Math Reference

Eran Guendelman

Last revised: November 16, 2002

1 Vectors

1.1 Cross Product

1.1.1 Basic Properties

- The cross product is *not* associative

1.1.2 The "star" operator [Baraff, "Rigid Body Simulation"]:

- Let $\mathbf{u}^* = \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix}$. Then

$$\begin{aligned} \mathbf{u}^*\mathbf{v} &= \mathbf{u} \times \mathbf{v} \\ \mathbf{v}^T\mathbf{u}^* &= (\mathbf{v} \times \mathbf{u})^T \end{aligned}$$

- $(\mathbf{u}^*)^T = -\mathbf{u}^*$
- $(\mathbf{u}^*)^2 = \mathbf{u}\mathbf{u}^T - \mathbf{u}^T\mathbf{u}I$ (using $\mathbf{u} \times (\mathbf{u} \times \mathbf{v}) = \left(\mathbf{u}\mathbf{u}^T - \mathbf{u}^T\mathbf{u}I\right)\mathbf{v}$ from 1.1.4)
- If $A = (\mathbf{a}_1 \mid \mathbf{a}_2 \mid \mathbf{a}_3)$ then $\mathbf{u}^*A = (\mathbf{u} \times \mathbf{a}_1 \mid \mathbf{u} \times \mathbf{a}_2 \mid \mathbf{u} \times \mathbf{a}_3)$. We could write $\mathbf{u} \times A$ but I'm not sure how standard this notation is...
- In summation notation $(\mathbf{u}^*)_{ij} = -\varepsilon_{ijk}u_k$

1.1.3 Scalar triple product [Weisstein, Scalar triple product]:

$$\begin{aligned} [\mathbf{u}, \mathbf{v}, \mathbf{w}] &= \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) \\ &= \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} \end{aligned}$$

1.1.4 Vector triple product [Weisstein, Vector triple product]:

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w})=\mathbf{v}(\mathbf{u}\cdot\mathbf{w})-\mathbf{w}(\mathbf{u}\cdot\mathbf{v})$$

- Note that

$$\begin{aligned}\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) &= \left(\mathbf{v}\mathbf{w}^T-\mathbf{w}\mathbf{v}^T\right)\mathbf{u}\\ &= \left(\mathbf{u}^T\mathbf{w}I-\mathbf{w}\mathbf{u}^T\right)\mathbf{v}\\ &= \left(\mathbf{v}\mathbf{u}^T-\mathbf{u}^T\mathbf{v}I\right)\mathbf{w}\end{aligned}$$

- As a special case, we get

$$\begin{aligned}\mathbf{u}\times(\mathbf{v}\times\mathbf{u}) &= \mathbf{v}(\mathbf{u}\cdot\mathbf{u})-\mathbf{u}(\mathbf{u}\cdot\mathbf{v})\\ &= \left(\mathbf{u}^T\mathbf{u}I-\mathbf{u}\mathbf{u}^T\right)\mathbf{v}\end{aligned}$$

- You can also look at it as decomposing $\mathbf{v}$ into components along orthogonal vectors $\mathbf{u}$ and $\mathbf{u}\times(\mathbf{v}\times\mathbf{u})$:

$$\begin{aligned}\mathbf{v} &= \frac{\mathbf{v}\cdot\mathbf{u}}{|\mathbf{u}|^2}\mathbf{u}+\frac{\mathbf{v}\cdot(\mathbf{u}\times(\mathbf{v}\times\mathbf{u}))}{|\mathbf{u}\times(\mathbf{v}\times\mathbf{u})|^2}\mathbf{u}\times(\mathbf{v}\times\mathbf{u})\\ &= \left(\frac{\mathbf{u}^T\mathbf{v}}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{u}+\left(\frac{1}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{u}\times(\mathbf{v}\times\mathbf{u})\end{aligned}$$

  where we used $\mathbf{v}\cdot(\mathbf{u}\times(\mathbf{v}\times\mathbf{u}))=(\mathbf{v}\times\mathbf{u})\cdot(\mathbf{v}\times\mathbf{u})=|\mathbf{v}\times\mathbf{u}|^2$ and $|\mathbf{u}\times(\mathbf{v}\times\mathbf{u})|^2=|\mathbf{u}|^2|\mathbf{v}\times\mathbf{u}|^2$

- Yet another way to look at it (related to above) is to note that $P_{\mathbf{u}}(\mathbf{v})=\left(\frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{v}$ is the linear operator projecting $\mathbf{v}$ onto the vector $\mathbf{u}$. (Since $P_{\mathbf{u}}(\mathbf{v})=\frac{\mathbf{u}\mathbf{u}^T\mathbf{v}}{\mathbf{u}^T\mathbf{u}}=\left(\frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|}\right)\left(\frac{\mathbf{u}}{|\mathbf{u}|}\right)$). Then $P_{\mathbf{u}}^{\perp}(\mathbf{v})=\left(I-\frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{v}$ gives the component of $\mathbf{v}$ perpendicular to $\mathbf{u}$. Comparing this to the triple vector product we see that

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{u})=|\mathbf{u}|^2P_{\mathbf{u}}^{\perp}(\mathbf{v})$$

- Finally, we also deduce that $(\mathbf{u}^*)^2=\mathbf{u}\mathbf{u}^T-\mathbf{u}^T\mathbf{u}I$

1.1.5 Quadruple product

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) &= \mathbf{c}\cdot(\mathbf{d}\times(\mathbf{a}\times\mathbf{b}))\\ &= \mathbf{c}\cdot((\mathbf{d}\cdot\mathbf{b})\mathbf{a}-(\mathbf{d}\cdot\mathbf{a})\mathbf{b})\\ &= (\mathbf{d}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{a})-(\mathbf{d}\cdot\mathbf{a})(\mathbf{c}\cdot\mathbf{b})\end{aligned}$$

or in summation notation

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) &= (\varepsilon_{ijk}a_ib_j\mathbf{e}_k)\cdot(\varepsilon_{lmn}c_ld_m\mathbf{e}_n)\\ &= \varepsilon_{ijk}\varepsilon_{lmn}a_ib_jc_ld_m\delta_{kn}\\ &= \varepsilon_{ijk}\varepsilon_{lmk}a_ib_jc_ld_m\\ &= (\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl})\,a_ib_jc_ld_m\\ &= a_ib_jc_id_m - a_ib_jc_jd_i\end{aligned}$$

1.1.6 Derivative of cross product

- If $\mathbf{w}(\mathbf{x}) = \mathbf{u}(\mathbf{x})\times\mathbf{v}(\mathbf{x})$ then $\frac{\partial\mathbf{w}}{\partial x_i} = \mathbf{u}\times\frac{\partial\mathbf{v}}{\partial x_i} + \frac{\partial\mathbf{u}}{\partial x_i}\times\mathbf{v} = \mathbf{u}\times\frac{\partial\mathbf{v}}{\partial x_i} - \mathbf{v}\times\frac{\partial\mathbf{u}}{\partial x_i}$
- The derivative $\mathbf{Dw} = \left(\frac{\partial\mathbf{w}}{\partial x_1}\mid\frac{\partial\mathbf{w}}{\partial x_2}\mid\cdots\mid\frac{\partial\mathbf{w}}{\partial x_n}\right)$ can be expressed succinctly using the star operator as

$$\mathbf{Dw} = \mathbf{u}^*(\mathbf{Dv}) - \mathbf{v}^*(\mathbf{Du})$$

- *e.g.* $\frac{d}{d\mathbf{v}}(\mathbf{u}\times\mathbf{v}) = \mathbf{u}^*$ (as expected, since $\mathbf{u}\times\mathbf{v} = \mathbf{u}^*\mathbf{v}$)

1.1.7 Misc Cross Product Properties...

- $(\mathbf{u}\times\mathbf{a})\times(\mathbf{u}\times\mathbf{b}) = (\mathbf{u}\cdot(\mathbf{a}\times\mathbf{b}))\,\mathbf{u}$
- If $A = V\mathbf{u}^*$ with V orthonormal and $\mathbf{u}$ a unit vector, then letting $\mathbf{v}_i^T$ be the ith row of V and $\mathbf{a}_i^T$ the ith row of A we can recover $\mathbf{v}_i$ as follows:
  - $\mathbf{v}_i = -\mathbf{a}_i\times\mathbf{u} + (\mathbf{a}_j\times\mathbf{u})\times(\mathbf{a}_k\times\mathbf{u})$ where (i,j,k) are cyclic permutations of $(1,2,3)$
  - Proof:
    * $A^T = \mathbf{u}^{*T}V^T = -\mathbf{u}^*V^T$ so $\mathbf{a}_i = -\mathbf{u}\times\mathbf{v}_i = \mathbf{v}_i\times\mathbf{u}$
    * $\mathbf{a}_j\times\mathbf{a}_k = (\mathbf{v}_j\times\mathbf{u})\times(\mathbf{v}_k\times\mathbf{u}) = (\mathbf{u}\times\mathbf{v}_j)\times(\mathbf{u}\times\mathbf{v}_k) = (\mathbf{u}\cdot(\mathbf{v}_j\times\mathbf{v}_k))\,\mathbf{u} = (\mathbf{u}\cdot\mathbf{v}_i)\,\mathbf{u}$
    * $(\mathbf{a}_j\times\mathbf{u})\times(\mathbf{a}_k\times\mathbf{u}) = (\mathbf{u}\times\mathbf{a}_j)\times(\mathbf{u}\times\mathbf{a}_k) = (\mathbf{u}\cdot(\mathbf{a}_j\times\mathbf{a}_k))\,\mathbf{u} = (\mathbf{u}\cdot(\mathbf{u}\cdot\mathbf{v}_i)\,\mathbf{u})\,\mathbf{u} = (\mathbf{u}\cdot\mathbf{v}_i)(\mathbf{u}\cdot\mathbf{u})\,\mathbf{u} = (\mathbf{u}\cdot\mathbf{v}_i)\,\mathbf{u}$
    * $\mathbf{a}_i\times\mathbf{u} = (\mathbf{v}_i\times\mathbf{u})\times\mathbf{u} = \mathbf{u}\times(\mathbf{u}\times\mathbf{v}_i) = (\mathbf{u}\cdot\mathbf{v}_i)\,\mathbf{u} - (\mathbf{u}\cdot\mathbf{u})\,\mathbf{v}_i = (\mathbf{u}\cdot\mathbf{v}_i)\,\mathbf{u} - \mathbf{v}_i$
    * So $-\mathbf{a}_i\times\mathbf{u} + (\mathbf{a}_j\times\mathbf{u})\times(\mathbf{a}_k\times\mathbf{u}) = \mathbf{v}_i - (\mathbf{u}\cdot\mathbf{v}_i)\,\mathbf{u} + (\mathbf{u}\cdot\mathbf{v}_i)\,\mathbf{u} = \mathbf{v}_i$

1.2 Norms

- A norm is a function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ satisfying
  - $\|\mathbf{x}\| \geq 0$ ($\|\mathbf{x}\| = 0$ iff $\mathbf{x} = 0$)
  - $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

- $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$

- **Hölder inequality**: $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q$ where $\frac{1}{p} + \frac{1}{q} = 1$
  - **Cauchy-Schwartz inequality**: A special case for $p = q = 2$: $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$

- All norms on $\mathbb{R}^n$ are equivalent. ($\exists c_1, c_2 > 0$ such that $c_1 \|\mathbf{x}\|_\alpha \leq \|\mathbf{x}\|_\beta \leq c_2 \|\mathbf{x}\|_\alpha$)

1.3 Misc

- Orthogonal $\implies$ Linearly independent

2 Multivariable

Jacobian Matrix and Gradient

$\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ with components $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ has derivative (Jacobian) matrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

If $m = 1$ then this is simply the gradient transposed

$$(\nabla f)^T = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

- The derivative of $\mathbf{f}$ at x_0 is the linear map $D\mathbf{f}(x_0)$ which satisfies

$$\lim_{x \to x_0} \frac{|\mathbf{f}(x) - \mathbf{f}(x_0) - D\mathbf{f}(x_0)(x - x_0)|}{|x - x_0|} = 0$$

- The derivative (Jacobian) matrix is the matrix of $D\mathbf{f}(x)$ with respect to the standard bases.

Derivatives (see [Horn, Appendix])

- $\nabla(\mathbf{f} \cdot \mathbf{g}) = (D\mathbf{g})^T \mathbf{f} + (D\mathbf{f})^T \mathbf{g}$ (really abusing notation!)
- $D(\alpha \mathbf{v}) = \alpha (D\mathbf{v}) + \mathbf{v} (\nabla \alpha)^T$
- $D\left(\frac{\mathbf{v}}{\alpha}\right) = \frac{1}{\alpha}(D\mathbf{v}) + \mathbf{v}\left(-\frac{\nabla \alpha}{\alpha^2}\right)^T = \frac{\alpha(D\mathbf{v}) - \mathbf{v}(\nabla \alpha)^T}{\alpha^2}$

- $\nabla \cdot (\alpha \mathbf{v}) = \alpha \nabla \cdot \mathbf{v} + (\nabla \alpha) \cdot \mathbf{v}$
- $\frac{d}{d\mathbf{x}} \mathbf{f}(\mathbf{y}) = \left(\frac{d\mathbf{f}}{d\mathbf{y}}\right)\left(\frac{d\mathbf{y}}{d\mathbf{x}}\right)$
- $\nabla \cdot (\mathbf{f}(\mathbf{y})) = \operatorname{tr}\left(\left(\frac{d\mathbf{f}}{d\mathbf{y}}\right)\left(\frac{d\mathbf{y}}{d\mathbf{x}}\right)\right)$
- For a vector field $\mathbf{v}$, $\nabla^2 \mathbf{v} = \nabla(\nabla \cdot \mathbf{v}) - \nabla \times (\nabla \times \mathbf{v})$
- **Solenoidal**: divergence free $\nabla \cdot \mathbf{v} = 0$
- **Irrotational**: curl free $\nabla \times \mathbf{v} = 0$

Examples

- $D(A\mathbf{x}) = A$
- $\nabla\left(\mathbf{x}^T A \mathbf{x}\right) = \left(A^T + A\right)\mathbf{x}$

2.1 Some useful derivatives

- $\nabla |\mathbf{x}| = \frac{\mathbf{x}}{|\mathbf{x}|}$
- $D\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) = \frac{|\mathbf{x}| I - \mathbf{x}\frac{\mathbf{x}^T}{|\mathbf{x}|}}{|\mathbf{x}|^2} = \frac{1}{|\mathbf{x}|}\left(I - \frac{\mathbf{x}\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}}\right) = \frac{\mathbf{x}^T\mathbf{x} I - \mathbf{x}\mathbf{x}^T}{|\mathbf{x}|^3}$
- $D\left(\frac{\mathbf{v}}{|\mathbf{v}|}\right) = \left(\frac{d}{d\mathbf{v}}\left(\frac{\mathbf{v}}{|\mathbf{v}|}\right)\right)(D\mathbf{v}) = \frac{1}{|\mathbf{v}|}\left(I - \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\right)(D\mathbf{v})$
- If $\mathbf{x}$ is a function of t then using the vector triple product (section 1.1.4) we get the nice expression

$$\frac{d}{dt}\frac{\mathbf{x}}{|\mathbf{x}|} = \left(\frac{d}{d\mathbf{x}}\frac{\mathbf{x}}{|\mathbf{x}|}\right)\mathbf{x}' = \frac{\mathbf{x}'(\mathbf{x}\cdot\mathbf{x}) - \mathbf{x}(\mathbf{x}\cdot\mathbf{x}')}{|\mathbf{x}|^3} = \frac{\mathbf{x}\times(\mathbf{x}'\times\mathbf{x})}{|\mathbf{x}|^3}$$

Derivatives

$f, g : \mathbb{R}^n \to \mathbb{R}^m$ multivariable functions of independent variables $(x_1, \ldots, x_n)$.

