\Delta\lambda_i = \pm \sqrt{\left(\frac{2\lambda_3(2\lambda_i - \lambda_3)}{(K-1)}\right)} \quad (86)$$

## 5 David Oliver [*Oliver, 2004*]

### 5.1 The Action Principle

Action is mathematically defined by the integral:

$$S = \int_{t_1}^{t_2} L dt \quad (87)$$

where,  $L$  is defined as the Lagrangian<sup>1</sup>. 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:

$$H = \sum_{\alpha} \mathbf{p}_{\alpha} \cdot \dot{\mathbf{x}}_{\alpha} - L(\mathbf{x}_{\alpha}, \dot{\mathbf{x}}_{\alpha}) \quad (88)$$

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

$$\mathbf{p}_{\alpha} \equiv \frac{\partial L(\mathbf{x}_{\alpha}, \dot{\mathbf{x}}_{\alpha})}{\partial \dot{\mathbf{x}}_{\alpha}} \quad (89)$$

The kinetic energy can be defined as:

$$T \equiv \frac{1}{2} \sum_{\alpha} \frac{\partial L(\mathbf{x}_{\alpha}, \dot{\mathbf{x}}_{\alpha})}{\partial \dot{\mathbf{x}}_{\alpha}} \cdot \dot{\mathbf{x}}_{\alpha} \quad (90)$$

<sup>1</sup>Note: Oliver refers to the Lagrangian as the *the gene of motion*.

and the angular momentum is:

$$\mathbf{J} \equiv \mathbf{x}_\alpha \times \frac{\partial L(\mathbf{x}_\alpha, \dot{\mathbf{x}}_\alpha)}{\partial \dot{\mathbf{x}}_\alpha} . \quad (91)$$

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,  $f(q_1, q_2, q_3, \dots, q_n)$ , the following statement defines what it means to have *no slope* at some point,  $q_o$ , in space-time:

$$\left. \frac{\partial f}{\partial q_i} \right|_{q_i=q_o} = 0 , \quad (92)$$

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

$$\delta f = \frac{\partial f}{\partial q_i} \delta q_i = 0 , \quad (93)$$

where  $\delta q_i$  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  $\delta q_i$ ). Since the variation of the arguments,  $\delta q_i$ , 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*  $\equiv$  a quantity that has a single value corresponding to an entire function. Thus, the variation of the action is defined as:

$$\delta S \equiv \int_{t_1}^{t_2} \delta L dt = \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) dt \quad (94)$$

where the second term in the equation is defined as:

$$\frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i = \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} (\delta q_i) \quad (95a)$$

$$= \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) - \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i . \quad (95b)$$

An important thing to note is that the path variations,  $\delta q$ , vanish at the end points,  $(\delta q, t)_1$  and  $(\delta q, t)_2$ . So we look at the first term in Equation 95b and notice the following:

$$\int_{t_1}^{t_2} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) dt = \left( \frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) \Big|_{t_1}^{t_2} = 0 . \quad (96)$$

So we then have the following from Equation 95b:

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i dt = - \int_{t_1}^{t_2} \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i dt \quad (97)$$

which means we've now transformed the second term in Equation 94 into a variation of  $\delta q_i$ , instead of  $\delta \dot{q}_i$ . This implies that we can do the following:

$$\delta S = - \int_{t_1}^{t_2} \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} \right) \delta q_i dt . \quad (98)$$

## 5.2 Other important definitions

### The Total Time Derivative

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \left( \frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i \right) \quad (99a)$$

$$= \frac{\partial f}{\partial t} + \{f, \mathcal{H}\} \quad (99b)$$

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

$$\{f, g\} \equiv \left( \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i} \right). \quad (100)$$

