Liouville's Theorem (K. Steffan's method)

In the local region of a particle, the particle density in phase space is constant, provided that the particles move in a general field consisting of magnetic fields and of fields whose forces are independent of velocity.

6-d density function: $f(x, y, z,\ p_x, p_y, p_z;\ t)$.

$$\text{i. e.}\quad d^6N = f(x, y, z,\ p_x, p_y, p_z;\ t)\, dx\, dy\, dz\, dp_x\, dp_y\, dp_z.$$

- Define a 6-d current of the particles moving in phase space:

$$\vec{J}_6 = (f\dot{x}, f\dot{y}, f\dot{z},\ f\dot{p}_x, f\dot{p}_y, f\dot{p}_z) = (f\vec{v}, f\vec{F}).$$

Assuming no particle creation or decay, expect continuity equation:

$$\frac{\partial f}{\partial t} + \nabla_6 \cdot \vec{J}_6 = 0.$$

- To prove we must show $\displaystyle \frac{df}{dt} = \frac{\partial f}{\partial t} + \nabla_6 \cdot \vec{J}_6.$

The rate of change of particles inside an arbitrary volume, V, is equal to the negative of the total flux leaving the volume:

$$\frac{\partial}{\partial t}\int_V f\,d^6X = \int_V \frac{\partial f}{\partial t} d^6X = -\int_{\partial V} \vec{J}_6 \cdot d\vec{S} = -\int_V (\nabla_6 \cdot \vec{J}_6)\,d^6X,$$

where the last equality is the six-dimensional version of the divergence theorem which is a special case of Stokes' theorem in n-dimensions. Since this is true for any arbitrary volume, the integrands of the second and fourth integrals must be identically equal, thus proving the continuity equation.

∇_6 splits into

$$\nabla = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}, \qquad \text{and} \qquad \nabla_p = \hat{p}_x\frac{\partial}{\partial p_x} + \hat{p}_y\frac{\partial}{\partial p_y} + \hat{p}_z\frac{\partial}{\partial p_z}.$$

We now have

$$\begin{aligned}\nabla_6 \cdot \vec{J}_6 &= \nabla \cdot (f\vec{v}) + \nabla_p \cdot (f\vec{F}) \\ &= (\nabla f) \cdot \vec{v} + f(\nabla \cdot \vec{v}) + (\nabla_p f) \cdot \vec{F} + f(\nabla_p \cdot \vec{F}).\end{aligned}$$

But

$$\vec{v} = (p^2c^2 + m^2c^4)^{-\frac{1}{2}}\vec{p}c^2,$$

is independent of the spacial coordinates, so

$$\nabla \cdot \vec{v} = 0.$$

$$\nabla_6 \cdot \vec{J}_6 = (\nabla f) \cdot \vec{v} + f(\nabla \cdot \vec{v}) + (\nabla_p f) \cdot \vec{F} + f(\nabla_p \cdot \vec{F})$$

Note: Grayed-out terms vanish.

The force may be written as

$$\vec{F} = \frac{d\vec{p}}{dt} = q\vec{v} \times \vec{B}(\vec{r}) + \vec{g}(\vec{r}),$$

where $\vec{g}(\vec{r})$ is any velocity independent force on the particle , e. g. $q\vec{E}$.

$$\begin{aligned} \nabla_p \cdot \vec{F} &= q\nabla_p \cdot (\vec{v} \times \vec{B}) + \nabla_p \cdot \vec{g}, \\ &= q\vec{B} \cdot (\nabla_p \times \vec{v}) - q\vec{v} \cdot (\nabla_p \times \vec{B}) \qquad \text{(see Appendix F)} \end{aligned}$$

Examine one component of $\nabla_p \times \vec{v}$:

$$\begin{aligned} \left[\nabla_p \times \left(\frac{\vec{p}}{\sqrt{p^2c^2+m^2c^4}}\right)\right]_z &= \frac{\partial}{\partial p_x}\left(\frac{p_y}{\sqrt{p^2c^2+m^2c^4}}\right) - \frac{\partial}{\partial p_y}\left(\frac{p_x}{\sqrt{p^2c^2+m^2c^4}}\right) \\ &= \frac{-\frac{1}{2}\times 2p_xp_y}{(p^2c^2+m^2c^4)^{-3/2}} - \frac{-\frac{1}{2}\times 2p_yp_x}{(p^2c^2+m^2c^4)^{-3/2}} = 0, \end{aligned}$$

so by the cyclic symmetry of variables, we find

$$\nabla_p \cdot \vec{F} = 0.$$

$$
\begin{aligned}
0 &= \nabla_6 \cdot \vec{J}_6 + \frac{\partial f}{\partial t} \\
&= (\nabla f) \cdot \vec{v} + (\nabla_p f) \cdot \vec{F} \\
&= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} + \frac{\partial f}{\partial p_x}\frac{dp_x}{dt} + \frac{\partial f}{\partial p_y}\frac{dp_y}{dt} + \frac{\partial f}{\partial p_z}\frac{dp_z}{dt} + \frac{\partial f}{\partial t} \\
&= \frac{df}{dt}.
\end{aligned}
$$