Want to find $\nabla(f \cdot g) = \begin{pmatrix} \frac{\partial}{\partial x_1}(f \cdot g) \\ \frac{\partial}{\partial x_2}(f \cdot g) \\ \vdots \\ \frac{\partial}{\partial x_n}(f \cdot g) \end{pmatrix}$

2.2 Curvilinear Coordinates

- Suppose you use curvilinear coordinates u_1, u_2, u_3.
- Let $h_i = \left|\frac{\partial \mathbf{x}}{\partial u_i}\right|$. *i.e.* the magnitude of the tangent to the u_i curve at $\mathbf{x}$.
- Then $\nabla \phi = \left(\frac{1}{h_1}\frac{\partial \phi}{\partial u_1}, \frac{1}{h_2}\frac{\partial \phi}{\partial u_2}, \frac{1}{h_3}\frac{\partial \phi}{\partial u_3}\right)$ where these components are in the curvilinear coordinates.

2.2.1 Cylindrical coordinates

- We have $r = \sqrt{x^2 + y^2}$, $\phi = \tan^{-1}(y/x)$, z
- Then $h_1 = 1$, $h_2 = \frac{1}{r}$, $h_3 = 1$.

2.3 Physics-related Theorems

2.3.1 Gradient

- [Schey, p140]discusses deriving the gradient for cylindrical and spherical coordinates. See also [Weisstein, Gradient]

2.3.2 Divergence

- Defined as

$$\nabla \cdot \mathbf{F} = \lim_{V \to 0} \frac{\oint_S \mathbf{F} \cdot d\mathbf{a}}{\mathbf{V}}$$

  see [Schey, p37] and [Weisstein, Divergence]

2.3.3 Curl

- Defined(?) as

$$(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} = \lim_{A \to 0} \frac{\oint_C \mathbf{F} \cdot d\mathbf{s}}{A}$$

  where C is a curve bounding area A which has normal $\hat{\mathbf{n}}$. *i.e.* the right hand side is the limiting value of circulation per unit area.

- See [Schey, p80] and [Weisstein, Curl]

2.3.4 Gradient Theorem

From [Weisstein]

Looks like a generalization of the fundamental theorem of Calculus:

$$\int_a^b (\nabla f) \cdot d\mathbf{s} = f(b) - f(a)$$

- Integral is a line integral
- This is what makes scalar potential functions so useful in gravitation/electromagnetism.
- Necessary and sufficient condition that permits a vector function to be represented by the gradient of a scalar function is that the curl of the vector function vanishes. [Marion and Thornton, p79] (This is Poincaré's theorem, see [Weisstein, Line Integral]).
- See also [Weisstein, Conservative Field]

2.3.5 Divergence Theorem

- Equivalent to Green's Theorem (?)
- AKA Gauss's Theorem (in [Marion and Thornton, p43])
- Volume in space:

$$\int_V (\nabla \cdot \mathbf{F})\, dV = \int_{\partial V} \mathbf{F} \cdot d\mathbf{a}$$

- Region in the plane:

$$\int_S (\nabla \cdot \mathbf{F})\, dA = \int_{\partial S} \mathbf{F} \cdot \mathbf{n} ds$$

- Informal proof in [Schey, p45]:
  - Divide volume into voxels
  - Flux of vector function $\mathbf{F}$ through surface S roughly equals the sum of fluxes through surfaces of voxels:
    - $\int\int_S \mathbf{F} \cdot \hat{\mathbf{n}}\, dS = \sum \int\int_{S_l} \mathbf{F} \cdot \hat{\mathbf{n}}\, dS$
  - Write $\int\int_S \mathbf{F} \cdot \hat{\mathbf{n}}\, dS = \sum \left[\frac{1}{\Delta V_l} \int\int_{S_l} \mathbf{F} \cdot \hat{\mathbf{n}}\, dS\right] \Delta V_l$
  - As $\Delta V_l \to 0$, quantity in [] goes to $\nabla \cdot \mathbf{F}$ (by definition).
  - At the same time the number of voxels goes to infinity so really we get the $\sum$ turning into the triple intergral: $\int\int\int_V \nabla \cdot \mathbf{F}\, dV$, which is the divergence theorem.

- Results:
  - If A is a matrix (second-order tensor) then $\int_{\partial V} A\mathbf{n}\, da = \int_V (\nabla \cdot A)\, dV$ where $\nabla \cdot A$ is row-wise divergence.
  - If A is constant then we see that $\int_{\partial V} A\mathbf{n}\, da = 0$
  - If $A = \mathbf{x}^*$ (*i.e.* $\int_{\partial V} \mathbf{x} \times \mathbf{n}\, da$) then since $\nabla \cdot \mathbf{x}^* = 0$ we also get $\int_{\partial V} \mathbf{x} \times \mathbf{n}\, da = 0$

2.3.6 Stokes' Theorem

- From [Schey, p96]

$$\oint_C \mathbf{F} \cdot \hat{\mathbf{t}}\, ds = \int\int_S \hat{\mathbf{n}} \cdot (\nabla \times \mathbf{F})\; dS$$

  where S is a capping surface of curve C.

- Informal proof:

– Divide capping surface into polygons (*i.e.* approximate it by a polyhedron)

– Consider sum of circulations around all of the polyhedron's faces. The contribution by internal edges will cancel out in the circulation of the adjacent faces, so the sum telescopes to simply be the circulation across the boundary curve.

$$\int_C \mathbf{F} \cdot \hat{\mathbf{n}} \, ds = \sum_l \oint_{C_l} \mathbf{F} \cdot \hat{\mathbf{t}} \, ds = \sum_l \left[\frac{1}{\Delta S_l} \oint_{C_l} \mathbf{F} \cdot \hat{\mathbf{t}} \, ds \right] \Delta S_l$$

and as you take the limit of the right expression you get $\int \int_S \hat{\mathbf{n}} \cdot (\nabla \times \mathbf{F}) \, dS$ as desired.

2.3.7 Green's Theorem

From [Weisstein, Green's Theorem]

f, g functions of x, y, then

$$\int_{\partial D} (f \, dx + g \, dy) = \int \int_D \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx \, dy$$

- So for example if $f = -\frac{1}{2}y, g = \frac{1}{2}x$, then get

$$\frac{1}{2} \int_{\partial D} (x \, dy - y \, dx) = \text{Area}\,(D)$$

- Equivalent to the curl theorem in the plane

2.3.8 Curl Theorem

From [Weisstein, Curl Theorem]

- Special case of Stoke's Theorem
- $\mathbf{F}$ a vector field, ∂S the boundary of a 2-manifold in $\mathbb{R}^3$

$$\int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{a} = \int_{\partial S} \mathbf{F} \cdot d\mathbf{s}$$

2.3.9 Relationship between these things

From [Schey, p118] (also see [Weisstein, Poincare's Theorem])

- Assume $\mathbf{F}$ is smooth. The following are almost equivalent:

  1. $\int_C \mathbf{F} \cdot \hat{\mathbf{t}} \, ds$ independent of path
  2. $\mathbf{F} = \nabla \psi$
  3. $\nabla \times \mathbf{F} = 0$

- The "almost" is because 3 $\implies$ 2 and 3 $\implies$ 1 only in a simply connected region.
- 1 $\implies$ 2 because you can define ψ for all (x,y,z) by picking some fixed (x_0,y_0,z_0), taking a curve C that goes from (x_0,y_0,z_0) to (x,y,z), and defining $\psi(x,y,z)=\int_C \mathbf{F}\cdot\hat{\mathbf{t}}\,ds$. Then to prove *e.g.* $\frac{\partial\psi}{\partial x}=F_x$ you consider a particular curve C that goes from (x_0,y_0,z_0) to (x,y,z) in a special way: it first goes from (x_0,y_0,z_0) to (x_0,y,z) and then to (x,y,z): the first part of this curve has path integral independent of x, and when you differentiate the second with respect to x you get F_x as required...
- 2 $\implies$ 1 because $\mathbf{F}\cdot\hat{\mathbf{t}}=\frac{\partial\psi}{\partial x}\frac{dx}{ds}+\frac{\partial\psi}{\partial y}\frac{dy}{ds}+\frac{\partial\psi}{\partial z}\frac{dz}{ds}=\frac{d\psi}{ds}$so $\int_C\mathbf{F}\cdot\hat{\mathbf{t}}\,ds=\int_C d\psi=\psi(x,y,z)-\psi(x_0,y_0,z_0)$ which is independent of the path taken.
- 1 $\implies$ 3 by definition of curl because the path integral is zero for closed paths
- 3 $\implies$ 1 in a simply connected region using Stokes theorem.

2.4 Divergence and Gradient are "adjoints" of each other

This is often referred to with regards to the Hodge decomposition.

From [Colella and Puckett] I gathered the following:

Define the following two inner products:

$$\begin{aligned}\langle f,g\rangle_s &= \int_D f(\mathbf{x})\,g(\mathbf{x})\;dV\\ \langle \mathbf{u},\mathbf{v}\rangle_v &= \int_D \mathbf{u}(\mathbf{x})\cdot\mathbf{v}(\mathbf{x})\;dV\end{aligned}$$

Then $\langle\phi,\nabla\cdot\mathbf{u}\rangle_s=\int_D\phi\nabla\cdot\mathbf{u}\,dV=\int_D\nabla\cdot(\phi\mathbf{u})\;dV-\int_D\nabla\phi\cdot\mathbf{u}\,dV=\int_{\partial D}\phi\mathbf{u}\cdot\hat{\mathbf{n}}\,dA-\int_D\nabla\phi\cdot\mathbf{u}\,dV$

If $\mathbf{u}\cdot\hat{\mathbf{n}}=0$ on ∂D then this gives

$$\langle\phi,\nabla\cdot\mathbf{u}\rangle_s=-\langle\nabla\phi,\mathbf{u}\rangle_v$$

so more precisely $\nabla\cdot$ is the negative adjoint of ∇ (in this case of no-flow boundary condition).

In the context of the Hodge decomposition , given the vector field $\mathbf{u}$ on a simply connected domain D satisfying $\mathbf{u}\cdot\hat{\mathbf{n}}$ =0 on ∂D, we can uniquely decompose $\mathbf{u}$as

$$\mathbf{u}=\mathbf{v}+\nabla\phi$$

with $\nabla\cdot\mathbf{v}=0$. Now, this is an orthogonal decomposition in the sense of $\langle\cdot,\cdot\rangle_v$ since

$$\langle\nabla\phi,\mathbf{v}\rangle_v=-\langle\phi,\nabla\cdot\mathbf{v}\rangle_s=0$$

(which presumebly requires assuming $\mathbf{v}\cdot\hat{\mathbf{n}}=0$)

Note that $\langle\nabla\phi,\mathbf{v}\rangle_v=0$ says that divergence free fields are orthogonal to gradients of scalars. Also,

- Orthogonality is used to guarantee that the projection is well defined
- We also find $\langle \mathbf{u}, \mathbf{u} \rangle_v = \langle \mathbf{v}, \mathbf{v} \rangle_v + \langle \nabla\phi, \nabla\phi \rangle_v \geq \langle \mathbf{v}, \mathbf{v} \rangle_v$ which shows that the projection is norm reducing (which according to [Minion, 1996] was used by Chorin to prove that the overall projection method is stable)

2.5 Continuity Equation

2.5.1 Derivation in [Schey, p50]

- Fix a volume V. The amount of stuff in V at time t is

$$\int\int\int_V \rho(x, y, z, t)\, dV$$

- The rate at which this mass is changing is

$$\frac{d}{dt}\int\int\int_V \rho dV = \int\int\int_V \frac{\partial\rho}{\partial t} dV$$

where we can take the derivative inside because $\frac{\partial\rho}{\partial t}$ is continuous

- The rate at which stuff flows outside V through its surface S is

$$\int\int_S \rho\mathbf{v}\cdot\hat{\mathbf{n}}\, dS$$

($\mathbf{v}$ is velocity; stuff flows past cross-sectional area ΔS at the rate $\rho\mathbf{v}\Delta S$ if the cross-sectional area is perpendicular to flow – the $\mathbf{v}\cdot\hat{\mathbf{n}}$ adds a cosine term in case the cross section is not perpendicular).

- Using the divergence theorem

$$\int\int_S (\rho\mathbf{v})\cdot\hat{\mathbf{n}}\, dS = \int\int\int_V \nabla\cdot(\rho\mathbf{v})\, dV$$

- We end up with

$$\int\int\int_V \frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v})\ dV = 0$$

and since this holds for all V we get

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0$$

as required.

2.5.2 Derivation in [Schaum, 5.1]

- Here they also look at $\int\int\int_V \rho dV$ but the volume V is *not* fixed, but moves together with the underlying mass. So they write

$$\frac{d}{dt}\int\int\int_V \rho dV = 0$$

the trick is that when they take the derivative inside they get $\int\int\int_V \frac{d}{dt}(\rho dV) = 0$ and they write

$$\frac{d}{dt}(\rho dV) = \frac{d\rho}{dt}dV + \rho\frac{d}{dt}(dV)$$

with

$$\frac{d}{dt}(dV) = (\nabla\cdot\mathbf{v})\,dV$$

and

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \mathbf{v}\cdot\nabla\rho$$

and $\frac{d}{dt}(\rho dV)$ becomes $\left(\frac{\partial\rho}{\partial t} + \nabla(\mathbf{v}\cdot\rho)\right)dV$, and we get the same result as earlier.

3 Calculus

3.1 Differentiation and Integration [Marsden and Hoffman, sec. 9.7]

- If $f, f' \in C[a,b]$ and $a < x < b$ then

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x) = f(a) + \int_a^x f'(t)\,dt$$

- $f : [a,b]\times[c,d] \to \mathbb{R}$ continuous, $\partial f/\partial y$ exists for $c < y < d$ and extends to be continuous on $[a,b]\times[c,d]$. Let

$$F(y) = \int_a^b f(x,y)\,dx$$

Then F is differentiable and

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx$$

3.2 Theorems

3.2.1 Inverse Function Theorem

- **Suppose** $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$ is C^p, $p \geq 1$, A is an open set, $x_0 \in A$, $\det(\mathbf{D}f(x_0)) \neq 0$.

- **Then** there are open neighbourhoods $x_0 \in U, f(x_0) \in W$ such that $f(U) = W$ and the restriction of f to U has a C^p inverse $f^{-1} : W \to U$. And $\mathbf{D}f^{-1}(y) = [\mathbf{D}f(x)]^{-1}$ where $x = f^{-1}(y)$.

3.2.2 Implicit Function Theorem

- $F : A \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is C^p, $p \geq 1$, $F(x_0, y_0) = 0$. $\Delta = \left|\frac{\partial F_i}{\partial y_j}(x_0, y_0)\right| \neq 0$. Then there are open neighbourhoods $x_0 \in U \subset \mathbb{R}^n$ and $y_0 \in V \subset \mathbb{R}^m$ and a unique function $f : U \to V$ such that $F(x, f(x)) = 0$ and f is C^p.