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:

$$\dot{q}_i = \{q_i, \mathcal{H}\} \quad (101a)$$

$$\dot{p}_i = \{p_i, \mathcal{H}\} \quad (101b)$$

$$\{q_i, q_j\} = 0 \quad (101c)$$

$$\{p_i, p_j\} = 0 \quad (101d)$$

$$\{q_i, p_j\} = \delta_{ij} \quad (101e)$$

$$\{J_i, J_j\} = \varepsilon_{ijk} J_k$$

$$\begin{aligned} \mathbf{J} \cdot \{\mathbf{J}, J_j\} &= J_i \{J_i, J_j\} \\ &= \varepsilon_{ijk} J_i J_k \\ &\equiv 0 \end{aligned} \quad (102a)$$

$$\begin{aligned} \{J^2, J_j\} &= 0 \\ &= \{J_i J_i, J_j\} \\ &= 2J_i \{J_i, J_j\} \end{aligned} \quad (103a)$$

$$\{x_i, J_j\} = \varepsilon_{ijk} x_k \quad (104a)$$

$$\{p_i, J_j\} = \varepsilon_{ijk} p_k \quad (104b)$$

$$\{F_i, J_j\} = \varepsilon_{ijk} F_k \quad (104c)$$

$$\{f, J_j\} = 0 \quad (104d)$$

(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<sup>2</sup>. If we define the following as the single  $2s$  state vector of phase space:

$$\xi = (q, p) \quad (105)$$

where the first  $s$ -components of  $\xi$  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:

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \quad (106)$$

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:

$$J = -J^\dagger = -J^{-1} \quad (107)$$

where  $J^\dagger$  is defined by:

$$J_{ij}^\dagger = -J_{ji}^* \quad (108)$$

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

$$J^2 = - \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} = -J \quad (109)$$

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

$$\xi_i J_{ij} \xi_j = 0 \text{ (often seen as } \xi J \xi = 0 \text{)}. \quad (110)$$

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

$$\{f, g\} = \frac{\partial f}{\partial \xi} J \frac{\partial g}{\partial \xi} \quad (111)$$

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

$$\{\xi, f\} = J \frac{\partial f}{\partial \xi}. \quad (112)$$

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

$$\dot{\xi} = \left\{ \xi, \mathcal{H} \right\} = J \frac{\partial \mathcal{H}}{\partial \xi}. \quad (113)$$

---

<sup>2</sup>*Symplectic structure* is named after the Greek word,  $\pi\lambda\epsilon\kappa\omicron\varsigma$ , meaning *twined* or *braided*. It is an antisymmetric pairing of coordinates induced by the action principle.



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

$$\{F, G\} = \frac{\partial F}{\partial \xi} J \frac{\partial G}{\partial \xi} = \frac{\partial F}{\partial \xi} \cdot \{\xi, G\} = -\frac{\partial G}{\partial \xi} \cdot \{\xi, F\}. \quad (114)$$

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  $\{F, G\} \neq 0$ , 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  $\{F, G\} = 0$ , 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,  $F(\xi)$ , has an image in phase space as *sheets* of level surfaces filled with streamlines generated by  $\{\xi, F\} \equiv$  the flow.  $\Rightarrow$  The Poisson bracket describes the intersection of the two flows produced by  $F(\xi)$  and  $G(\xi)$ .
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  $Q = (Q_1, Q_2, \dots, Q_s)$  and  $P = (P_1, P_2, \dots, P_s)$ . Though it may appear that Q and P are actual position and momentum coordinates, they need not be. Now Hamilton's equations are:

$$\dot{Q}_i = \frac{\partial \mathcal{H}}{\partial P_i} \quad (115a)$$

$$\dot{P}_i = -\frac{\partial \mathcal{H}}{\partial Q_i} \quad (115b)$$

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

$$\delta S = \delta \int_{t_1}^{t_2} (p_i dq_i - \mathcal{H} dt) = \delta \int_{t_1}^{t_2} (P_i dQ_i - \mathcal{H}' dt). \quad (116)$$

These two integrals may differ by any function, F, which has a vanishing variation (i.e.  $\delta F = 0$ ). Thus the difference between the integrals in Equation 116 must be the total differential of F:

$$dF = p_i dq_i - P_i dQ_i - (\mathcal{H} - \mathcal{H}') dt \quad (117)$$

which defines what is referred to as *the generating function*. The generating function, from Equation 117, is  $F = F(q, Q, t)$ . The relationship of any coordinate can be described as follows:

$$p_i = \frac{\partial F}{\partial q_i} \quad (118a)$$

$$P_i = -\frac{\partial F}{\partial Q_i} \quad (118b)$$

$$(\mathcal{H} - \mathcal{H}') = \frac{\partial F}{\partial t} \quad (118c)$$

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

$$dG = d\left(F + P_i Q_i\right) = p_i dq_i + P_i dQ_i - \left(\mathcal{H} - \mathcal{H}'\right) dt . \quad (119)$$