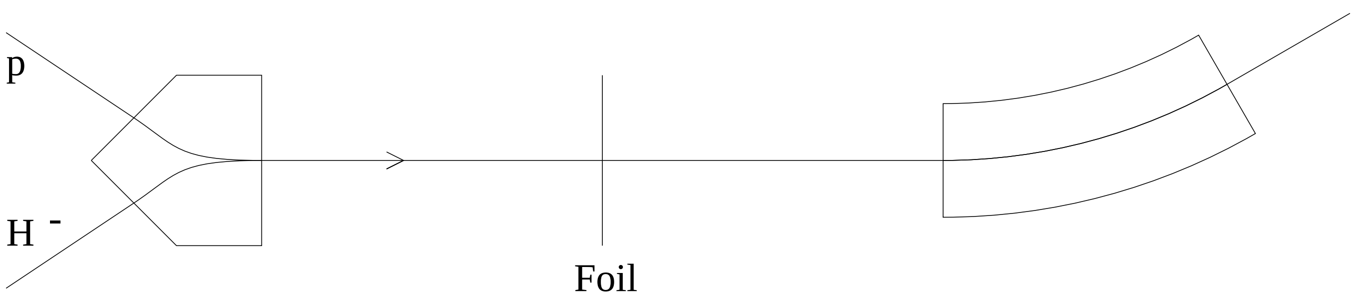
QED

Comments

- Can extend to particle-particle interactions: requires a $6N$-d formalism.
- Smooth space-charge potential:

$$V(\vec{x}) = \frac{q}{4\pi\epsilon_0} \int \frac{f(\vec{x}', \vec{p}')}{|\vec{x} - \vec{x}'|} \, d^3x' \, d^3p',$$

- Liouville's theorem not valid for nonconservative forces:
  - Radiation damping.
  - Passage of beam through materials (foils, targets, gas in beam pipe).
  - H^- injection: electrons stripped by foil.



  - Electron cooling: mixing cold electron beam with hot ion beam.
- Stochastic cooling does not conflict with Liouville's theorem:
  Just removes some holes between particles in the distribution.

General Transformations

Relative to the design particle's trajectory, we may define the initial positon of a test particle by

$$\vec{X} = (\Delta x_0, \Delta p_{x0}, \Delta y_0, \Delta p_{y0}, \Delta z_0, \Delta p_{z0}).$$

where $\Delta x_0 = x_{0,\mathrm{test}} - x_{0,\mathrm{design}}$, etc.
At a later time, the location will be

$$\vec{Y} = (\Delta x_1, \Delta p_{x1}, \Delta y_1, \Delta p_{y1}, \Delta z_1, \Delta p_{z1}).$$

with the similar $\Delta x_1 = x_{1,\mathrm{test}} - x_{1,\mathrm{design}}$, etc. Given the transport function $\vec{T}(\vec{X})$ determined by the force equations on the particles, e. g., $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$, we may write

$$\vec{Y} = \vec{T}(\vec{X}).$$

Expand $\vec{T}(\vec{X})$ in a Taylor series about the design trajectory:

$$\Delta\vec{Y} = \vec{T}(\vec{0}) + \sum_{j=1}^{6} \frac{\partial\vec{T}}{\partial\Delta X_j}(\vec{0})\,\Delta X_j + \frac{1}{2}\sum_{j,k=1}^{6} \frac{\partial^2\vec{T}}{\partial\Delta X_j\,\partial\Delta X_k}(\vec{0})\Delta X_j\,\Delta X_k + \cdots.$$

$\vec{T}(\vec{0}) = \vec{0}$ since it corresponds to the path of the design particle.

Keeping only the linear term

$$\Delta Y_i = \sum_{i=1}^{6} M_{ij}\,\Delta X_j = \begin{pmatrix} \frac{\partial\Delta x_1}{\partial\Delta x_0} & \frac{\partial\Delta x_1}{\partial\Delta p_{x0}} & \cdots \\ \frac{\partial\Delta p_{x1}}{\partial\Delta x_0} & \frac{\partial\Delta p_{x1}}{\partial\Delta p_{x0}} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \Delta x_0 \\ \Delta p_{x0} \\ \vdots \end{pmatrix}.$$

M is simply the Jacobian matrix of the transformation map $\vec{T}$.

- Liouville's theorem tells us that $|\mathbf{M}| = 1$ (i. e., volume is preserved).

Canonical Momentum and Vector Potential

- Work done by a conservative force is independent of path take:

$$\oint \vec{F} \cdot d\vec{r} = 0.$$

  Equivalently by Stoke's theorem: $\nabla \times \vec{F} = 0$.

- Forces depending only on position but not velocity are conservative.

The Lorentz force depends on velocity

$$\frac{d\vec{p}}{dt} = \vec{F} = q(\vec{E} + \vec{v} \times \vec{B}),$$

and is not, in general, conservative.

$$
\begin{aligned}
\nabla \times \vec{F} = \nabla \times \frac{d\vec{p}}{dt} &= q(\nabla \times \vec{E} + \nabla \times (\vec{v} \times \vec{B})) \\
&= -q\frac{\partial \vec{B}}{\partial t} + q\left[(\vec{B}\cdot\nabla)\vec{v} - (\vec{v}\cdot\nabla)\vec{B} + (\nabla\cdot\vec{B})\vec{v} - (\nabla\cdot\vec{v})\vec{B}\right] \\
&= -q\frac{\partial \vec{B}}{\partial t} - q\left[(\vec{v}\cdot\nabla)\vec{B}\right] \\
&= -q\left[\frac{\partial \vec{B}}{\partial t} + \frac{\partial \vec{B}}{\partial x}\frac{dx}{dt} + \frac{\partial \vec{B}}{\partial y}\frac{dy}{dt} + \frac{\partial \vec{B}}{\partial z}\frac{dz}{dt}\right] \\
&= -q\frac{d\vec{B}}{dt} = -\frac{d}{dt}(\nabla \times q\vec{A}),
\end{aligned}
$$

where is a vector potential such that $\vec{B} = \nabla \times \vec{A}$. Rearranging and swapping the order of differentiations:

$$
\nabla \times \left[\frac{d}{dt}\left(\vec{p} + q\vec{A}\right)\right] = 0.
$$

If we define a new *canonical momentum* by

$$\boxed{\vec{P} = \vec{p} + q\vec{A}}$$

then the corresponding *canonical force*

$$\vec{F}_{\rm can} = \frac{d\vec{P}}{dt}$$

is conservative.