- Proof using the inverse function theorem applied to the map $G(x, y) = (x, F(x, y))$. Then take $f(x)$ determined by $(x, f(x)) = G^{-1}(x, 0)$

- Differentiating $F(x, f(x)) = 0$ using the chain rule gives $0 = \frac{dF_i}{dx_j} = \frac{\partial F_i}{\partial x_j} + \frac{\partial F_i}{\partial y_k}\frac{\partial f_k}{\partial x_j}$ so that $0 = \mathbf{D}_x F + (\mathbf{D}_y F)(\mathbf{D}f)$ and we get $\mathbf{D}f = -(\mathbf{D}_y F)^{-1}(\mathbf{D}_x F)$.

4 Calculus of Variations

Want to find extreme points of

$$J = \int_{x_1}^{x_2} f(y, y'; x)\, dx$$

where y is a function of x, and $y' = \frac{dy}{dx}$. f is known as a *functional* because it depends on the functional form of dependent variable y.

Suppose $y(\alpha, x) = y(x) + \alpha\eta(x)$ where $\eta(x_1) = \eta(x_2) = 0$ (so $y(\alpha, x)$ matches $y(x)$ at the endpoints). To find the extreme/critical/stationary points we want

$$\left.\frac{\partial J}{\partial \alpha}\right|_{\alpha=0} = 0$$

for all functions $\eta(x)$. This is only a *necessary* condition; it is not sufficient.

So

$$\begin{aligned} \frac{\partial J}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\int_{x_1}^{x_2} f(y, y'; x) \\ &= \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha}\right) dx \\ &= \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta'\right) dx \end{aligned}$$

using integration by parts

$$\int_{x_1}^{x_2}\frac{\partial f}{\partial y'}\eta' dx = \underbrace{\left.\frac{\partial f}{\partial y'}\eta\right|_{x_1}^{x_2}}_{0\ (\eta(x_1)=\eta(x_2)=0)} - \int_{x_1}^{x_2}\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\eta dx$$

so

$$\frac{\partial J}{\partial \alpha} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) \eta(x)\, dx$$

Note that here the y and y' are functions of α as well as x. We want $\frac{\partial J}{\partial \alpha}\big|_{\alpha=0} = 0$ for *arbitrary* $\eta(x)$, hence the integrand itself must vanish for $\alpha = 0$. That is, we get **Euler's equation**

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} = 0$$

where since $\alpha = 0$, y and y' are now the original functions independent of α.

- When applied to mechanical sysytems, this is known as the **Euler-Lagrange equation**.
- We can also derive the equation

$$\frac{\partial f}{\partial x} - \frac{d}{dx} \left(f - y' \frac{\partial f}{\partial y'} \right) = 0$$

  so that, when f does not depend on x ($\frac{\partial f}{\partial x} = 0$) we get

$$f - y' \frac{\partial f}{\partial y'} = \text{constant}$$

- More generally, for independent functions $y_1, \ldots, y_n$, the extreme values of

$$J = \int_{x_1}^{x_2} f(y_1, y_1', \ldots, y_n, y_n'; x)\, dx$$

  occur when $y_1, \ldots, y_n$ satisfy the independent equations

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial y_i'} = 0$$

- [Horn, Robot Vision, p284] And for $J = \int \int F(u, u_x, u_y; x, y)\, dxdy$ get $F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} = 0$.

Lagrange Multiplier

If have $y_1, \ldots, y_m$ *dependent* functions, with a functional $f\left(y_1, y_1', \ldots, y_m, y_m'; x\right)$, and n constraint equations $g_j(y_1, \ldots, y_m; x) = 0$ $1 \le j \le n$, then you solve the following set of equations

$$\begin{aligned} \frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial y_i'} + \sum_j \lambda_j(x) \frac{\partial g_j}{\partial y_i} &= 0 \\ g_j(y_i; x) &= 0 \end{aligned}$$

5 Eikonal Equation

$$|\phi(\mathbf{x})| = f(\mathbf{x})$$

Method of characteristics (my derivation):

Let $\mathbf{p} = \nabla\phi$, and let $H(\mathbf{p},\mathbf{x}) = |\mathbf{p}|^2 - f^2$, then $H \equiv 0$. Now suppose we're following a ray $\mathbf{x}(\mathbf{t})$ parameterized by t. Then (using $\frac{df}{dt} = \nabla f \cdot \frac{d\mathbf{x}}{dt}$) $\frac{dH}{dt} = 2\mathbf{p} \cdot \frac{d\mathbf{p}}{dt} - 2f\nabla f \cdot \frac{d\mathbf{x}}{dt} = 0$. Letting $\frac{d\mathbf{p}}{dt} = \alpha f \nabla f$ and $\frac{d\mathbf{x}}{dt} = \alpha\mathbf{p}$ satisfies this equation. If we additionally desire $1 = \frac{d\phi}{dt} = \nabla\phi \cdot \frac{d\mathbf{x}}{dt} = \alpha\mathbf{p}\cdot\mathbf{p} = \alpha f^2$ then $\alpha = \frac{1}{f^2}$.

Hence the curve is defined by:

- $\frac{d\mathbf{x}}{dt} = \frac{\mathbf{p}}{f^2}$ (note $\left|\frac{d\mathbf{x}}{dt}\right| = \frac{1}{f}$ hence the speed of the curve is $\frac{1}{f}$)
- The change in $\mathbf{p}$ as we follow the curve is $\frac{d\mathbf{p}}{dt} = \frac{\nabla f}{f}$
- Since the change in ϕ as we follow the curve is $\frac{d\phi}{dt} = 1$, if the ray starts at $\phi_0 = 0$ then $\phi(\mathbf{x}(t)) = t$. *i.e.* ϕ is the time of arrival

Hamiltonian (see Cheng et al. "Level set based Eulerian methods for multivalued...")

- They call $H(\mathbf{x},\mathbf{p}) = \frac{1}{2}\left(\frac{|\mathbf{p}|^2}{f^2} - 1\right)$ a Hamiltonian, and then we get the same equation as above using the method of characteristics (which I think is related to Hamilton's canonical equations) which are $\frac{d\mathbf{x}}{dt} = \frac{\partial H}{\partial \mathbf{p}}$ and $\frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{x}}$.

6 Matrices

6.1 Definitions

6.1.1 Matrix Norms[Golub and van Loan, 2.3]

- Definition of norm analogous to vector norm
- **Frobenius norm**: $\|A\|_F = \sqrt{\sum_{ij} |a_{ij}|^2}$
- p-**norms**: $\|A\|_p = \sup_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p = 1} \|A\mathbf{x}\|_p$
  - $\|A\mathbf{x}\|_p \leq \|A\|_p \|\mathbf{x}\|_p$
  - Submultiplicative property: $\|AB\|_p \leq \|A\|_p \|B\|_p$ A, B arbitrary shapes (but not all norms have this property)
  - $\|A\|_1 = \max_j \sum_i |a_{ij}|$
  - $\|A\|_\infty = \max_i \sum_j |a_{ij}|$

  - [Golub and van Loan, p57]$A \in \mathbb{R}^{m \times n}$, then $\exists \mathbf{z} \in \mathbb{R}^n$ with $\|\mathbf{z}\|_2 = 1$ such that $A^T A\mathbf{z} = \|A\|_2^2 \mathbf{z}$
    * $\|A\|_2^2$ is the largest eigenvalue of $A^T A$
  - [Golub and van Loan, p57]$A \in \mathbb{R}^{m \times n}$, then $\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty}$

- All norms on $\mathbb{R}^{m \times n}$ are equivalent

6.1.2 Special Matrices

- **Symmetric**: $A^T = A$
- **Self-adjoint (Hermitian):** $A^* = A$
- **Skew-symmetric**: $A^T = -A$
- **Normal matrix**: $AA^* = A^* A$
- **Unitary (over $\mathbb{C}$)** or **Orthogonal (over $\mathbb{R}$)**: $AA^* = A^* A = I$.
  - In finite dimensions, equivalent to being an isometry $\|Ax\| = \|x\|$.
- **Diagonally dominant**:

$$|a_{ii}| \geq \sum_{j \neq i}^{n} |a_{ij}| \text{ for } 1 \leq i \leq n$$

  - **Strictly diagonally dominant** if $\geq$ replaced by $>$.

6.1.3 Other Definitions

- **Similar matrices**: $B = Q^{-1} AQ$
- **Unitarily equivalent (or Orthogonally equivalent in $\mathbb{R}$):** $B = Q^* AQ$
- **Condition number**: $\kappa(A) = \|A\| \left\|A^{-1}\right\|$ (for some fixed norm; $\kappa(A)$ defined as ∞ for singular A).
  - Measures sensitivity of solution of $A\mathbf{x} = \mathbf{b}$ to perturbations of A or $\mathbf{b}$.
  - **Ill-conditioned**: κ large
  - **Perfectly conditioned**: $\kappa = 1$

6.2 Determinant

The determinant of a 3x3 matrix has a nice form (as seen in 1.1.3):

$$\text{If } A = (\mathbf{c}_1 \mid \mathbf{c}_2 \mid \mathbf{c}_3) = \begin{pmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{pmatrix} \text{ then } |A| = \mathbf{c}_1 \cdot (\mathbf{c}_2 \times \mathbf{c}_3) = \mathbf{r}_1 \cdot (\mathbf{r}_2 \times \mathbf{r}_3).$$

6.3 Misc. Properties

- $trace\,(AB) = trace\,(BA)$

6.4 Diagonalization

6.4.1 Definitions

- Diagonalizable: similar to a diagonal matrix

6.4.2 Properties / Theorems

- If $\lambda_1, \ldots, \lambda_k$ *distinct* eigenvalues and $v_1, \ldots, v_k$ associated eigenvectors then $\{v_1, \ldots, v_k\}$ are linearly independent. [Friedberg et al., Thm. 5.10]
  - Proof: By induction. Base case: $\{v_1\}$ is linearly independent. Inductive step: Assuming $\{v_1, \ldots, v_{k-1}\}$ linearly independent, suppose $a_1 v_1 + \cdots + a_k v_k = 0$. Multiply both sides by $(A - \lambda_k I)$ to get $a_1 (\lambda_1 - \lambda_k) + \cdots + a_{k-1} (\lambda_{k-1} - \lambda_k) = 0$ which gives $a_1 = \cdots = a_{k-1} = 0$ by inductive hypothesis, and we must also then have $a_k = 0$ proving linear independence.
- Dimension of eigenspace is $\leq$ multiplicity of corresponding eigenvalue
- Diagonalizable iff multiplicity of λ_i equals dimension of eigenspace for all i. Then take a basis for each E_λ and combine to get basis for whole space consisting of eigenvectors.
- Diagonalizable iff space is the direct sum of the eigenspaces.
- $\dim (E_\lambda) = \text{nullity}\,(A - \lambda I) = n - \text{rank}\,(A - \lambda I)$
- If λ is an eigenvalue of A with associated eigenvector x, and f is a polynomial (with coefficients in the same field as A), then $f(\lambda)$ is an eigenvalue of $f(A)$ with associated eigenvector x. [Friedberg et al., p250#22]

6.4.3 Diagonalization of special matrices

- From [Friedberg et al., p351-353,p360]
  - A complex matrix is normal iff it is unitarily equivalent to a diagonal matrix
  - A real matrix is symmetric iff it is orthogonally equivalent to a diagonal matrix
- **Schur's Theorem** [Friedberg et al., p363]: A a matrix whose characteristic polynomial splits over F. If $F = \mathbb{C}$ then A is unitarily equivalent to a complex upper triangular matrix. If $F = \mathbb{R}$ then A is orthogonally equivalent to a real upper triangular matrix

- **Spectral Theorem** [Friedberg et al., p377]: T a linear operator on fin.dim. i.p.s. V over F with distinct eigenvalues. Assume T normal (if $\mathbb{C}$) or self-adjoint (if $\mathbb{R}$). Then $T = \sum \lambda_i T_i$ where T_i is the orthogonal projection onto the eigenspace corresponding to λ_i.
  - On a sort of related note, if A is a unitarily/orthogonally diagonalizable with eigenvalues λ_i and eigenvectors v_i, then $A = VDV^*$ where $V = (v_1 \mid v_2 \mid \cdots \mid v_n)$ and you can write

$$A = \sum \lambda_i v_i v_i^* = \sum \lambda_i v_i \otimes v_i$$

- $|A| = \prod \lambda_i$, $\text{trace}\,(A) = \sum A_{ii} = \sum \lambda_i$
  - **Proof**: Compare the characteristic polynomial $C_A(\lambda) = |A - \lambda I|$ to $\prod (\lambda_i - \lambda)$. In particular compare the constant terms and the λ^{n-1} terms.
- A symmetric then $\rho(A) = \|A\|_2$

6.5 Adjoint, etc.

- [Friedberg et al., 6.3]
- V finite dimensional inner product space over F, $g : V \to F$ linear transformation. Then $\exists y \in V$ such that $g(x) = \langle x, y \rangle$.
- V finite dimensional i.p.s., T a linear operator on V. Then there exists a unique linear operator T^* such that $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$.
  - In the infinite dimensional i.p.s. case, T^* may not exist, but when it does it is unique and linear.
- If $A \in M_{m \times n}(F)$ then $\text{rank}\,(A^*A) = \text{rank}\,(A)$. In particular if A has rank n then A^*A is invertible.

6.5.1 Significance of self-adjoint matrices:

- Just like all eigenvalues of a real matrix are real (by definition), all eigenvalues of a self-adjoint complex matrix are real
- Just like the characteristic polynomial of a complex matrix splits (by fundamental theorem of algebra), it also splits for a real self-adjoint matrix

6.6 Canonical Forms

- Alternative matrix representations for nondiagonalizable operators
- **Jordan Canonical Form**
  - Requires that the characteristic polynomial splits
  - Get matrix

$$\begin{pmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_k \end{pmatrix}$$

    where each block is of the form

$$\begin{pmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

- Basically if the characteristic polynomial of A looks like $(\lambda_1 - \lambda)^{k_1} \cdots (\lambda_m - \lambda)^{k_m}$ then we look at the generalized eigenspaces $E_{\lambda_i} = \text{null}\left((A - \lambda_i I)^{k_i}\right)$. The claim is that a generalized eigenspace has a basis consisting of a union of cycles. *e.g.* if you find an element $\mathbf{v}$ in E_{λ_i} for which $(A - \lambda_i I)^{k_i - 1} \mathbf{v} \neq \mathbf{0}$ then we can look at the cycle $\left\{\mathbf{v}, (A - \lambda_i I)\mathbf{v}, \ldots, (A - \lambda_i I)^{k_i - 1}\mathbf{v}\right\}$ which would be all distinct and so would give you a Jordan block... [Friedberg et al., ch.7]

6.7 Symmetric Positive Definite

- Definition in [Friedberg et al., p355] (for a linear operator)
  - Linear operator T on fin.dim. i.p.s. is **positive definite** if T is self-adjoint and $\langle T(x), x \rangle > 0$ for all $x \neq 0$.
- If A is symmetric, diagonally dominant, with positive diagonal elements, then A is positive definite. (mentioned in Multigrid tutorial book)
- T is positive definite (semidefinite) iff all its eigenvalues are positive (non-negative)
- T is positive semidefinite iff $A = B^* B$ ($A = [T]_\beta$ where β an orthonormal basis of V)
- If $T^2 = U^2$ where both are positive semidefinite operators then $T = U$

- L. Adams (1985). "m-step Preconditioned Conjugate Gradient Methods," SIAM J. Sci. and Stat. Comp. 6, 452-463. (Lemma 1)
  - If $A = BC$ is symmetric, B is symmetric positive definite and C has positive eigenvalues then A is positive definite

6.8 Factorization

6.8.1 Polar Decomposition [Golub and van Loan, p149]

$A \in \mathbb{R}^{m \times n}$ with $m \geq n$. Then we can write $A = ZP$ where $Z \in \mathbb{R}^{m \times n}$ has orthonormal columns and P is symmetric positive semidefinite.