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

$$p_i = \frac{\partial G}{\partial q_i} \quad (120a)$$

$$Q_i = \frac{\partial G}{\partial P_i} \quad (120b)$$

$$\left(\mathcal{H} - \mathcal{H}'\right) = -\frac{\partial G}{\partial t} \quad (120c)$$

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 121a and 121b.
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,  $f(\mathbf{x}(t))$  would not explicitly depend on  $t$  UNLESS  $t$  existed independent of  $\mathbf{x}(t)$  in the function.
4. *Implicit Dependence*  $\equiv$  An indirect relationship between two quantities, e.g.  $f(\mathbf{x}(t))$  implicitly depends on the variable  $t$ , but  $\mathbf{x}$  explicitly depends on  $t$

$$\frac{\partial}{\partial q_i} G(q, P, t) = p_i \quad (121a)$$

$$Q_i = \frac{\partial}{\partial P_i} G(q, P, t) \quad (121b)$$

where the new coordinates,  $Q(q, p)$ , are given **explicitly** by Equation 121b and the coordinates,  $P(q, p)$ , are given **implicitly** by Equation 121a. The two following examples illustrate some trivial transformations:

$$G(q, P, t) = q_i P_i \quad (122a)$$

$$F(q, Q, t) = q_i Q_i \quad (122b)$$

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

$$Q_i = q_i \quad (123a)$$

$$P_i = p_i \quad (123b)$$

$$\mathcal{H}' = \mathcal{H} \quad (123c)$$

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

$$Q_i = p_i \quad (124a)$$

$$P_i = -q_i \quad (124b)$$

$$\mathcal{H}' = \mathcal{H} . \quad (124c)$$

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  $P_i \equiv I_i$ , then  $\dot{I}_i = 0$  and  $Q_i \equiv \alpha_i$ ). Now Hamilton's equations simplify dramatically to:

$$\dot{\alpha}_i = \frac{\partial \mathcal{H}'}{\partial I_i} \quad (125a)$$

$$\dot{I}_i = -\frac{\partial \mathcal{H}'}{\partial \alpha_i} = 0 \quad (125b)$$

which simplifies our job of solving Hamilton's equations even more than just having Equation 125b 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  $\mathcal{H}' = \mathcal{H}'(I)$ . We may now rewrite the R.H.S. of Equation 126 as a function of only  $I$  also:

$$\omega_i(I) \equiv \frac{\partial \mathcal{H}'}{\partial I_i} . \quad (126)$$

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

$$\alpha_i = \int dt \left( \frac{\partial \mathcal{H}'}{\partial I_i} \right) = \omega_i t + \beta_i \quad (127)$$

where  $\beta_i$  ( $i = 1, 2, \dots, s$ ) are the integration constants.

1. The elementary flow ONLY "flows" along the  $\alpha$ -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  $\beta$ , which make up the two sets of  $s$  quantities.
3. We can define the *Phase Vector* (Equation 128a), the phase vector's *Phase Velocity* (Equation 128b), and the *Integration Constants* (Equation 128c).

$$\Xi = (\alpha, I) \quad (128a)$$

$$\Omega = (\omega, 0) \quad (128b)$$

$$\mathcal{I} = (\beta, I) \quad (128c)$$

Thus the elementary flow can be described by:

$$\Xi(t) = \Omega t + \mathcal{I} . \quad (129)$$

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,  $\Xi = (\alpha, I)$ , are NOT connected to their image phase space coordinates,  $\xi = (q, p)$ , 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:

$$\Delta(\mathcal{I}) = 0 . \quad (130)$$

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

$$\frac{d\mathcal{I}}{dt} = \frac{\partial \mathcal{I}}{\partial t} + \{ \mathcal{I}, \mathcal{H} \} = 0. \quad (131)$$

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:

$$\{ \mathcal{I}, \mathcal{H} \} = 0, \quad (132)$$

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:

$$dS = -\mathcal{H}dt + p_i dq_i = \frac{\partial S}{\partial t} dt + \frac{\partial S}{\partial q_i} dq_i. \quad (133)$$