We should note that the choice of vector potential is not unique, since the gradient of any scalar function $\phi(x, y, z)$ of spatial coordinates may be added to $\vec{A}$ without changing $\vec{B}$, i. e.,

$$\nabla \times (\nabla\phi) = 0.$$

Note: I tend to refer to the momentum $\vec{p} = \gamma\vec{\beta}mc$ as the "kinetic momentum" when it differs from the canonical momentum.

Relativistic Hamiltonian

For a free (no fields) particle with rest mass, m, charge, q, and kinetic momentum $\vec{p} = \gamma m \vec{v}$, we may write the relativistic Hamiltonian as

$$H = \sqrt{p^2c^2 + m^2c^4}.$$

In an electromagnetic field with potentials $\vec{A}(\vec{x},t)$ and $\phi(\vec{x},t)$ such that

$$\vec{B} = \nabla \times \vec{A}, \qquad \text{and} \qquad \vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t},$$

the Hamiltonian can be given interms of canonical momentum, $\vec{P} = \vec{p} + q\vec{A}(\vec{x},t)$, by

$$H = \sqrt{(\vec{P} - q\vec{A})^2c^2 + m^2c^4} + q\phi.$$

Hamilton's Equations

Using Hamilton's equations, the equations of motion can be written as

$$\frac{d\vec{P}}{dt} = -\nabla H, \quad \text{and} \quad \frac{d\vec{x}}{dt} = \nabla_P H.$$

Coordinates revisited

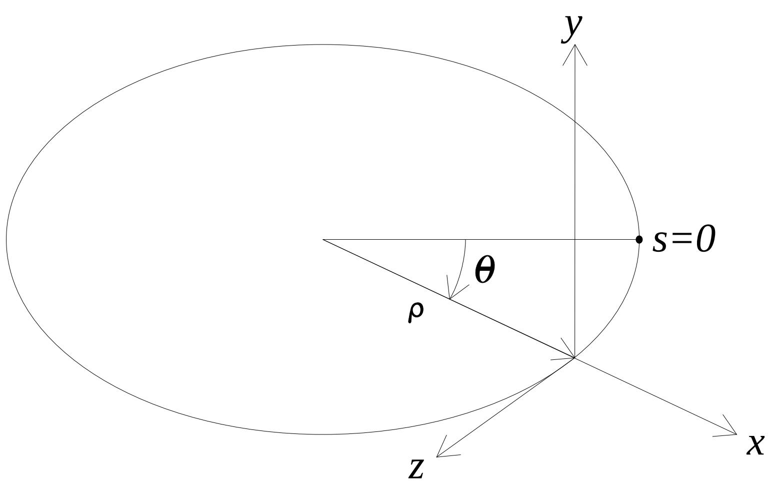


Fig. 2.1: $(\rho, y, \theta) \rightarrow (x, y, z)$

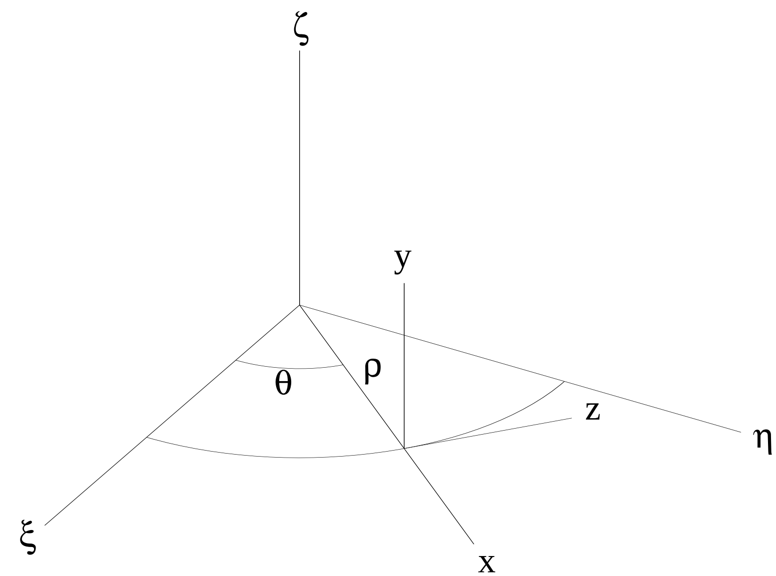


Fig. 3.2: $(\rho, \theta, y) \rightarrow (x, y, z)$

Normally cylindrical coordinates are written in the order $(\rho, \theta, \mathrm{up} = y)$ for a right-handed orientation, whereas Cartesian coordinates are written in the order $(x, \mathrm{up}{=}y, z)$ for a right-handed system.

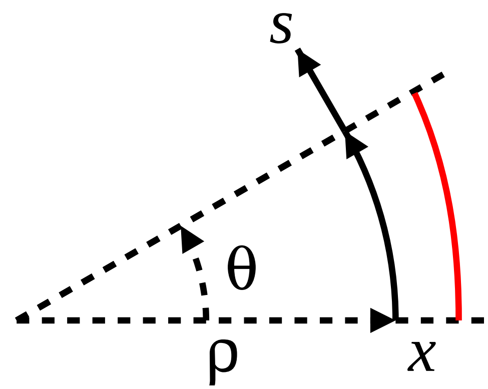
- Could use z for "up", but lots of accelerator physicists use y.
- Either figures requires a left-handed orientation for one of the two systems (cylindrical or Cartesian).