6.8.2 Cholesky [Golub and van Loan, p143]

$A \in \mathbb{R}^{n \times n}$ is symmetric positive definite. Then can write $A = GG^T$ with G lower triangular.

6.8.3 Square Root [Golub and van Loan, p149]

$A \in \mathbb{R}^{n \times n}$ is symmetric positive semidefinite. Then there exists a unique symmetric positive semidefinite matrix X such that $A = X^2$.

6.9 Misc

6.9.1 Cramer's Rule [Friedberg et al., p213]

Given the system $Ax = b$, with $|A| \neq 0$. The solution is $x_i = \frac{|M_i|}{|A|}$where M_i is the matrix $A$with column i replaced by b**.**

Proof: $M_i = AX_i$ where X_i is the identity with column i replaced by x. One can show that $|X_i| = x_i$. Hence $|M_i| = |A|\,|X_i| = |A|\,x_i$.

6.9.2 Cofactor Matrix (?)

A is an $n \times n$ invertible matrix. The minor matrix M_{ij} is the submatrix of $A$with row i and column j deleted. The cofactor c_{ij} corresponding to element a_{ij} is $c_{ij} = (-1)^{i+j}\,|M_{ij}|$. Then we have the property

$$A^{-1} = \frac{1}{|A|} C^T$$

where $(C)_{ij} = c_{ij}$.

This is related to Cramer's rule in that $C^T b = \begin{pmatrix} |M_1| \\ |M_2| \\ \vdots \\ |M_n| \end{pmatrix}$ and $Ax = b \iff x = A^{-1} b \iff x = \frac{C^T b}{|A|}$.

Special case for 3x3 matrix

If A is a 3x3 matrix of the form $A = \begin{pmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{pmatrix}$ then the cofactor matrix is $C = \begin{pmatrix} (\mathbf{r}_2 \times \mathbf{r}_3)^T \\ (\mathbf{r}_3 \times \mathbf{r}_1)^T \\ (\mathbf{r}_1 \times \mathbf{r}_2)^T \end{pmatrix}$. Indeed, one can see that

$$AC^T = \begin{pmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{pmatrix} ((\mathbf{r}_2 \times \mathbf{r}_3) \mid (\mathbf{r}_3 \times \mathbf{r}_1) \mid (\mathbf{r}_1 \times \mathbf{r}_2)) = |A|\, I$$

(*e.g.* $\left(AC^T\right)_{1,1} = \mathbf{r}_1 \cdot (\mathbf{r}_2 \times \mathbf{r}_3) = |A|$**,** $\left(AC^T\right)_{2,1} = \mathbf{r}_2 \cdot (\mathbf{r}_2 \times \mathbf{r}_3) = 0$**)**

6.9.3 Matrix multiplication in outer product form

- If $A = (\mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n)$ and $B = (\mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_n)^T$. *i.e.* $\mathbf{a}_i$ is the ith column of A, and $\mathbf{b}_j$ is the jth row of B. Then $C = AB$ can be written as

$$C = \sum_{k=1}^{n} \mathbf{a}_k \mathbf{b}_k^T$$

6.10 Extra cross product properties

Claim

If A is a 3x3 matrix with cofactor matrix C (*i.e.* $AC^T = |A|\, I$), then we claim:

$$(A\mathbf{u}) \times (A\mathbf{v}) = C\, (\mathbf{u} \times \mathbf{v})$$

If A is invertible then this becomes

$$(A\mathbf{u}) \times (A\mathbf{v}) = |A|\, A^{-T}\, (\mathbf{u} \times \mathbf{v})$$

In fact, if A is invertible then the following also holds

$$(A\mathbf{u}) \times \mathbf{v} = C\left(\mathbf{u} \times \left(A^{-1}\mathbf{v}\right)\right)$$

Proof

Well, here's a proof that works if A is invertible: Let $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ be the rows of C. Then $AC^T = A\,(\mathbf{r}_1 \mid \mathbf{r}_2 \mid \mathbf{r}_3) = |A|\, I$ so in particular

$$A\,(\mathbf{r}_i \mid \mathbf{u} \mid \mathbf{v}) = (|A|\, \mathbf{e}_i \mid A\mathbf{u} \mid A\mathbf{v})$$

for $i = 1, 2, 3$. Taking determinants of both sides we see that

$$|A\,(\mathbf{r}_i \mid \mathbf{u} \mid \mathbf{v})| = |A|\, \mathbf{r}_i \cdot (\mathbf{u} \times \mathbf{v})$$

and

$$|(|A|\,\mathbf{e}_i \mid A\mathbf{u} \mid A\mathbf{v})| = |A|\,\mathbf{e}_i \cdot ((A\mathbf{u}) \times (A\mathbf{v}))$$

Dividing by the factor of $|A|$ we get

$$\mathbf{r}_i \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{e}_i \cdot ((A\mathbf{u}) \times (A\mathbf{v}))$$

which can be written compactly as

$$C\,(\mathbf{u} \times \mathbf{v}) = (A\mathbf{u}) \times (A\mathbf{v})$$

Note that we could have started with

$$A\left(\mathbf{r}_i \mid \mathbf{u} \mid A^{-1}\mathbf{v}\right) = (|A|\,\mathbf{e}_i \mid A\mathbf{u} \mid \mathbf{v})$$

and gotten

$$(A\mathbf{u}) \times \mathbf{v} = C\left(\mathbf{u} \times \left(A^{-1}\mathbf{v}\right)\right)$$

Proof using summation notation that works for singular A:

Using

$$C = \frac{1}{2}\varepsilon_{mni}\varepsilon_{pqj}A_{mp}A_{nq}\mathbf{e}_i \otimes \mathbf{e}_j$$

(see summation notation section), and $\mathbf{u} \times \mathbf{v} = \varepsilon_{rst}u_r v_s \mathbf{e}_t$, we get

$$\begin{aligned}
C\,(\mathbf{u} \times \mathbf{v}) &= \left(\frac{1}{2}\varepsilon_{mni}\varepsilon_{pqj}A_{mp}A_{nq}\mathbf{e}_i \otimes \mathbf{e}_j\right)(\varepsilon_{rst}u_r v_s \mathbf{e}_t) \\
&= \frac{1}{2}\varepsilon_{mni}\varepsilon_{pqj}\varepsilon_{rst}A_{mp}A_{nq}u_r v_s \delta_{jt}\mathbf{e}_i \\
&= \frac{1}{2}\varepsilon_{mni}\varepsilon_{pqj}\varepsilon_{rsj}A_{mp}A_{nq}u_r v_s \mathbf{e}_i \\
&= \frac{1}{2}\varepsilon_{mni}\left(\delta_{pr}\delta_{qs} - \delta_{ps}\delta_{qr}\right)A_{mp}A_{nq}u_r v_s \mathbf{e}_i \\
&= \frac{1}{2}\varepsilon_{mni}\delta_{pr}\delta_{qs}A_{mp}A_{nq}u_r v_s \mathbf{e}_i - \frac{1}{2}\varepsilon_{mni}\delta_{ps}\delta_{qr}A_{mp}A_{nq}u_r v_s \mathbf{e}_i \\
&= \frac{1}{2}\varepsilon_{mni}A_{mp}A_{nq}u_p v_q \mathbf{e}_i - \frac{1}{2}\varepsilon_{mni}A_{mp}A_{nq}u_q v_p \mathbf{e}_i
\end{aligned}$$

Now renaming $m \leftrightarrow n$ and $p \leftrightarrow q$ in the second term

$$\begin{aligned}
\frac{1}{2}\varepsilon_{mni}A_{mp}A_{nq}u_q v_p \mathbf{e}_i &= \frac{1}{2}\varepsilon_{nmi}A_{nq}A_{mp}u_p v_q \mathbf{e}_i \\
&= -\frac{1}{2}\varepsilon_{mni}A_{nq}A_{mp}u_p v_q \mathbf{e}_i
\end{aligned}$$

so we get

$$C\,(\mathbf{u} \times \mathbf{v}) = \varepsilon_{mni}A_{mp}A_{nq}u_p v_q \mathbf{e}_i$$

Now we perform the easier task of writing out$(A\mathbf{u}) \times (A\mathbf{v})$:

$$\begin{aligned}(A\mathbf{u}) \times (A\mathbf{v}) &= (A_{mp}u_p\mathbf{e}_m) \times (A_{nq}v_q\mathbf{e}_n) \\ &= A_{mp}A_{nq}u_pv_q\mathbf{e}_m \times \mathbf{e}_n \\ &= \varepsilon_{mni}A_{mp}A_{nq}u_pv_q\mathbf{e}_i\end{aligned}$$

which is identical to the expression for $C(\mathbf{u} \times \mathbf{v})$.

6.10.1 Consequence

Suppose A is (3x3) invertible. Then we want to find an expression for $(A\mathbf{u})^*$. Well,

$$(A\mathbf{u})^* A\mathbf{v} = (A\mathbf{u}) \times (A\mathbf{v}) = C(\mathbf{u} \times \mathbf{v}) = C\mathbf{u}^*\mathbf{v}$$

Since this holds for all $\mathbf{v}$ we get $(A\mathbf{u})^* A = C\mathbf{u}^*$, or

$$(A\mathbf{u})^* = C\mathbf{u}^*A^{-1} = \frac{1}{|A|}C\mathbf{u}^*C^T$$

6.11 Misc Matrix Properties

- $trace(AB) = trace(BA)$

7 Least Squares

Given $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Want to find $x \in \mathbb{R}^n$ that minimizes

$$\|Ax - b\|_2$$

x is found by solving the associated **Normal Equations**

$$A^TAx = A^Tb$$

7.1 Deriving the normal equations

7.1.1 Calculus method

We can seek to minimize $\|Ax - b\|_2^2$ instead. Let

$$\begin{aligned}f(x) = \|Ax - b\|_2^2 &= (Ax - b)^T(Ax - b) \\ &= x^TA^TAx - x^TA^Tb - b^TAx + b^Tb \\ &= x^TA^TAx - 2b^TAx + b^Tb\end{aligned}$$

where $x^TA^Tb = b^TAx$ because, in general, $x \cdot y = x^Ty = y^Tx$.

Now, the derivative of f is

$$\nabla f(x) = 2A^TAx - 2b^TA = 2A^TAx - 2A^Tb$$

so a solution x_0 to our least squares problem would satisfy $\nabla f(x_0) = 0$. That is

$$A^TAx_0 = A^Tb$$

7.1.2 Linear Algebra method

In [Friedberg et al., p343].

Let $W = \{Ax \mid x \in \mathbb{R}^n\}$. Then W is a closed subspace of $\mathbb{R}^m$, so we can uniquely write $b = u + v$ with $u \in W, v \in W^\perp$. Say $u = Ax_0$. Then we're looking for x_0 because $\|Ax_0 - b\|_2$ is minimized. We have $b - Ax_0 \in W^\perp$ so $\langle Ax, b - Ax_0\rangle = 0$. But $\langle Ax, b - Ax_0\rangle = \left\langle x, A^T\left(b - Ax_0\right)\right\rangle = 0$. Since this holds for all x we get $A^T\left(b - Ax_0\right) = 0 \implies A^TAx_0 = A^Tb$.

8 Curves and Surfaces

8.1 Curvature

8.1.1 Arclength parameterization

For a curve $\mathbf{x} : [a, b] \to \mathbb{R}^3$ which is *parameterized by arclength*, curvature is

$$\kappa\left(s\right) = \left|\mathbf{x}''\left(s\right)\right|$$

- The curvature measures the rate of change of curve direction (unit tangent vector) with respect to arclength.

8.1.2 Arbitrary parameterization

If we're given a curve $\mathbf{y} : [c, d] \to \mathbb{R}^3$with arbitrary parameterization, we can derive an expression for curvature as follows:

- [Notation: $\mathbf{x}', \mathbf{x}''$ are derivatives with respect to s, $\mathbf{y}', \mathbf{y}''$ are derivatives with respect to t]
- Define curve $\mathbf{x}$ implicitly by $\mathbf{x}\left(s\left(t\right)\right) = \mathbf{y}\left(t\right)$where $s\left(t\right) = \text{Arclength}_{\mathbf{y}}\left(t\right) = \int_c^t \left|\mathbf{y}'\left(u\right)\right| du$.
  - $\mathbf{x}\left(t\right)$ is a reparameterization of $\mathbf{y}\left(s\right)$ by arclength.
  - Think of t as being time, and s as being distance. So $\mathbf{y}\left(t\right)$ tells you your position along the path at time t, and $\mathbf{x}\left(s\right)$ tells you your position along the path after you have travelled a distance s.
- $\frac{ds}{dt} = \left|\mathbf{y}'\left(t\right)\right|$by the fundamental theorem of calculus.
  - This makes sense because $\frac{ds}{dt} = \frac{d(\text{distance})}{d(\text{time})} = \text{speed}$
- $\frac{d\mathbf{x}}{ds} = \frac{d\mathbf{y}}{dt}\frac{dt}{ds} = \frac{\mathbf{y}'}{|\mathbf{y}'|}$ (as expected, $\mathbf{x}$ has constant speed)
- $\frac{d^2\mathbf{x}}{ds^2} = \frac{d}{ds}\frac{\mathbf{y}'}{|\mathbf{y}'|} = \left(\frac{d}{d\mathbf{y}'}\frac{\mathbf{y}'}{|\mathbf{y}'|}\right)\left(\frac{d\mathbf{y}'}{dt}\right)\left(\frac{dt}{ds}\right) = \frac{\mathbf{y}'\times\left(\mathbf{y}''\times\mathbf{y}'\right)}{|\mathbf{y}'|^4}$ (using the formula from 2.1)

- We therefore get

$$\kappa = \frac{|\mathbf{y}' \times (\mathbf{y}'' \times \mathbf{y}')|}{|\mathbf{y}'|^4} = \frac{|\mathbf{y}' \times \mathbf{y}''|}{|\mathbf{y}'|^3}$$

- Note that if $\mathbf{y}$ has constant speed around some parameter then since $\mathbf{y}' \cdot \mathbf{y}' = |\mathbf{y}'|^2 = \text{constant}$ we get $0 = \frac{d}{dt}\mathbf{y}' \cdot \mathbf{y}' = 2\mathbf{y}' \cdot \mathbf{y}''$ hence $\mathbf{y}' \perp \mathbf{y}''$ and we get $\kappa = \frac{|\mathbf{y}''|}{|\mathbf{y}'|^2}$.

8.1.3 Alternative formulation

Have a curve $\mathbf{x}(t)$. Let $s(t)$ be the arclength up to $\mathbf{x}(t)$.