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

$$\mathcal{H} = -\frac{\partial S}{\partial t} \quad (134a)$$

$$p_i = \frac{\partial S}{\partial q_i} dq_i \quad (134b)$$

and since  $\mathcal{H} = \mathcal{H}(q, p)$ , we can rewrite Equation 134a as:

$$\frac{\partial S}{\partial t} + \mathcal{H}\left(q, \frac{\partial S}{\partial q}\right) = 0. \quad (135)$$

We now know that  $S = S(q, t)$  is a solution to the 1<sup>st</sup>-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  $S(q, t)$  is a solution of the Hamilton-Jacobi equation, then  $S(q, t) + A$  is too (assuming  $A$  is an additive constant)<sup>3</sup>. 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:

$$S(q, I, t) = -\mathcal{E}t + S_o(q, I) \quad (136)$$

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

$$p_i = \frac{\partial}{\partial q_i} S_o(q, I) \quad (137a)$$

$$\alpha_i = \frac{\partial}{\partial I_i} S_o(q, I) \quad (137b)$$

$$(137c)$$

---

<sup>3</sup>So, it's not entirely clear why, but the remaining constants must invariants of motion.

or they may also be expressed as:

$$p_i = \frac{\partial}{\partial q_i} S(q, I, t) \quad (138a)$$

$$\beta_i = \frac{\partial}{\partial I_i} S(q, I, t) \quad (138b)$$

$$(138c)$$

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

$$\mathcal{H}'(I) = \mathcal{H}(q, p) . \quad (139)$$

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

$$S = \int (p_i dq_i - \mathcal{H} dt) \quad (140)$$

but this form is *clearly* not an invariant<sup>4</sup>, so we reconsider this case as a contour integral over a closed contour,  $\gamma$ , in phase space as:

$$S = \oint_{\gamma} (p_i dq_i - \mathcal{H} dt) . \quad (141)$$

One should note that the contour,  $\gamma$ , 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<sup>5</sup>. This closed integral is the Poincaré 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<sup>6</sup> involved a force corresponding to the potential:

$$V(q) = \frac{\kappa q^2}{2} , \quad (142)$$

where  $\kappa$  is a constant. It gives rise to a force, denoted by:

$$-\frac{\partial}{\partial q} V(q) = -\kappa q \equiv \mathcal{F} , \quad (143)$$

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)<sup>7</sup>. The Hamiltonian can be written as:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{\kappa q^2}{2} , \quad (144)$$

---

<sup>4</sup>Due the indefinite nature of the integral and the lack of rotational invariance in the  $p_i dq_i$  terms, the integral can't be said to be an invariant of motion.

<sup>5</sup>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

<sup>6</sup>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.*

<sup>7</sup>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.

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:

$$I = \mathcal{H} / \omega_o \quad (145a)$$

$$\alpha_i = \omega_i + \beta_i \quad (145b)$$

which gives us a new Hamiltonian,  $\mathcal{H}'(I) = \omega_o I$ , and the coordinates,  $(\alpha, I)$ . The generating function for this case is the action found in Equation 136. This can be found from solving Equation 135 for one degree of freedom, which reduces to:

$$\frac{dS_o}{dq} = p(q) \quad (146)$$

After some algebraic manipulation, the integral for  $S_o$  can be found<sup>8</sup> to be:

$$S_o(q, I) = \int dq \sqrt{2m\omega_o \left( I - \frac{m\omega_o q^2}{2} \right)}. \quad (147)$$

## References

- Gurnett, D. A., and A. Bhattacharjee (2005), *Introduction to Plasma Physics: With Space and Laboratory Applications*, ISBN 0521364833. Cambridge, UK: Cambridge University Press.
- 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.
- Khrabrov, A. V., and B. U. Ö. Sonnerup (1998), Error estimates for minimum variance analysis, *J. Geophys. Res.*, 103, 6641–6652.
- Lang, K. R. (2000a), *The Sun From Space*, Astronomy and Astrophysics Library, Springer, Verlag Berlin, Germany.
- Oliver, David, *The Shaggy Steed of Physics: Mathematical Beauty in the Physical World*, Springer-Verlag New York, Inc., New York, NY.

---

<sup>8</sup>Solve for p(q) in Equation 144 by replacing  $\kappa$  with  $m\omega_o^2$ , and  $\mathcal{H}$  with  $\mathcal{H}' = I\omega_o$ .