Care must be taken when transforming from the cylindrical system to the local system. The velocity of the particle is

$$\vec{v} = \frac{d\vec{x}}{dt} = \frac{d}{dt}(r\hat{r} + y\hat{y}) = \dot{r}\hat{r} + r\dot{\theta}\hat{\theta} + \dot{y}\hat{y}.$$

The kinetic momentum is then

$$\begin{aligned}\vec{p} = \gamma m\vec{v} &= \gamma m(\dot{r}\hat{r} + r\dot{\theta}\hat{\theta} + \dot{y}\hat{y}) \\ &= \quad p_r\hat{r} + p_\theta\hat{\theta} + p_y\hat{y}.\end{aligned}$$



In terms of the design trajectory radius, ρ, we may write

$$r = \rho + x, \qquad \text{and} \qquad s = \rho\theta,$$

with x being the excess distance of our particle from the origin. The momentum conjugate to x is just $p_x = p_r$, since ρ may be taken to be constant in piecewise sections of the accelerator. It can be shown that a momentum coordinate which is conjugate to the coordinate s is given by

$$p_s = \left(1 + \frac{x}{\rho}\right)(\vec{p}\cdot\hat{s}).$$

From Fig. 3.2 we have for old coordinates in terms of the new ones ($\vec{X}$):

$$\theta = \frac{s}{\rho}, \qquad \xi = (\rho + x)\cos\theta, \qquad \text{and} \qquad \eta = (\rho + x)\sin\theta,$$

and we want to construct $F_3(\vec{p}, \vec{X})$ (see CM: Appendix C) with

$$\xi = (\rho + x)\cos\frac{s}{\rho} = -\frac{\partial F_3}{\partial p_\xi},$$
$$\eta = (\rho + x)\sin\frac{s}{\rho} = -\frac{\partial F_3}{\partial p_\eta},$$

Integrating the two previous partial derivatives and comparing constants yields:

$$F_3(p_\xi, p_\eta, x, s) = -(\rho + x)\left(p_\xi \cos\frac{s}{\rho} + p_\eta \sin\frac{s}{\rho}\right),$$

so

$$P_\xi = -\frac{\partial F_3}{\partial x} = p_\xi \cos\theta + p_\eta \sin\theta = p_r = \vec{p}\cdot\hat{x},$$
$$P_s = -\frac{\partial F_3}{\partial s} = \left(1 + \frac{x}{\rho}\right)(-p_\xi \sin\theta + p_\eta \cos\theta) = \left(1 + \frac{x}{\rho}\right)\vec{p}\cdot\hat{s}. \qquad \checkmark$$

Modifying the Hamiltonian

In our curvilinear system

$$H = \sqrt{m^2c^4 + c^2\left[p_x^2 + p_y^2 + \left(\frac{p_s}{1+x/\rho}\right)^2\right]}.$$

Writing this in terms of the canonical momenta,

$$H = c\sqrt{(P_x - qA_x)^2 + (P_y - qA_y)^2 + \left(\frac{P_s - qA_s}{1+x/\rho}\right)^2 + m^2c^2} \; + q\phi,$$

where we define

$$A_s = \left(1 + \frac{x}{\rho}\right)(\vec{A}\cdot\hat{\theta}).$$

Interchange $(t, -H)$ with (s, p_s):

> A canonical transformation from the variables $(\vec{q}, \vec{p})$ to the variables $(\vec{Q}, \vec{P})$ preserves the *integral invariant of Poincaré–Cartan* $\vec{p} \cdot d\vec{q} - H\,dt = \vec{P} \cdot d\vec{Q} - K\,dt$, where $K(\vec{Q}, \vec{P}; t)$ is the Hamiltonian expressed in terms of the new coordinates. If some coordinate q_j increases in time, then we may rewrite this invariant as
>
> $$\left(\sum_{i \neq j} p_i dq_i + (-H)dt\right) - (-p_j)dq_j.$$

The conjugate momentum corresponding to t is now $-H$, which is just the negative of the total energy U, so the new Hamiltonian is

$$\begin{aligned} \mathrm{H} &= -P_s(x, P_x, y, P_y, t, -U; s) \\ &= -qA_s - \left(1 + \frac{x}{\rho}\right)\sqrt{\left(\frac{U - q\phi}{c}\right)^2 - m^2c^2 - (P_x - qA_x)^2 - (P_y - qA_y)^2}\,. \end{aligned}$$

In the cylindrical coordinates (see Appendix F) (r, θ, y), the curl of $\vec{A}$ is

$$\nabla \times \vec{A} = \left(\frac{1}{r}\frac{\partial A_y}{\partial \theta} - \frac{\partial A_\theta}{\partial y}\right)\hat{r} + \left(\frac{\partial A_r}{\partial y} - \frac{\partial A_y}{\partial r}\right)\hat{\theta} + \frac{1}{r}\left(\frac{\partial}{\partial r}(rA_\theta) - \frac{\partial A_r}{\partial \theta}\right)\hat{y}.$$

Transforming to the local coordinate system, (x, y, s), gives

$$\nabla \times \vec{A} = \frac{1}{1+x/\rho}\left(\frac{\partial A_s}{\partial y} - \frac{\partial A_y}{\partial s}\right)\hat{x} + \frac{1}{1+x/\rho}\left(\frac{\partial A_x}{\partial s} - \frac{\partial A_s}{\partial x}\right)\hat{y}$$
$$+ \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\hat{s}.$$

Note: an overall minus sign is applied since we have interchanged the position of the θ-like coordinate with the vertical coordinate (as mentioned back on p. 14).