$\mathbf{x}'(t)$ is the velocity at time t, $s'(t) = |\mathbf{x}'(t)|$ is the speed.

Define $T(t) = \frac{\mathbf{x}'(t)}{s'(t)}$. T is the (unit) tangent vector at time t.

Then first note that $|T|^2 = T \cdot T = 1$ so differentiating w.r.t s gives $2\frac{dT}{ds} \cdot T = 0$. Hence $\frac{dT}{ds}$ is perpendicular to T. Define

$$N = \frac{dT/ds}{|dT/ds|}$$

and verify that $\kappa = |dT/ds|$. (This holds because $T(s)$ is an arclength parameterized velocity curve, so $\frac{dT}{ds}$ is the arclength parameterized acceleration). Hence $\frac{dT}{ds} = \kappa N$

Then

$$\begin{aligned} \mathbf{x}'(t) &= T(t)\, s'(t) \\ \mathbf{x}''(t) &= T'(t)\, s'(t) + T(t)\, s''(t) \end{aligned}$$

and $T' = \frac{dT}{dt} = \frac{dT}{ds}\frac{ds}{dt} = s'(t)\,\kappa N$ so that the final result is

$$\mathbf{x}''(t) = s''(t)\, T + (s'(t))^2\, \kappa N$$

That is, the acceleration has a tangential component of $s''(t)$ (the rate of change of speed), and a normal component of $s'(t)^2\,\kappa$.

For example, if we had uniform circular motion along a circle of radius r and uniform speed v then the centripetal acceleration (pointing towards the center of the circle) would have magnitude $s'(t)^2\,\kappa$ which (since $\kappa = \frac{1}{r}$) is $\frac{v^2}{r}$.

8.2 Principal Curvatures of an implicit surface

- Suppose $\phi(\mathbf{x}) = 0$ defines the implicit surface.
- Pick a point $\mathbf{x}_0$ on the surface ($\phi(\mathbf{x}_0) = 0$).
- Let $\mathbf{y}(t) = \mathbf{x}_0 + t\mathbf{T} + a(t)\,\mathbf{N}$ be a curve on surface with $\mathbf{N} = \nabla\phi(\mathbf{x}_0) / |\nabla\phi(\mathbf{x}_0)|$ the unit normal, $\mathbf{T} \perp \mathbf{N}$ a unit tangent direction. (This is the general form of the intersection curve you'd get if you intersected a plane through $\mathbf{x}_0$ with normal $\mathbf{T} \times \mathbf{N}$ with the implicit surface).

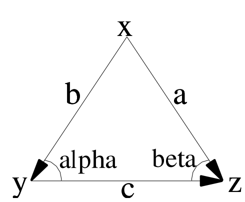
- Note $\mathbf{y}'(t) = \mathbf{T} + a'(t)\mathbf{N}$, $\mathbf{y}'(0) = \mathbf{T}$ (by construction, so $a'(0) = 0$), $\mathbf{y}''(t) = a''(t)\mathbf{N}$. One can check that using our definition of curvature (of a curve) as $\kappa = \frac{|\mathbf{y}'\times\mathbf{y}''|}{|\mathbf{y}'|^3}$ we get the curvature of $\mathbf{y}$ at the point $\mathbf{x}_0$ is $\kappa = a''(0)$.
- Now $\phi(\mathbf{y}(t)) = 0$ since $\mathbf{y}$ lies on the surface.
  - Differentiate once: $\nabla\phi(\mathbf{y}(t))\cdot\mathbf{y}'(t) = 0$. With $t = 0$ we get $\nabla\phi(\mathbf{x}_0)\cdot\mathbf{T} = 0$. (We already know this)
  - Differentiate again: $\left[D^2\phi(\mathbf{y}(t))\,\mathbf{y}'(t)\right]\cdot\mathbf{y}'(t) + \nabla\phi(\mathbf{y}(t))\cdot\mathbf{y}''(t) = 0$. With $t = 0$ we get

    $$\mathbf{T}^T H\mathbf{T} + a''(0)\,|\nabla\phi(\mathbf{x}_0)| = 0$$

    where $H = D^2\phi(\mathbf{x}_0)$, which gives $\kappa = -\frac{\mathbf{T}^T H\mathbf{T}}{|\nabla\phi(\mathbf{x}_0)|}$. [Not sure what's up with the negative sign...]
- We subsequently drop the $\mathbf{x}_0$ and assume quantities are evaluated there...
- Let $P = I - \frac{(\nabla\phi)(\nabla\phi)^T}{|\nabla\phi|^2}$. P is the projection onto the tangent plane at $\mathbf{x}_0$.
- See [http://www.magic-software.com/Documentation/PrincipalCurvature.pdf] for proof that the maximum of κ occurs at the max eigenvalue of $-P\frac{H}{|\nabla\phi|}$...
- Mean curvature: divergence of normal: $\nabla\cdot\left(\frac{\nabla\phi}{|\nabla\phi|}\right)$
- The eigenvalues of $\frac{PHP}{|\nabla\phi|}$ are 0 and the two principal curvatures.
- Suppose you have $\phi(\mathbf{x})$, not necessarily signed distance.
- Claim: $\nabla\cdot\left(\frac{\nabla\phi}{|\nabla\phi|}\right) = \mathrm{tr}\left(\frac{PHP}{|\nabla\phi|}\right)$, $P = I - \frac{(\nabla\phi)(\nabla\phi)^T}{|\nabla\phi|^T}$ is projection onto tangent plane, $H = D^2\phi$ Hessian
  - Note $D\left(\frac{\nabla\phi}{|\nabla\phi|}\right) = \frac{1}{|\nabla\phi|}\left(I - \frac{(\nabla\phi)(\nabla\phi)^T}{|\nabla\phi|^2}\right)D(\nabla\phi) = \frac{PH}{|\nabla\phi|}$.
  - And $\nabla\cdot(\mathbf{f}) = \mathrm{tr}(D\mathbf{f})$
  - And $\mathrm{tr}(PHP) = \mathrm{tr}\left(P^2H\right) = \mathrm{tr}(PH)$ since P is a projection.

9 Miscellaneous

9.1 Derivative of triangle area



- Given triangle A with vertices $\mathbf{x}, \mathbf{y}, \mathbf{z}$ (in ccw order), $\text{Area}(A) = \frac{1}{2}\left|(\mathbf{x}-\mathbf{y}) \times (\mathbf{x}-\mathbf{z})\right|$, and

$$\begin{aligned}
\frac{d\text{Area}(A)}{d\mathbf{x}} &= \frac{1}{2}\frac{((\mathbf{x}-\mathbf{y}) \times (\mathbf{x}-\mathbf{z}))^T}{|(\mathbf{x}-\mathbf{y}) - (\mathbf{x}-\mathbf{z})|}\frac{d}{d\mathbf{x}}((\mathbf{x}-\mathbf{y}) \times (\mathbf{x}-\mathbf{z})) \\
&= \frac{1}{2}(\hat{\mathbf{n}})^T\left((\mathbf{x}-\mathbf{y})^* I - (\mathbf{x}-\mathbf{z})^* I\right) \\
&= \frac{1}{2}(\hat{\mathbf{n}})^T(\mathbf{z}-\mathbf{y})^* \\
&= \frac{1}{2}(\hat{\mathbf{n}} \times (\mathbf{z}-\mathbf{y}))^T
\end{aligned}$$

  where $\hat{\mathbf{n}}$ is the triangle normal, and we've used the cross product derivative formula from section 1.1.6.

- Note that the change in triangle area is *zero* if $\mathbf{x}$ is moved in the direction of the normal or edge $\mathbf{c}$.

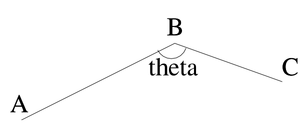
- Note that $\hat{\mathbf{n}} \times (\mathbf{z}-\mathbf{y}) = \hat{\mathbf{n}} \times \mathbf{c}$is edge $\mathbf{c}$ rotated ccw 90 degrees on the plane of the triangle (*i.e.* in the direction of the altitude of the triangle). This is the direction along which $\mathbf{x}$ must be moved to maximize the change in triangle area.

- **Claim**: $\hat{\mathbf{n}} \times \mathbf{c} = -(\cot\alpha)\,\mathbf{a} - (\cot\beta)\,\mathbf{b}$

  - **Proof**: Using $\hat{\mathbf{n}} = \frac{\mathbf{b}\times\mathbf{a}}{|\mathbf{b}\times\mathbf{a}|}$ and using the fact that $\mathbf{b}\times\mathbf{a} = \mathbf{b}\times\mathbf{c} = \mathbf{a}\times\mathbf{c}$

$$\begin{aligned}
\hat{\mathbf{n}} \times \mathbf{c} &= \frac{(\mathbf{b}\times\mathbf{a})\times\mathbf{c}}{|\mathbf{b}\times\mathbf{a}|} \\
&= \frac{\mathbf{a}(\mathbf{c}\cdot\mathbf{b}) - \mathbf{b}(\mathbf{c}\cdot\mathbf{a})}{|\mathbf{b}\times\mathbf{a}|} \\
&= \frac{\mathbf{a}(\mathbf{b}\cdot\mathbf{c})}{|\mathbf{b}\times\mathbf{c}|} - \frac{\mathbf{b}(\mathbf{a}\cdot\mathbf{c})}{|\mathbf{a}\times\mathbf{c}|} \\
&= \frac{\cos(180^\circ-\alpha)}{\sin(180^\circ-\alpha)}\mathbf{a} - \frac{\cos\beta}{\sin\beta}\mathbf{b} \\
&= -(\cot\alpha)\,\mathbf{a} - (\cot\beta)\,\mathbf{b}
\end{aligned}$$

9.2 Circle through three points



- Radius of a circle going through A, B, C is

$$r = \frac{\|AC\|}{2\sin\theta}$$

9.3 Vector Triangle Area

- Given triangle ABC (vertices listed in counterclockwise order), and given any point P, we have

$$AB \times BC = PA \times PB + PB \times PC + PC \times PA$$

(since $PB \times PC = PB \times BC = (PA + AB) \times BC = PA \times BC + AB \times BC$)

- I consider $\frac{1}{2}(AB \times AC)$ to be the "vector triangle area" because it's a vector in the direction of the triangle normal and has magnitude equal to the triangle's area.

9.3.1 Vector area of a triangulated surface

- The vector area of a triangulated surface is the sum of the vector areas of the triangles. To make notation easier, I will avoid the factor of $\frac{1}{2}$ by cheating and looking at *two times* the vector area instead. Using the formula above, we get the total vector area to be:

$$\sum_{(A,B) \in \text{all directed edges}} PA \times PB$$

Now, every interior edge will show up twice (in two different orientations). *e.g.* interior edge AB has two directions: $(A, B)$and (B, A). The cross products for the two orientations cancel each other out ($PA \times PB + PB \times PA = 0$). Hence the result of the sum is

$$\sum_{(A,B) \in \text{directed boundary edges}} PA \times PB$$

- In particular if the surface is closed, there is no boundary and the sum equals 0.
- So the sum of the vector areas of a closed triangulated surface.
- This can also be derived from the Divergence Theorem (or other such theorem) because really what we're doing is taking the surface integral of the surface normal. (The vector area of a triangle equals the normal of that triangle integrated over the triangle's surface).

9.4 Factorial

- Integral definition:

$$n! = \int_0^\infty e^{-x} x^n \, dx$$

9.5 Area of spherical triangle

- From `http://www.math.niu.edu/~rusin/known-math/99/spher_area`
- Does anyone know of a (simple) proof for the formula for the surface area of a spherical triangle Area = R^2 * (A + B + C - Pi).
- The proof that the area of a spherical triangle is the "spherical excess": Extend the sides and note that the sphere is divided into three sets of "orange slice" pairs with area 4*(A+B+C), where A,B, and C are the angles of the triangle. These cover the sphere, but overlap in the triangle and in its inverse image, so the triangle's area is counted four extra times. Hence 4*pi + 4*Area = 4*(A+B+C) and Area = A+B+C - pi.

10 Summation Notation

Matrix/Second-order tensor representation

If A has rows $\mathbf{r}_i$ and columns $\mathbf{c}_j$ then

$$\begin{aligned} A &= A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j \\ &= \mathbf{e}_i \otimes (A_{ij}\mathbf{e}_j) = \mathbf{e}_i \otimes \mathbf{r}_i \\ &= (A_{ij}\mathbf{e}_i) \otimes \mathbf{e}_j = \mathbf{c}_j \otimes \mathbf{e}_j \end{aligned}$$

Epsilon-Delta

$$\begin{aligned} \varepsilon_{ijk}\varepsilon_{lmk} &= \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \\ \varepsilon_{ijk}\varepsilon_{ljk} &= 2\delta_{il} \end{aligned}$$

$$\mathbf{e}_i \times \mathbf{e}_j = \varepsilon_{ijk}\mathbf{e}_k$$

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

To prove this just observe how $(\mathbf{a} \otimes \mathbf{b})(\mathbf{c} \otimes \mathbf{d})$ acts on a vector: $(\mathbf{a} \otimes \mathbf{b})(\mathbf{c} \otimes \mathbf{d})\mathbf{v} = (\mathbf{a} \otimes \mathbf{b})(\mathbf{d} \cdot \mathbf{v})\mathbf{c} = (\mathbf{d} \cdot \mathbf{v})(\mathbf{b} \cdot \mathbf{c})\mathbf{a} = (\mathbf{b} \cdot \mathbf{c})(\mathbf{d} \cdot \mathbf{v})\mathbf{a} = (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \otimes \mathbf{d})\mathbf{v}$

Cofactor:

$$cof\,(A)_{ij} = \frac{1}{2}\varepsilon_{mni}\varepsilon_{pqj}A_{mp}A_{nq}$$

One way to derive this is to recall that if A has rows $\mathbf{r}_i$ then $C = \begin{pmatrix} (\mathbf{r}_2 \times \mathbf{r}_3)^T \\ (\mathbf{r}_3 \times \mathbf{r}_1)^T \\ (\mathbf{r}_1 \times \mathbf{r}_2)^T \end{pmatrix}$.

That is

$$\begin{aligned} C &= \mathbf{e}_1 \otimes (\mathbf{r}_2 \times \mathbf{r}_3) + \mathbf{e}_2 \otimes (\mathbf{r}_3 \times \mathbf{r}_1) + \mathbf{e}_3 \otimes (\mathbf{r}_1 \times \mathbf{r}_2) \\ &= (\mathbf{e}_2 \times \mathbf{e}_3) \otimes (\mathbf{r}_2 \times \mathbf{r}_3) + (\mathbf{e}_3 \times \mathbf{e}_1) \otimes (\mathbf{r}_3 \times \mathbf{r}_1) + (\mathbf{e}_1 \times \mathbf{e}_2) \otimes (\mathbf{r}_1 \times \mathbf{r}_2) \\ &= \frac{1}{2} (\mathbf{e}_i \times \mathbf{e}_j) \otimes (\mathbf{r}_i \times \mathbf{r}_j) \end{aligned}$$

where the factor of $\frac{1}{2}$ is introduced because the summation will count $(\mathbf{e}_i \times \mathbf{e}_j) \otimes (\mathbf{r}_i \times \mathbf{r}_j)$ also as $(\mathbf{e}_j \times \mathbf{e}_i) \otimes (\mathbf{r}_j \times \mathbf{r}_i)$. With $\mathbf{r}_i = A_{ik}\mathbf{e}_k$, $\mathbf{r}_j = A_{jl}\mathbf{e}_l$ we get

$$\begin{aligned} C &= \frac{1}{2} A_{ik} A_{jl} (\mathbf{e}_i \times \mathbf{e}_j) \otimes (\mathbf{e}_k \times \mathbf{e}_l) \\ &= \frac{1}{2} A_{ik} A_{jl} \varepsilon_{ijm} \varepsilon_{kln} \mathbf{e}_m \otimes \mathbf{e}_n \end{aligned}$$

as required to show.

Determinant:

$$det(A) = \frac{1}{6} \varepsilon_{ijk} \varepsilon_{pqr} A_{ip} A_{jq} A_{kr}$$

11 Optimization

11.1 Linear Complementary Problem (LCP)

From Baraff's "Issues in Computing Contact Forces for Non-Penetrating Rigid Bodies"

Find $\mathbf{x}$ satisfying

$$\mathbf{y} = A\mathbf{x} - \mathbf{b} \geq 0,\ \mathbf{x} \geq 0,\ \mathbf{x}^T \mathbf{y} = 0$$

where $A, \mathbf{y}, \mathbf{b}$ are given.

11.2 Quadratic Programming (QP)

From Baraff's "Issues in Computing Contact Forces for Non-Penetrating Rigid Bodies"

Minimize $\mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x}$ subject to $A\mathbf{x} \geq c$ and $\mathbf{x} \geq 0$.

11.3 A note about LCP and QP

Determining whether an LCP has a solution or whether a QP can achieve a certain minimum are NP-complete problems.

However, if A is positive semi-definite, then the problem are known as *convex* problems and they have polynomial time solutions.

12 Quaternions

12.1 Properties

- **Quaternion**: $(s, \mathbf{u}) = s + u_x I + u_y J + u_z K$
- **Multiplication**: $(s, \mathbf{u})(t, \mathbf{v}) = (st - \mathbf{u} \cdot \mathbf{v}, s\mathbf{v} + t\mathbf{u} + \mathbf{u} \times \mathbf{v})$
  - Equivalent to multiplying $(s + u_x I + u_y J + u_z K)(t + v_x I + v_y J + v_z K)$ with the rules $I^2 = J^2 = K^2 = -1$, $IJ = K, JK = I, KI = J$, $JI = -K, KJ = -I, IK = -J$.
  - $(p + q) r = pr + qr$
- Analog to Euler's formula:
  - Note that if $|\mathbf{u}| = 1$ then $(0, \mathbf{u})(0, \mathbf{u}) = (-1, \mathbf{0}) = -\mathbf{1}$ so writing the quaternion as $q = \cos\Theta + \mathbf{u}\sin\Theta$ with $\mathbf{u}^2 = -\mathbf{1}$ (using quaternion multiplication) is analogous to $e^{i\Theta} = \cos\Theta + i\sin\Theta$ with $i^2 = -1$
- **Magnitude**: $|q|^2 = s^2 + \mathbf{u}^2$
  - $|pq| = |p|\,|q|$
- **Conjugate**: $\bar{q} = (s, -\mathbf{u})$
- **Inverse**: $q^{-1} = \frac{\bar{q}}{|q|^2}$
- **Derivative**: $\frac{dq}{dt} = \frac{1}{2}\omega q$
- **Conjugate scaling**(?): (derivation 13.1.1)

$$
\begin{aligned}
qpq^{-1} &= \frac{1}{|q|^2}\left(t\,|q|^2, s^2\mathbf{v} + 2s\,(\mathbf{u} \times \mathbf{v}) + (\mathbf{v} \cdot \mathbf{u})\,\mathbf{u} + \mathbf{u} \times (\mathbf{u} \times \mathbf{v})\right) \\
&= \left(t, \frac{1}{|q|^2}\left(|q|^2 I - 2\left(\mathbf{u}^T\mathbf{u}I - \mathbf{u}\mathbf{u}^T\right) + 2s\mathbf{u}^*\right)\mathbf{v}\right)
\end{aligned}
$$

- **Matrix** corresponding to linear transformation $p \mapsto qpq^{-1}$:

$$
\frac{1}{|q|^2}\begin{pmatrix}
|q|^2 & 0 & 0 & 0 \\
0 & |q|^2 - 2\left(u_y^2 + u_z^2\right) & 2\left(u_x u_y - s u_z\right) & 2\left(u_x u_z + s u_y\right) \\
0 & 2\left(u_x u_y + s u_z\right) & |q|^2 - 2\left(u_x^2 + u_z^2\right) & 2\left(u_y u_z - s u_x\right) \\
0 & 2\left(u_x u_z - s u_y\right) & 2\left(u_y u_z + s u_x\right) & |q|^2 - 2\left(u_x^2 + u_y^2\right)
\end{pmatrix}
$$

- **Rotation**: Let $s = \cos(\theta/2)$ and $|\mathbf{u}| = \sin(\theta/2)$.

12.2 Derivatives

- Let $q = \begin{pmatrix} s \\ \mathbf{u} \end{pmatrix}$ be a quaternion viewed as an element of $\mathbb{R}^4$
- Given a *rotation vector* ω, let $\theta = \frac{|\omega|}{2}, \hat{\omega} = \frac{\omega}{|\omega|}$. Then $s = \cos\theta, \mathbf{u} = \sin\theta\,\hat{\omega}$.

$$\frac{dq}{d(\theta, \hat{\omega})} = \begin{pmatrix} -\sin\theta & 0 \\ \cos\theta\,\hat{\omega} & \sin\theta\, I \end{pmatrix}$$

$$\frac{d(\theta, \hat{\omega})}{d\omega} = \begin{pmatrix} \frac{1}{2}\hat{\omega}^T \\ \frac{1}{|\omega|}\left(I - \hat{\omega}\hat{\omega}^T\right) \end{pmatrix}$$

$$\frac{dq}{d\omega} = \frac{dq}{d(\theta, \hat{\omega})}\frac{d(\theta, \hat{\omega})}{d\omega} = \begin{pmatrix} -\frac{1}{2}\sin\theta\,\hat{\omega}^T \\ \frac{1}{2}\cos\theta\,\hat{\omega}\hat{\omega}^T + \frac{\sin\theta}{|\omega|}\left(I - \hat{\omega}\hat{\omega}^T\right) \end{pmatrix}$$

- Note that this is only valid when the derivative is evaluated away from $\omega = \mathbf{0}$. For $\left.\frac{dq}{d\omega}\right|_{\omega=\mathbf{0}}$, note that $s = \cos\theta = 1 - \frac{(|\omega|/2)^2}{2!} + \ldots$ and $\mathbf{u} = \sin\theta\,\frac{\omega}{|\omega|} = \left(\frac{|\omega|}{2} - \frac{(|\omega|/2)^3}{3!} + \ldots\right)\frac{\omega}{|\omega|} = \left(\frac{1}{2}I - O\left(|\omega|^2\right)\right)\omega$. Hence at $\omega = \mathbf{0}$, $\frac{ds}{d\omega} = \mathbf{0}^T$ and $\frac{d\mathbf{u}}{d\omega} = \frac{1}{2}I$. Thus

$$\left.\frac{dq}{d\omega}\right|_{\omega=\mathbf{0}} = \begin{pmatrix} \mathbf{0}^T \\ \frac{1}{2}I \end{pmatrix}$$

13 Derivations

13.1 Quaternions

13.1.1 Conjugate scaling

We have

$$
\begin{aligned}
qpq^{-1} &= \frac{1}{|q|^2}(s,\mathbf{u})(t,\mathbf{v})(s,-\mathbf{u}) \\
&= \frac{1}{|q|^2}(s,\mathbf{u})(ts+\mathbf{v}\cdot\mathbf{u},-t\mathbf{u}+s\mathbf{v}+\mathbf{u}\times\mathbf{v}) \\
&= \frac{1}{|q|^2}\big(s(ts+\mathbf{v}\cdot\mathbf{u})-\mathbf{u}\cdot(-t\mathbf{u}+s\mathbf{v}+\mathbf{u}\times\mathbf{v}), \\
&\qquad s(-t\mathbf{u}+s\mathbf{v}+\mathbf{u}\times\mathbf{v})+(ts+\mathbf{v}\cdot\mathbf{u})\mathbf{u}+\mathbf{u}\times(-t\mathbf{u}+s\mathbf{v}+\mathbf{u}\times\mathbf{v})\big) \\
&= \frac{1}{|q|^2}\big(ts^2+s(\mathbf{v}\cdot\mathbf{u})+t|\mathbf{u}|^2-s(\mathbf{u}\cdot\mathbf{v})-\mathbf{u}\cdot(\mathbf{u}\times\mathbf{v}), \\
&\qquad -st\mathbf{u}+s^2\mathbf{v}+s(\mathbf{u}\times\mathbf{v})+st\mathbf{u}+(\mathbf{v}\cdot\mathbf{u})\mathbf{u}+s(\mathbf{u}\times\mathbf{v})+\mathbf{u}\times(\mathbf{u}\times\mathbf{v})\big) \\
&= \frac{1}{|q|^2}\left(t|q|^2, s^2\mathbf{v}+2s(\mathbf{u}\times\mathbf{v})+(\mathbf{v}\cdot\mathbf{u})\mathbf{u}+\mathbf{u}\times(\mathbf{u}\times\mathbf{v})\right) \\
&= \frac{1}{|q|^2}\left(t|q|^2, \left(s^2-|\mathbf{u}|^2\right)\mathbf{v}+2s(\mathbf{u}\times\mathbf{v})+2(\mathbf{v}\cdot\mathbf{u})\mathbf{u}\right) \\
&= \left(t, \frac{1}{|q|^2}\left(\left(s^2-|\mathbf{u}|^2\right)I+2\mathbf{u}\mathbf{u}^T+2s\mathbf{u}^*\right)\mathbf{v}\right) \\
&= \left(t, \frac{1}{|q|^2}\left(|q|^2 I-2\left(\mathbf{u}^T\mathbf{u}I-\mathbf{u}\mathbf{u}^T\right)+2s\mathbf{u}^*\right)\mathbf{v}\right)
\end{aligned}
$$

where we used $\mathbf{u}\times(\mathbf{u}\times\mathbf{v}) = (\mathbf{v}\cdot\mathbf{u})\mathbf{u} - |\mathbf{u}|^2\mathbf{v}$ and in the last step we added and subtracted $2|\mathbf{u}|^2$.

13.1.2 3x3 matrix corresponding to quaternion rotation

- Using $qpq^{-1} = \left(t, \frac{1}{|q|^2}\left(|q|^2 I-2\left(\mathbf{u}^T\mathbf{u}I-\mathbf{u}\mathbf{u}^T\right)+2s\mathbf{u}^*\right)\mathbf{v}\right)$ we have
- $\mathbf{u}^T\mathbf{u}I-\mathbf{u}\mathbf{u}^T = \begin{pmatrix} u_y+u_z^2 & -u_xu_y & -u_xu_z \\ -u_yu_x & u_x^2+u_z^2 & -u_yu_z \\ -u_zu_x & -u_zu_y & u_x^2+u_y^2 \end{pmatrix}$
- $\mathbf{u}^* = \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix}$
- So

$$
\begin{aligned}
&\frac{1}{|q|^2}\left(|q|^2 I-2\left(\mathbf{u}^T\mathbf{u}I-\mathbf{u}\mathbf{u}^T\right)+2s\mathbf{u}^*\right) \\
&= \frac{1}{|q|^2}\begin{pmatrix} |q|^2-2\left(u_y^2+u_z^2\right) & 2u_xu_y-2su_z & 2u_xu_z+2su_y \\ 2u_xu_y+2su_z & |q|^2-2\left(u_x^2+u_z^2\right) & 2u_yu_z-2su_x \\ 2u_xu_z-2su_y & 2u_yu_z+2su_x & |q|^2-2\left(u_x^2+u_y^2\right) \end{pmatrix}
\end{aligned}
$$

- Note that when q is normalized this becomes

$$I-2\left(\mathbf{u}^T\mathbf{u}I-\mathbf{u}\mathbf{u}^T\right)+2s\mathbf{u}^* = \begin{pmatrix} 1-2\left(u_y^2+u_z^2\right) & 2u_xu_y-2su_z & 2u_xu_z+2su_y \\ 2u_xu_y+2su_z & 1-2\left(u_x^2+u_z^2\right) & 2u_yu_z-2su_x \\ 2u_xu_z-2su_y & 2u_yu_z+2su_x & 1-2\left(u_x^2+u_y^2\right) \end{pmatrix}$$

13.1.3 Rotation of a vector

- (Assuming a right handed system): Suppose you have a quaternion $q = (s, \mathbf{u})$ which represents a rotation θ about axis $\mathbf{u}$ with $s = \cos(\theta/2)$, and you want to apply this rotation to vector $\mathbf{v}$. Then you compute

$$q\mathbf{v}q^{-1} = \left(0, s^2\mathbf{v} + 2s(\mathbf{u}\times\mathbf{v}) + (\mathbf{v}\cdot\mathbf{u})\mathbf{u} + \mathbf{u}\times(\mathbf{u}\times\mathbf{v})\right)$$

Consider the frame created by $\mathbf{u}, \mathbf{u}\times\mathbf{v}, \mathbf{u}\times(\mathbf{u}\times\mathbf{v})$.