For static magnetic fields which are transverse to the direction of motion, we may write:

$$\phi = 0, \qquad , A_x = 0, \qquad \text{and} \qquad A_y = 0,$$

$$\therefore \qquad P_x = p_x = \gamma\beta_x mc, \qquad \text{and} \qquad P_y = p_y = \gamma\beta_x mc.$$

- For the transverse components of magnetic field

$$B_x = \frac{1}{1+x/\rho}\frac{\partial A_s}{\partial y}, \qquad \text{and} \qquad B_y = -\frac{1}{1+x/\rho}\frac{\partial A_s}{\partial x}.$$

- For nonzero longitudinal components (as in solenoids), life becomes a bit more complicated as we shall see.

Symplectic Transforms and Matrices

In accelerator physics it is usually the convention to order the canonical coordinates in a vector of coordinate–momentum pairs:

$$\mathbf{X} = \begin{pmatrix} x \\ P_x \\ y \\ P_y \\ t \\ -U \end{pmatrix}.$$

Recall Hamilton's equations for

$$\mathrm{H} = \mathrm{H}(x, P_x, y, P_y, t, -U; s):$$

$$\frac{dx_i}{ds} = \frac{\partial \mathrm{H}}{\partial P_i}, \qquad \text{and} \qquad \frac{dP_i}{ds} = -\frac{\partial \mathrm{H}}{\partial x_i}.$$

We can write this in a matrix form as

$$\frac{dX_i}{dt} = \sum_{i=1}^{6} \mathbf{S}_{ij} \frac{\partial \mathrm{H}}{\partial X_j}, \qquad \text{with}$$

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

Note that:

- $|\mathbf{S}| = 1$,
- $\mathbf{S}^2 = -\mathbf{I}$,
- $\mathbf{S}^{\mathrm{T}} = -\mathbf{S}$,
- $\mathbf{S}^{-1} = -\mathbf{S}$.

Consider the transport of a beam from s_0 to s_1, $\mathbf{T}: \mathbf{X}_0 \to \mathbf{X}_1$.
The slopes of the components may be written as

$$\frac{dX_{1i}}{ds} = \sum_j \mathbf{M}_{ij} \frac{dX_{0j}}{ds}, \qquad \text{where} \qquad \mathbf{M}_{ij} = \frac{\partial X_{1i}}{\partial X_{0j}}$$

is the Jacobian matrix for $\mathbf{T}$.
Since the particle motion is reversible, we also have $\mathbf{T}^{-1}: \mathbf{X}_1 \to \mathbf{X}_0$ with slopes

$$\frac{dX_{1i}}{ds} = \sum_j (\mathbf{M})^{-1}_{ij} \frac{dX_{0j}}{ds}, \qquad \text{where} \qquad (\mathbf{M})^{-1}_{ij} = \frac{\partial X_{0i}}{\partial X_{1j}}$$

is the Jacobian matrix for $\mathbf{T}^{-1}$.

$$\frac{dX_{1i}}{ds} = \sum_j \mathbf{S}_{ij} \frac{\partial \mathrm{H}}{\partial X_{1j}} = \sum_{j,k} \mathbf{S}_{ij} \frac{\partial \mathrm{H}}{\partial X_{0k}} \frac{\partial X_{0k}}{\partial X_{1j}} = \sum_{j,k} \mathbf{S}_{ij} \left((\mathbf{M}^{-1})^{\mathrm{T}} \right)_{jk} \frac{\partial \mathrm{H}}{\partial X_{0k}}.$$

In matrix notation this looks like $$\frac{\partial \mathbf{X}_1}{ds} = \mathbf{S}(\mathbf{M}^{-1})^{\mathrm{T}} \nabla_{\vec{\mathrm{X}}_0} \mathrm{H}.$$

But we also have

$$\frac{dX_{0i}}{ds} = \sum_k \mathbf{S}_{ij} \frac{\partial \mathrm{H}}{\partial X_{0j}}, \qquad \text{or in matrix notation} \qquad \frac{d\mathbf{X}_0}{ds} = \mathbf{S}\nabla_{\vec{\mathrm{X}}_0}\mathrm{H}$$

or solving for the divergence of H:

$$\nabla_{\vec{\mathrm{X}}_0}\mathrm{H} = \mathbf{S}^{-1}\frac{d\mathbf{X}_0}{ds}$$

$$\frac{d\mathbf{X}_1}{ds} = \mathbf{M}\,\frac{d\mathbf{X}_0}{ds} = \mathbf{S}(\mathbf{M}^{-1})^{\mathrm{T}}\mathbf{S}^{-1}\frac{d\mathbf{X}_0}{ds}$$

So it must be that

$$\begin{aligned}\mathbf{M} &= \mathbf{S}(\mathbf{M}^{-1})^{\mathrm{T}}\mathbf{S}^{-1} = \mathbf{S}(\mathbf{M}^{\mathrm{T}})^{-1}\mathbf{S}^{-1},\\ &= \mathbf{S}^{-1}(\mathbf{M}^{\mathrm{T}})^{-1}\mathbf{S},\end{aligned}$$

- Symplectic condition: $\boxed{\mathbf{S} = \mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M} = \mathbf{M}\mathbf{S}\mathbf{M}^{\mathrm{T}},}$

- A matrix $\mathbf{M}$ which satisfies $\mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M} = \mathbf{S}$ is called a *symplectic* matrix.
- Note also that $\mathbf{S}$ is symplectic since $\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{S} = (-\mathbf{S})(-\mathbf{I}) = \mathbf{S}$.