  - Normalizing, let $\hat{\mathbf{i}} = \left(\frac{\mathbf{u}}{|\mathbf{u}|}\right)$, $\hat{\mathbf{j}} = \left(\frac{\mathbf{u}\times\mathbf{v}}{|\mathbf{u}\times\mathbf{v}|}\right)$, $\hat{\mathbf{k}} = \left(\frac{\mathbf{u}\times(\mathbf{u}\times\mathbf{v})}{|\mathbf{u}\times(\mathbf{u}\times\mathbf{v})|}\right)$.
  - So our resultant vector is $\mathbf{v}' = s^2\mathbf{v} + (\mathbf{v}\cdot\mathbf{u})|\mathbf{u}|\hat{\mathbf{i}} + 2s|\mathbf{u}\times\mathbf{v}|\hat{\mathbf{j}} + |\mathbf{u}\times(\mathbf{u}\times\mathbf{v})|\hat{\mathbf{k}}$
  - $\mathbf{v}$ only has components in the $\hat{\mathbf{i}}$ and $\hat{\mathbf{k}}$ directions:

$$\begin{aligned} \mathbf{v} &= \frac{\mathbf{v}\cdot\mathbf{u}}{|\mathbf{u}|}\left(\frac{\mathbf{u}}{|\mathbf{u}|}\right) + \frac{\mathbf{v}\cdot(\mathbf{u}\times(\mathbf{u}\times\mathbf{v}))}{|\mathbf{u}\times(\mathbf{u}\times\mathbf{v})|}\left(\frac{\mathbf{u}\times(\mathbf{u}\times\mathbf{v})}{|\mathbf{u}\times(\mathbf{u}\times\mathbf{v})|}\right) \\ &= \frac{\mathbf{v}\cdot\mathbf{u}}{|\mathbf{u}|}\hat{\mathbf{i}} - \frac{|\mathbf{u}\times\mathbf{v}|^2}{|\mathbf{u}||\mathbf{u}\times\mathbf{v}|}\hat{\mathbf{k}} \\ &= \frac{\mathbf{v}\cdot\mathbf{u}}{|\mathbf{u}|}\hat{\mathbf{i}} - \frac{|\mathbf{u}\times\mathbf{v}|}{|\mathbf{u}|}\hat{\mathbf{k}} \end{aligned}$$

    where we used $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$
  - So

$$\begin{aligned} \mathbf{v}' &= s^2\left(\frac{\mathbf{v}\cdot\mathbf{u}}{|\mathbf{u}|}\hat{\mathbf{i}} - \frac{|\mathbf{u}\times\mathbf{v}|}{|\mathbf{u}|}\hat{\mathbf{k}}\right) + (\mathbf{v}\cdot\mathbf{u})|\mathbf{u}|\hat{\mathbf{i}} + 2s|\mathbf{u}\times\mathbf{v}|\hat{\mathbf{j}} + |\mathbf{u}\times(\mathbf{u}\times\mathbf{v})|\hat{\mathbf{k}} \\ &= \frac{(\mathbf{v}\cdot\mathbf{u})}{|\mathbf{u}|}\left(s^2+|\mathbf{u}|^2\right)\hat{\mathbf{i}} + 2s|\mathbf{u}|\frac{|\mathbf{u}\times\mathbf{v}|}{|\mathbf{u}|}\hat{\mathbf{j}} + \left(|\mathbf{u}|^2 - s^2\right)\frac{|\mathbf{u}\times\mathbf{v}|}{|\mathbf{u}|}\hat{\mathbf{k}} \end{aligned}$$

    and substituting $s = \cos(\theta/2)$, $|\mathbf{u}| = \sin(\theta/2)$, we get

$$\mathbf{v}' = \frac{\mathbf{v}\cdot\mathbf{u}}{|\mathbf{u}|}\hat{\mathbf{i}} + \sin(\theta)\frac{|\mathbf{u}\times\mathbf{v}|}{|\mathbf{u}|}\hat{\mathbf{j}} - \cos(\theta)\frac{|\mathbf{u}\times\mathbf{v}|}{|\mathbf{u}|}\hat{\mathbf{k}}$$

  - Notice that $\mathbf{v}'$ has the same $\hat{\mathbf{i}}$ component as $\mathbf{v}$, and has rotated $\mathbf{v}$'s $\hat{\mathbf{k}}$ component about $\hat{\mathbf{i}}$ in a right-handed fashion

13.1.4 Derivative of a quaternion

Suppose the unit quaternion $q = (s, \mathbf{u})$ represents the orientation of an object. Suppose the object is rotating with angular velocity ω. That is, it is rotating at $|\omega|\, rad/s$ about the axis $\omega/|\omega|$. After a time interval Δt, the object has rotated $\theta = |\omega|\Delta t$ radians about $\hat{\omega} = \omega/|\omega|$. This rotation is represented by the quaternion $(\cos(\theta/2), \sin(\theta/2)\hat{\omega})$. The resulting orientation is $q' = (\cos(\theta/2), \sin(\theta/2)\hat{\omega})\, q$ and we get

$$\frac{q'-q}{\Delta t} = \frac{(\cos(\theta/2)-1, \sin(\theta/2)\hat{\omega})}{\Delta t} q = \left(\frac{\cos(\theta/2)-1}{\Delta t}, \frac{\sin(\theta/2)}{\Delta t}\hat{\omega}\right) q$$

Now $\lim_{\Delta t\to 0} \frac{\cos(|\omega|\Delta t/2)-1}{\Delta t} = \frac{d(\cos(|\omega|t/2)-1)}{dt}(t=0) = 0$ and $\lim_{\Delta t\to 0} \frac{\sin(|\omega|\Delta t/2)}{\Delta t} = \frac{d(\sin(|\omega|t/2))}{dt}(t=0) = \frac{|\omega|}{2}$ so

$$\begin{aligned}\frac{dq}{dt} = \lim_{\Delta t\to 0}\frac{q'-q}{\Delta t} &= \left(0, \frac{|\omega|}{2}\hat{\omega}\right) q \\ &= \frac{1}{2}(0,\omega)\, q\end{aligned}$$

14 Integration Schemes

According to [http://discuss.foresight.org/~pcm/nanocad/0146.html]:

- original Verlet integrator eliminated velocities completely. Leap-frog put them back in, but in between steps. Velocity-Verlet gives them at the "right" time.
- Original Verlet more succeptible to roundoff errors, but hardly any difference between velocity Verlet and leapfrog.
- These three methods are common in Molecular Dynamics

14.1 Verlet [http://www.ch.embnet.org/MD_tutorial/pages/MD.Part1.html]

$$\mathbf{x}(t+\Delta t) = 2\mathbf{x}(t) - \mathbf{x}(t-\Delta t) + \mathbf{a}(t)\Delta t^2 + O\left(\Delta t^4\right)$$

- You can get the velocity using $\mathbf{v}(t) = \frac{\mathbf{x}(t+\Delta t)-\mathbf{x}(t-\Delta t)}{2\Delta t} + O\left(\Delta t^2\right)$ (but with only second order accuracy, this might be undesirable)
- Derived by taking sum of Taylor expansions of $\mathbf{x}(t+\Delta t)$ and $\mathbf{x}(t-\Delta t)$
- Time reversible (due to symmetry in expression)

14.2 Velocity Verlet [http://www.ch.embnet.org/MD_tutorial/pages/MD.Part1

$$\begin{aligned}\mathbf{x}(t+\Delta t) &= \mathbf{x}(t)+\mathbf{v}(t)\Delta t+\frac{1}{2}\mathbf{a}(t)\Delta t^2\\ \mathbf{v}(t+\Delta t) &= \mathbf{v}(t)+\frac{1}{2}(\mathbf{a}(t)+\mathbf{a}(t+\Delta t))\Delta t\end{aligned}$$

- Apparently better "precision" than regular Verlet because in regular Verlet you might get roundoff/precision issues due to taking difference between the two $\mathbf{x}$ values

14.3 Leap-Frog

$$\begin{aligned}\mathbf{x}(t+\Delta t) &= \mathbf{x}(t)+\mathbf{v}\left(t+\frac{1}{2}\Delta t\right)\Delta t\\ \mathbf{v}\left(t+\frac{1}{2}\Delta t\right) &= \mathbf{v}\left(t-\frac{1}{2}\Delta t\right)+\mathbf{a}(t)\Delta t\end{aligned}$$

- Disadvantage is you don't have $\mathbf{v}$ at same time points as $\mathbf{x}$.

15 Second Order Linear ODEs

- Say $\begin{pmatrix} x \\ y \end{pmatrix}_t = A\begin{pmatrix} x \\ y \end{pmatrix}$.

A diagonalizable

- Suppose $A = Q\Lambda Q^{-1}$ is the eigenvalue decomposition, with Q containing the eigenvectors and Λ the eigenvalues. Then $\begin{pmatrix} x \\ y \end{pmatrix}_t = Q\Lambda Q^{-1}\begin{pmatrix} x \\ y \end{pmatrix}$. Let $\begin{pmatrix} u \\ v \end{pmatrix} = Q^{-1}\begin{pmatrix} x \\ y \end{pmatrix}$.

Analytic solution

- So $\begin{pmatrix} u \\ v \end{pmatrix}_t = \Lambda\begin{pmatrix} u \\ v \end{pmatrix}$ which has solution $\begin{pmatrix} u \\ v \end{pmatrix}(t) = e^{t\Lambda}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$. Note that $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$, so we have

$$\begin{pmatrix} u \\ v \end{pmatrix}(t) = e^{t\Lambda}\begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$$

- Using $\begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = Q^{-1}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ we get

$$\begin{pmatrix} x \\ y \end{pmatrix}(t) = Qe^{t\Lambda}Q^{-1}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

Forward Euler

- Forward Euler is $\frac{\begin{pmatrix} x \\ y \end{pmatrix}^{n+1} - \begin{pmatrix} x \\ y \end{pmatrix}^{n}}{\Delta t} = A \begin{pmatrix} x \\ y \end{pmatrix}^n$ which gives $\begin{pmatrix} x \\ y \end{pmatrix}^{n+1} = (I + \Delta t A) \begin{pmatrix} x \\ y \end{pmatrix}^n$. Hence $\begin{pmatrix} x \\ y \end{pmatrix}^n = (I + \Delta t A)^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Now $I + \Delta t A = Q (I + \Delta t \Lambda) Q^{-1}$. So

$$\begin{pmatrix} x \\ y \end{pmatrix}^n = Q (I + \Delta t \Lambda)^n Q^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

Backward Euler

- Get $\begin{pmatrix} x \\ y \end{pmatrix}^n = (I - \Delta t A)^{-n} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = Q (I - \Delta t \Lambda)^{-n} Q^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ with $(I - \Delta t \Lambda)^{-n} = \text{diag}\left\{(1 - \Delta t \lambda_1)^{-n}, (1 - \Delta t \lambda_2)^{-n}\right\}$.

Comparing solutions

- For a general analysis, let $D = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}$, and suppose we have $\begin{pmatrix} x \\ y \end{pmatrix} = QDQ^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Then

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{|Q|} \begin{pmatrix} q_{11} d_{11} + q_{12} d_{21} & q_{11} d_{12} + q_{12} d_{22} \\ q_{21} d_{11} + q_{22} d_{21} & q_{21} d_{12} + q_{22} d_{22} \end{pmatrix} \begin{pmatrix} q_{22} x_0 - q_{12} y_0 \\ -q_{21} x_0 + q_{11} y_0 \end{pmatrix}$$

so

$$\begin{aligned} x = & \frac{1}{|Q|} \big(q_{11} (q_{22} x_0 - q_{12} y_0) d_{11} + q_{11} (-q_{21} x_0 + q_{11} y_0) d_{12} + \\ & q_{12} (q_{22} x_0 - q_{12} y_0) d_{21} + q_{12} (-q_{21} x_0 + q_{11} y_0) d_{22} \big) \end{aligned}$$

- Let $a = \frac{q_{11}(q_{22} x_0 - q_{12} y_0)}{|Q|}$, $b = \frac{q_{12}(-q_{21} x_0 + q_{11} y_0)}{|Q|}$. Note that $a + b = x_0$.
- In both cases we have $d_{12} = d_{21} = 0$. Both the analytic and numerical solutions are of the form

$$x = a d_{11} + b d_{22}$$

- In particular,
  - **Analytic**: $x(t) = a e^{t\lambda_1} + b e^{t\lambda_2}$
  - **Forward Euler**: $x^n = a (1 + \Delta t \lambda_1)^n + b (1 + \Delta t \lambda_2)^n$
  - **Backward Euler**: $x^n = a (1 - \Delta t \lambda_1)^{-n} + b (1 - \Delta t \lambda_2)^{-n}$

A not diagonalizable

- Suppose $A = QJQ^{-1}$ is the Jordan decomposition, with $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

Analytic solution

- End up with $\begin{pmatrix} u \\ v \end{pmatrix}_t = J \begin{pmatrix} u \\ v \end{pmatrix}$. That is, $u_t = \lambda u + v$ and $v_t = \lambda v$. Clearly $v(t) = v_0 e^{t\lambda}$. Then $u(t) = u_0 e^{t\lambda} + tv(t) = u_0 e^{t\lambda} + v_0 t e^{t\lambda}$. That is
$$\begin{pmatrix} u \\ v \end{pmatrix}(t) = \begin{pmatrix} e^{t\lambda} & te^{t\lambda} \\ 0 & e^{t\lambda} \end{pmatrix} \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$$

- Now $\begin{pmatrix} x \\ y \end{pmatrix}(t) = Q \begin{pmatrix} e^{t\lambda} & te^{t\lambda} \\ 0 & e^{t\lambda} \end{pmatrix} Q^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$.

- This has the solution
$$x(t) = x_0 e^{t\lambda} + cte^{t\lambda}$$
where $c = \frac{q_{11}(-q_{21}x_0 + q_{11}y_0)}{|Q|}$.

Forward Euler

- $\begin{pmatrix} x \\ y \end{pmatrix}^n = (I + \Delta t A)^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = Q (I + \Delta t J)^n Q^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$

- Now, $(I + \Delta t J)^n = \begin{pmatrix} 1 + \Delta t\lambda & \Delta t \\ 0 & 1 + \Delta t\lambda \end{pmatrix}^n = \begin{pmatrix} (1 + \Delta t\lambda)^n & n\Delta t (1 + \Delta t\lambda)^{n-1} \\ & (1 + \Delta t\lambda)^n \end{pmatrix}$ (can prove by induction), so in the forward Euler case we get
$$\begin{aligned} x^n &= a(1 + \Delta t\lambda)^n + b(1 + \Delta t\lambda)^n + cn\Delta t (1 + \Delta t\lambda)^{n-1} \\ &= x_0 (1 + \Delta t\lambda)^n + cn\Delta t (1 + \Delta t\lambda)^{n-1} \end{aligned}$$

Backward Euler

- $\begin{pmatrix} x \\ y \end{pmatrix}^n = (I - \Delta t A)^{-n} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = Q (I - \Delta t J)^{-n} Q^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$

- Now $(I - \Delta t J)^{-1} = \begin{pmatrix} (1 - \Delta t\lambda)^{-1} & \Delta t (1 - \Delta t\lambda)^{-2} \\ 0 & (1 - \Delta t\lambda)^{-1} \end{pmatrix}$ and $(I - \Delta t J)^{-n} = \begin{pmatrix} (1 - \Delta t\lambda)^{-n} & n\Delta t (1 - \Delta t\lambda)^{-(n+1)} \\ 0 & (1 - \Delta t\lambda)^{-n} \end{pmatrix}$

- So we get
$$x^n = x_0 (1 - \Delta t\lambda)^{-n} + cn\Delta t (1 - \Delta t\lambda)^{-(n+1)}$$

Summary

- Let $a = \frac{q_{11}(q_{22}x_0 - q_{12}y_0)}{|Q|}, b = \frac{q_{12}(-q_{21}x_0 + q_{11}y_0)}{|Q|}, c = \frac{q_{11}(-q_{21}x_0 + q_{11}y_0)}{|Q|}$.
  - Note $a + b = x_0$.
- A diagonalizable:
  - **Analytic**: $x(t) = ae^{t\lambda_1} + be^{t\lambda_2}$
  - **Forward Euler**: $x^n = a(1 + \Delta t\lambda_1)^n + b(1 + \Delta t\lambda_2)^n$
- A not diagonalizable:
  - **Analytic**: $x(t) = x_0 e^{t\lambda} + cte^{t\lambda}$
  - **Forward Euler**: $x^n = x_0(1 + \Delta t\lambda)^n + cn\Delta t(1 + \Delta t\lambda)^{n-1}$

Damped Harmonic Oscillator

- Suppose $A = \begin{pmatrix} 0 & 1 \\ -k & -b \end{pmatrix}$. This represents the second order ODE $\ddot{x} + b\dot{x} + kx = 0$.
- $\lambda = \frac{-b \pm \sqrt{b^2 - 4k}}{2} = \left(-\frac{b}{2}\right) \pm \sqrt{\left(-\frac{b}{2}\right)^2 - k}$. Let $\beta = -\frac{b}{2}$ and $\gamma = \sqrt{\beta^2 - k}$. Then $\lambda = \beta \pm \gamma$.

Critically damped

- $\gamma = 0$ (*i.e.* $4k = b^2$, or $k = \beta^2$). Then the eigenvalue is $\lambda = \beta$.
- To find Q in $A = QJQ^{-1}$ we want to find a generator for cycle corresponding to the Jordan block. To do this we find a $\mathbf{v}$ such that $(A - \beta I)^2 \mathbf{v} = 0$ but $(A - \beta I)\mathbf{v} \neq 0$. Note that $A = \begin{pmatrix} 0 & 1 \\ -\beta^2 & 2\beta \end{pmatrix}$, so $A - \beta I = \begin{pmatrix} -\beta & 1 \\ -\beta^2 & \beta \end{pmatrix}$. Let $\mathbf{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then $(A - \beta I)\mathbf{v} = \begin{pmatrix} 1 \\ \beta \end{pmatrix} \equiv \mathbf{u}$, and $(A - \beta I)^2 \mathbf{v} = (A - \beta I)\mathbf{u} = 0$. So $A(\mathbf{u} \mid \mathbf{v}) = (\mathbf{u} \mid \mathbf{v}) \begin{pmatrix} \beta & 1 \\ 0 & \beta \end{pmatrix}$.
- *i.e.* $Q = (\mathbf{u} \mid \mathbf{v}) = \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix}$, $|Q| = 1$ and $J = \begin{pmatrix} \beta & 1 \\ 0 & \beta \end{pmatrix}$.
- $c = -\beta x_0 + v_0$
- **Analytic**: $x(t) = x_0 e^{t\beta} + (-\beta x_0 + v_0)\, te^{t\beta}$
  - Note that we need $\beta \leq 0$ to get a solution that doesn't blow up (a physically meaningful solution). (Hence $b \geq 0$)

- **Forward Euler**: $x^n = x_0 (1+\Delta t\beta)^n + (-\beta x_0 + v_0) n\Delta t (1+\Delta t\beta)^{n-1}$
  - Note that we need $|1+\Delta t\beta| \le 1$ to ensure stability. (Combined with the $\beta \le 0$ requirement, we get $1+\Delta t\beta \ge -1$ which gives $\Delta t \le -\frac{2}{\beta}$)

Under/Overdamped

- We have $\lambda = \beta \pm \gamma$
- Note that $A = \begin{pmatrix} 0 & 1 \\ \gamma^2-\beta^2 & 2\beta \end{pmatrix}$. You can verify that $\mathbf{u} = \begin{pmatrix} 1 \\ \beta+\gamma \end{pmatrix}$ is the eigenvector corresponding to $\lambda_1 = \beta+\gamma$, and $\mathbf{v} = \begin{pmatrix} 1 \\ \beta-\gamma \end{pmatrix}$ corresponds to $\lambda_2 = \beta-\gamma$. Hence $Q = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}$ and $|Q| = -2\gamma$.
- So $a = \frac{-\lambda_2 x_0 + v_0}{2\gamma}, b = \frac{\lambda_1 x_0 - v_0}{2\gamma}$ and we get
- **Analytic**: $x(t) = \frac{1}{2\gamma}\left((-\lambda_2 x_0 + v_0) e^{t\lambda_1} + (\lambda_1 x_0 - v_0) e^{t\lambda_2}\right)$
  - $x(t) = \frac{e^{t\beta}}{2\gamma}\left((-\lambda_2 x_0 + v_0) e^{t\gamma} + (\lambda_1 x_0 - v_0) e^{-t\gamma}\right)$
  - In order not to blow up, in general we want $\lambda_1, \lambda_2 \le 0$. This implies $\beta \le 0$ ($b \ge 0$). It also implies $\beta+\gamma \le 0 \Rightarrow \sqrt{\beta^2-k} = \gamma \le -\beta = |\beta| \Rightarrow \beta^2 - k \le \beta^2$ and therefore $k \ge 0$.
  - If $-\lambda_2 x_0 + v_0 = 0$ then we have $(2\gamma - \lambda_1) x_0 + v_0 = 0$. Then $\lambda_1 x_0 - v_0 = 2\gamma x_0$ and we get $x(t) = x_0 e^{t\lambda_2}$. If $\lambda_2 \le 0$ then this gives us a stable solution (even though λ_1 might be > 0).
  - Similarly if $\lambda_1 x_0 - v_0 = 0$ then $(\lambda_2 + 2\gamma) x_0 - v_0 = 0$ so $-\lambda_2 x_0 + v_0 = 2\gamma x_0$ and we get $x(t) = x_0 e^{t\lambda_1}$. If $\lambda_1 \le 0$ then this gives us a stable solution (even though λ_2 might be > 0).
- **Forward Euler**: $x^n = \frac{1}{2\gamma}\left((-\lambda_2 x_0 + v_0)(1+\Delta t\lambda_1)^n + (\lambda_1 x_0 - v_0)(1+\Delta t\lambda_2)^n\right)$
  - Need $|1+\Delta t\lambda_1|, |1+\Delta t\lambda_2| \le 1$ for stability.
- Underdamped case:
  - Say $\gamma = \delta i$.
  - Note that for $|1+\Delta t(\beta+\delta i)| \le 1$ (with $\delta \ne 0$) we definitely need $\beta < 0$. SO WE CANNOT USE FORWARD EULER FOR AN OSCILLATOR WITHOUT DAMPING!!!
  - Then $a = \frac{-\beta x_0 + \delta x_0 i + v_0}{2\delta i} = \frac{x_0}{2} + \frac{\beta x_0 - v_0}{2\delta} i$ and $b = \frac{\beta x_0 + \delta x_0 i - v_0}{2\delta i} = \frac{x_0}{2} - \frac{\beta x_0 - v_0}{2} i$. So that $b = \bar{a}$. Also $\lambda_2 = \overline{\lambda_1}$
  - So $x(t) = ae^{t\lambda_1} + \bar{a}\overline{e^{t\lambda_1}} = 2\mathrm{Re}\left(ae^{t\lambda_1}\right) = 2e^{t\beta}\mathrm{Re}\left(ae^{t\delta i}\right) = 2e^{t\beta}\left(\frac{x_0}{2}\cos t\delta + \frac{(v_0-\beta x_0)}{2\delta}\sin t\delta\right) = e^{t\beta}\left(x_0 \cos t\delta + \frac{v_0 - \beta x_0}{\delta}\sin t\delta\right)$

  - This equals $x(t) = e^{t\beta}\sqrt{x_0^2 + \left(\frac{v_0-\beta x_0}{\delta}\right)^2} \sin\left(t\delta + \tan^{-1}\left(\frac{x_0\delta}{v_0-\beta x_0}\right)\right)$
  - The frequency is $\frac{\delta}{2\pi}$

- Forward Euler: Following a similar analysis we get $1+\Delta t\lambda_2 = \overline{1+\Delta t\lambda_1}$ and so $x^n = 2\mathrm{Re}\left(a\left(1+\Delta t\lambda_1\right)^n\right)$.
  - In polar form, $1+\Delta t\lambda_1 = (1+\Delta t\beta) + \Delta t\delta i = re^{\theta i}$ where $r = \sqrt{(1+\Delta t\beta)^2 + (\Delta t\delta)^2}$ and $\theta = \tan^{-1}\left(\frac{\Delta t\delta}{1+\Delta t\beta}\right)$.
  - Then $\left(1+\Delta t\lambda_1\right)^n = r^n e^{n\theta i}$.
  - The period of oscillation is the amount of time it takes for $n\theta = 2\pi$. Letting $n = T/\Delta t$ we get $T = \frac{2\pi\Delta t}{\theta}$.
  - Compare this to the actual period $T = \frac{2\pi}{\delta}$.
  - Note that for small Δt, $\frac{\Delta t\delta}{1+\Delta t\beta}$ is small and $\theta = \tan^{-1}\left(\frac{\Delta t\delta}{1+\Delta t\beta}\right) \approx \frac{\Delta t\delta}{1+\Delta t\beta}$. Then $T \approx \frac{2\pi}{\delta}(1+\Delta t\beta)$...

- Backward Euler: Similar to above we get $\left(1-\Delta t\lambda_2\right)^{-1} = \overline{\left(1-\Delta t\lambda_1\right)^{-1}}$ and so $x^n = 2\mathrm{Re}\left(a\left(1-\Delta t\lambda_1\right)^{-n}\right)$
  - In polar form we have $\left(1-\Delta t\lambda_1\right)^{-1} = \left(re^{\theta i}\right)^{-1}$ where $r = \sqrt{(1-\Delta t\beta)^2 + (\Delta t\delta)^2}$ and $\theta = \tan^{-1}\left(\frac{-\Delta t\delta}{1-\Delta t\beta}\right)$.
  - Then $\left(1-\Delta t\lambda_1\right)^{-n} = r^{-n}e^{-n\theta i}$
  - The period of oscillation is $T = \frac{2\pi\Delta t}{-\theta}$.

16 Trigonometry

$$A\cos\theta + B\sin\theta = \sqrt{A^2+B^2}\sin\left(\theta + \tan^{-1}(A/B)\right)$$

- Proof: Assume $A\cos\theta + B\sin\theta = C\sin(\theta+\phi)$. Setting $\theta = 0$ gives $A = C\sin\phi$, and setting $\theta = \pi/2$ gives $B = C\cos\phi$. So $C^2 = A^2+B^2$ and $\tan\phi = A/B$.

17 Interpolation

17.1 Divided Differences

- Define recursively $[x_0, x_1, \ldots, x_n] = \frac{[x_1,x_2,\ldots,x_n]-[x_0,x_1,\ldots,x_{n-1}]}{x_n-x_0}$ with $[x_i] = f(x_i)$.

- Then

$$\begin{aligned} P(x) &= [x_0] + (x-x_0)[x_0,x_1] + (x-x_0)(x-x_1)[x_0,x_1,x_2] + \cdots + (x-x_0)(x-x_1)\cdots(x-x_{n-1})[\\ &= \sum_{i=0}^{n} \pi_i(x)[x_0,x_1,\ldots,x_i] \end{aligned}$$

with $\pi_i(x) = \prod_{k=0}^{i-1}(x - x_k)$ interpolates f at the points $x_0, x_1, \ldots, x_n$

18 Fourier Transform

18.1 From [Marsden and Hoffman, ch.10]

- Basic idea:
  - An orthonormal family φ_k
  - A **complete** i.p.s. $\mathcal{V}$: every $f \in \mathcal{V}$ can be written as a **Fourier series** $f = \sum_{k=0}^{\infty} \langle f, \varphi_k \rangle \varphi_k$
- *e.g.* Let $\varphi_n = \frac{1}{\sqrt{2\pi}} e^{inx}$ for $n = 0, \pm 1, \pm 2, \ldots$,$x \in [-\pi, \pi]$ and $\langle f, g \rangle = \int_{-\pi}^{\pi} f\bar{g}$
- Fourier Series for f periodic on $[-\pi, \pi]$: $f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$ with $c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx$
- Fourier Series for f periodic on $[-L, L]$: $f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx\pi/L}$with $c_n = \frac{1}{2L} \int_{-L}^{L} f(x) e^{-inx\pi/L} dx$
  - Let $g(x) = f(Lx/\pi)$. Then g is periodic on $[-\pi, \pi]$ and $f(x) = g(\pi x/L) = \sum_{n=-\infty}^{\infty} c_n e^{i\pi n x/L}$
  - and $c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(Lx/\pi) e^{-inx} dx$
  - Let $x' = Lx/\pi$ then $dx = \pi dx'/L$ and $c_n = \frac{1}{2L} \int_{-L}^{L} f(x') e^{-i\pi n x'/L} dx'$
- Get $\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L}\right)$ with $a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos \frac{n\pi x}{L}\, dx$,$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin \frac{n\pi x}{L}\, dx$

18.2 From [Weisstein]

- **Dirichlet condition**: Piecewise regular function with (1) finite number of finite discontinuities and (2) finite number of extrema.
  - Under these conditions the Fourier series converges (where there's a discontinuity, it converges to average of left and right values)
- **Fourier transform:**
  - In the limit as $L \to \infty$ (replace $c_n$with $F(k)\, dk$ and $n/L \to k$)
  - Start with f periodic on $[-L/2, L/2]$ so that $f(x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x/L}$with $c_n = \frac{1}{L} \int_{-L/2}^{L/2} f(x) e^{-2\pi i n x/L} dx$
  - Let $k = n/L$, so $\Delta k = \frac{1}{L}$ and replace $c_n$with $\Delta k F(k) = \frac{1}{L} \int_{-L}^{L} f(x) e^{-2\pi i k x} dx$

- Then $f(x) = \sum_{n=-\infty}^{\infty} F(k) e^{2\pi ikx} \Delta k$
- As $L \to \infty$, $F(k) \to \int_{-\infty}^{\infty} f(x) e^{-2\pi ikx} dx$ and the sum approaches the integral $\int_{-\infty}^{\infty} F(k) e^{2\pi ikx} dk$.
- Fourier transform recovers original function assuming (1) $\int_{-\infty}^{\infty} |f(x)| dx$ exists, (2) finite number of discontinuities, (3) f is of bounded variation
- $F(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi ikx} dx$
- $f(x) = \int_{-\infty}^{\infty} F(k) e^{2\pi ikx} dk$

- **Discrete Fourier transform**:
  - $F_n = \sum_{k=0}^{N-1} f_k e^{-2\pi ikn/N}$
  - $f_k = \frac{1}{N} \sum_{n=0}^{N-1} F_n e^{2\pi ikn/N}$

19 Terminology

- **Hilbert space**: A vector space with an inner product s.t. the norm induced by the inner product turns it into a complete metric space. (An instance of a Banach space)
  - *e.g.* $\mathbb{R}^n$ or $\mathbb{C}^n$ with standard inner products; L^2 with $\langle f, g \rangle = \int_{-\infty}^{\infty} f\, g\, dx$
- **Banach space**: Complete vector space with a norm (not necc. induced by an inner product).
  - *e.g.* Continuous functions with $\|f\| = sup\,|f|$
- **Sobolev space**: Banach space where norm involves derivatives (or at least something other than just function value)
  - *e.g.* $H^1(a,b)$ is space of functions $\left\{f \mid f \in L^2(a,b), f' \in L^2(a,b)\right\}$ with $|f|_{H^1}^2 = |f|_{L^2}^2 + |f'|_{L^2}^2$
- **Complete metric space**: Every Cauchy sequence converges.
- **Cauchy sequence**: $\{x_n\}$ s.t. $|x_{n+m} - x_n|$ are uniformly small in m and $\to 0$ as $n \to \infty$.
- **Functional**: Map from function space to scalars.
- **Operator**: Map between two function spaces.

References

Phillip Colella and Elbridge Gerry Puckett. *Modern Numerical Methods for Fluid Flow.* URL http://www.rzg.mpg.de/ bds/numerics/cfd-lectures.html.

Friedberg, Insel, and Spence. *Linear Algebra, 3rd ed.*

Golub and van Loan. *Matrix Computations, 3rd ed.*

Horn. *Robot vision.*

Marion and Thornton. *Classical dynamics of particles and systems, 4th ed.*

Marsden and Hoffman. *Elementary classical analysis, 2nd ed.*

Michael L. Minion. A projection method for locally refined grids. *J. Comput. Phys.*, 127(1):158–178, 1996.

Schaum. *Continuum Mechanics.*

H. M. Schey. *Div grad curl and all that, 3rd ed.*

Eric Weisstein. World of mathematics. URL http://mathworld.wolfram.com.

Index