Generalize to $2n$ dimensions where $\mathbf{S}$ is a $2n \times 2n$ matrix with 2×2-blocks

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

along the diagonal and zeroes* everywhere else.

From the symplectic condition: $1 = |\mathbf{S}| = |\mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M}| = |\mathbf{M}|^2 \quad \Rightarrow \quad |\mathbf{M}| = \pm 1.$

- Actually, $|\mathbf{M}| = +1$ only.
  - "Simple" proof using Brioschi's theorem:

    For any antisymmetric matrix $\mathbf{A}$, $\mathrm{Pf}(\mathbf{M}^{\mathrm{T}}\mathbf{A}\mathbf{M}) = \det(\mathbf{M})\mathrm{Pf}(\mathbf{A})$.

    Taking $\mathbf{A} = \mathbf{S}$ we then $|\mathbf{M}| = \mathrm{Pf}(\mathbf{S}) = 1$.

    (Simple except, what the #&*!$ is a pfaffian?)

- A.Dragt, *Lie Methods for Dynamics with Applications to Accelerator Physics* (2011) for different proof. (`http://www.physics.umd.edu/dsat/`)

---

* With an "e" in honour to Dan Quayle aka Mr. Potatoe.

- Symplectic matrices form an algebraic group $\mathrm{Sp}(2n, \mathbb{R})$.
  - Since the representation depends on the form of $\mathbf{S}$, I usually write $\mathrm{Sp}(2n, \mathbb{R}; \mathbf{S})$ for our particular representation.
  - Another common representation in use for a different ordering of the canonical coordinate-momentum vector is $\mathrm{Sp}(2n, \mathbb{R}, \mathbf{E})$ for

$$\mathbf{X} = \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ P_x \\ P_y \\ P_z \end{pmatrix}, \qquad \text{with} \qquad \mathbf{E} = \begin{pmatrix} \mathbf{0}_3 & \mathbf{I}_3 \\ -\mathbf{I}_3 & \mathbf{0}_3 \end{pmatrix},$$

$$\text{where} \qquad \mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \text{and} \qquad \mathbf{0}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

  - Note that in general $\mathrm{Sp}(2n, \mathbb{R}; \mathbf{S}) \neq \mathrm{Sp}(2n, \mathbb{R}, \mathbf{E})$.

Symplectic inverse and conjugate

The inverse of a symplectic matrix $\mathbf{M}$ can easily be found by

$$\begin{aligned} \mathbf{S} &= \mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M}, \\ \mathbf{S}\mathbf{S}\mathbf{M}^{-1} &= \mathbf{S}\mathbf{M}^{\mathrm{T}}\mathbf{S}, \\ \mathbf{M}^{-1} &= -\mathbf{S}\mathbf{M}^{\mathrm{T}}\mathbf{S} = \mathbf{S}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{S}. \end{aligned}$$

It is convenient to define the *symplectic conjugate* of a square matrix $\mathbf{N}$ of the same size (same # of rows and columns) by

$$\widetilde{\mathbf{N}} = \mathbf{S}^{\mathrm{T}}\mathbf{N}^{\mathrm{T}}\mathbf{S}.$$

Clearly $\mathbf{M}^{-1} = \widetilde{\mathbf{M}}$, if $\mathbf{M}$ is symplectic.

Useful 2-d matrix info

Consider a general 2×2 matrix $\quad \mathbf{N} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$

$$\widetilde{\mathbf{N}} = \mathbf{S}^{\mathrm{T}}\mathbf{N}^{\mathrm{T}}\mathbf{S} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

- If $|\mathbf{N}| \neq 0$, then $\qquad \mathbf{N}^{-1} = \dfrac{\widetilde{\mathbf{N}}}{|\mathbf{N}|} = \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$
  - Note: While this simple inverse formula works for 2×2 matrices, it won't work for higher order matrices in general; only the symplectic matrices.
- If a 2×2 matrix has a determinant of 1, then it is symplectic.
  - Not in general true for 4×4 and higher order matrices.

More transformations of Hamiltonian

Let us rescale our previous Hamiltonian (from p. 18)

$$\mathrm{H} = -qA_s - \left(1+\frac{x}{\rho}\right)\sqrt{\left(\frac{U-q\phi}{c}\right)^2 - m^2c^2 - (P_x - qA_x)^2 - (P_y - qA_y)^2}\,,$$

by dividing by the design kinetic momentum p_0:

$$\frac{\mathrm{H}}{p_0} = -\frac{qA_s}{p_0} - \left(1+\frac{x}{\rho}\right)\sqrt{\left(\frac{U-q\phi}{p_0c}\right)^2 - \left(\frac{mc}{p_0}\right)^2 - \left(\frac{P_x - qA_x}{p_0}\right)^2 - \left(\frac{P_y - qA_y}{p_0}\right)^2}.$$

(This is not rigorously a cannonical transformation — just a change of units.)

Consider transport through transverse magnetic field only.
Then we may use $A_x = A_y = 0 \quad \Rightarrow \quad P_x = p_x$, and $P_y = p_y$.
Also we have $\phi = 0$ and $\vec{E} = 0$.

$$\frac{\mathrm{H}}{p_0} = -\frac{q}{p_0}A_s - \left(1+\frac{x}{\rho}\right)\sqrt{\left(\frac{p}{p_0}\right)^2 - \left(\frac{p_x}{p_0}\right)^2 - \left(\frac{p_y}{p_0}\right)^2}.$$

Define new rescaled Hamiltonian

$$\begin{aligned} &H_1(x, w_x, y, w_y, t, -U/p_0; s) \\ &\qquad = -\frac{q}{p_0}A_s - \left(1+\frac{x}{\rho}\right)\sqrt{\left(\frac{U}{p_0 c}\right)^2 - \left(\frac{mc}{p_0}\right)^2 - w_x^2 - w_y^2}, \end{aligned}$$

where $w_x = \frac{p_x}{p_0}$, and $w_y = \frac{p_y}{p_0}$.

Want to use coords relative to design traj.: $\Delta t = t - t_0$, and $-\Delta U = U_0 - U$.

Define the fractional momentum deviation: $\delta = \frac{\Delta p}{p_0} = \frac{p - p_0}{p_0}$,

which we would like to use as a new canonical momentum conjugate to some new coordinate z, yet to be determined.

Construct another generating function from Appendix C:

$$F_2(\text{old coord, new mom; indep var}): \qquad F_2(t,\delta;s).$$

$$\frac{\partial F_2}{\partial \delta} = z \qquad \text{and} \qquad \frac{\partial F_2}{\partial t} = -\frac{U}{p_0}.$$

We may write $\delta = \dfrac{\Delta p}{p_0} = \dfrac{U^2}{p_0^2 c^2}\dfrac{\Delta U}{U_0} = \dfrac{1}{\beta_0^2}\dfrac{\Delta U}{U_0}$, so then

$$\frac{\partial F_2}{\partial t} = -\frac{U}{p_0} = -\frac{c}{\beta_0}\frac{\Delta U}{U_0} = -\frac{c}{\beta_0}(1+\beta_0^2\delta).$$

Integrating with respect to t and setting $t_0 = s/v_0$ gives

$$F_2(t,\delta;s) = -\frac{ct}{\beta_0}(1+\beta_0^2\delta) + g(\delta;s) = \frac{c}{\beta_0}(1+\beta_0^2\delta)(t_0 - t) + g(\delta;s)$$
$$= \frac{1+\beta_0^2\delta}{\beta_0^2}(s - v_0 t) + g(\delta;s),$$

so our new longitudinal coordinate is $z = \dfrac{\partial F_2}{\partial \delta} = s - v_0 t + \dfrac{\partial g}{\partial \delta}$, but $g(\delta;s)$ is an arbitrary integration constant which we are free to choose.

$$\mathcal{H} = \mathcal{H}(x, w_x, y, w_y, z, \delta; s) = H_1(x, w_x, y, w_y, t, -U/p_0; s) + \frac{\partial F_2(t, \delta; s)}{\partial s}$$

$$\frac{\partial F_2}{\partial s} = \frac{1}{\beta_0^2} + \delta + \frac{\partial g}{\partial s}.$$

$$\begin{aligned}
\mathcal{H} &\simeq \mathcal{H}(x, w_x, y, w_y, z, \delta; s) \\
&= -\frac{qA_s}{p_0} - \left(1 + \frac{x}{\rho}\right)\sqrt{\left(\frac{U - mc^2}{p_0 c}\right)^2 - \left(\frac{mc}{p_0 c}\right)^2 - w_x^2 - w_y^2} \quad + \frac{\partial F_2}{\partial s} \\
&= -\frac{qA_s}{p_0} - \left(1 + \frac{x}{\rho}\right)\sqrt{(1+\delta)^2 - w_x^2 - w_y^2} \quad + \delta + \frac{1}{\beta_0^2} + \frac{\partial g}{\partial s}.
\end{aligned}$$

Expand square root up to quadratic terms:

$$\begin{aligned}
\sqrt{1 + 2\delta + \delta^2 - w_x^2 - w_y^2} &= 1 + \frac{1}{2}[2\delta + \delta^2 - w_x^2 - w_y^2 + \cdots] - \frac{1}{8}[4\delta^2 + \cdots] \\
&= 1 + \delta - \frac{1}{2}(w_x^2 + w_y^2) + \cdots.
\end{aligned}$$

$$
\begin{aligned}
\mathcal{H} &\simeq -\frac{qA_s}{p_0} - \left(1+\frac{x}{\rho}\right)\left[1+\delta-\frac{1}{2}(w_x^2+w_y^2)+\cdots\right] \quad +\delta+\frac{1}{\beta_0^2}+\frac{\partial g}{\partial s} \\
&= -\frac{qA_s}{p_0} - (1+\delta) - \frac{x}{\rho} - \frac{x\delta}{\rho} + \frac{1}{2}(w_x^2+w_y^2)+\cdots \quad +\delta+\frac{1}{\beta_0^2}+\frac{\partial g}{\partial s} \\
&= -\frac{qA_s}{p_0} - \frac{x}{\rho} - \frac{x\delta}{\rho} + \frac{1}{2}(w_x^2+w_y^2)+\cdots \quad -1+\frac{1}{\beta_0^2}+\quad\frac{\partial g}{\partial s}
\end{aligned}
$$

We can cancel the irrelevant constant terms by taking

$$
g(s) = \frac{s}{\beta_0^2} - s.
$$

Paraxial approximation:

$$
w_x = \frac{p_x}{p_0} \simeq \frac{p_x}{p_z} = \frac{dx}{ds} = x', \quad w_y = \frac{p_y}{p_0} \simeq \frac{p_y}{p_z} = \frac{dy}{ds} = y'.
$$

Symplectic generators (integration)

Consider a magnetic field which is independent of s, i. e. $A_s = A_s(x, y)$.

$$\mathbf{X}(s) = \mathbf{M}(s)\mathbf{X}_0.$$

For an infinitesimal step ds, we will have

$$\mathbf{M}(ds) = \mathbf{I} + \mathbf{G}\,ds. \qquad (\mathbf{G} \text{ called a generator})$$

In case of s-indenpendent field, $\mathbf{M}(ds)$ does not change along the trajectory.

Total transformation matrix $\mathbf{M}(s)$ may be written as

$$\mathbf{M}(s) = \lim_{n\to\infty} \left(\mathbf{I} + \mathbf{G}\,\frac{s}{n}\right)^n = e^{\mathbf{G}s}.$$

Integration without an integral sign. You may recall something like this from the treatment of spin rotation matrices in quantum mechanics.

Dipole magnet example

Uniform vertical field: $\vec{B} = B_0\hat{y}$.
Recall from p. 20:

$$B_y = -\frac{1}{1+x/\rho}\frac{\partial A_s}{\partial x}.$$

$$A_s = -B_0\left(x + \frac{x^2}{2\rho}\right).$$

$$\mathcal{H} \simeq \frac{qB_0}{p_0}\left(x+\frac{x^2}{2\rho}\right) - \frac{x}{\rho} - \frac{x\delta}{\rho} + \frac{1}{2}(x'^2+y'^2) = \frac{x^2}{2\rho^2} - \frac{x\delta}{\rho} + \frac{1}{2}(x'^2+y'^2).$$

Hamiton's equations of motion:

$$\frac{dx}{ds} = \frac{\partial\mathcal{H}}{\partial x'}, \quad \frac{dx'}{ds} = -\frac{\partial\mathcal{H}}{\partial x},$$
$$\frac{dy}{ds} = \frac{\partial\mathcal{H}}{\partial y'}, \quad \frac{dy'}{ds} = -\frac{\partial\mathcal{H}}{\partial y}$$
$$\frac{dz}{ds} = \frac{\partial\mathcal{H}}{\partial \delta}, \quad \frac{d\delta}{ds} = -\frac{\partial\mathcal{H}}{\partial z}.$$

The *longitudinal* time-like coordinate is a negative deviation in path length from the synchronous particle, $z = v(t_0 - t) = s - vt.$

$$\begin{aligned}
\frac{dx}{ds} &= x', \\
\frac{dx'}{ds} &= -\frac{x}{\rho^2} + \frac{\delta}{\rho}, \\
\frac{dy}{ds} &= y', \\
\frac{dy'}{ds} &= 0, \\
\frac{dz}{ds} &= -\frac{x}{\rho}, \\
\frac{d\delta}{ds} &= 0.
\end{aligned}$$

$$
\begin{pmatrix} x_1 \\ x_1' \\ y_1 \\ y_1' \\ z_1 \\ \delta_1 \end{pmatrix} = \begin{pmatrix} 1 & ds & 0 & 0 & 0 & 0 \\ \frac{-ds}{\rho^2} & 1 & 0 & 0 & 0 & \frac{ds}{\rho} \\ 0 & 0 & 1 & ds & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\frac{ds}{\rho} & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \\ y_0 \\ y_0' \\ z_0 \\ \delta_0 \end{pmatrix}
$$

$$
= (\mathbf{I} + \mathbf{G}\, ds)\, \mathbf{X}_0.
$$

$$
\mathbf{G}\, ds = \begin{pmatrix} 0 & \rho & 0 & 0 & 0 & 0 \\ -\frac{1}{\rho} & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \rho & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \frac{ds}{\rho} = \mathbf{K} \frac{ds}{\rho} = \mathbf{K}\, d\theta.
$$

$$\mathbf{M}(\theta) = \lim_{n\to\infty} \left(\mathbf{I} + \mathbf{K}\frac{\theta}{n}\right)^n = e^{\mathbf{K}\theta}$$

$$= \mathbf{I} + \mathbf{K}\theta + \frac{(\mathbf{K}\theta)^2}{2!} + \frac{(\mathbf{K}\theta)^3}{3!} + \ldots .$$

$$\mathbf{K}^2 = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & \rho \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{K}^3 = \begin{pmatrix} 0 & -\rho & 0 & 0 & 0 & 0 \\ \frac{1}{\rho} & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -\rho \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{K}^4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -\rho \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = -\mathbf{K}^2.$$

Only a finite number of powers of $\mathbf{K}$ must be evaluated, since every $2n \times 2n$ square matrix satisfies its own characteristic function by the Cayley-Hamilton theorem:

The eigenvalues λ of $\mathbf{K}$ satisfy the characteristic equation

$$0 = \sum_{j=0}^{2n} A_j \lambda^j \quad \text{with} \quad A_{2n} = 1,$$

so the theorem gives us

$$\mathbf{K}^{2n} = -\sum_{j=0}^{2n-1} A_j \mathbf{K}^j.$$

$$\mathbf{M} = \mathbf{I} + \mathbf{K}\theta + \mathbf{K}^2\left(\frac{\theta^2}{2!} - \frac{\theta^4}{4!} + \dots\right) + \mathbf{K}^3\left(\frac{\theta^3}{3!} - \frac{\theta^5}{5!} + \dots\right)$$
$$= \mathbf{I} + \mathbf{K}\theta + \mathbf{K}^2(1-\cos\theta) + \mathbf{K}^3(\theta - \sin\theta)$$
$$= \begin{pmatrix} \cos\theta & \rho\sin\theta & 0 & 0 & 0 & \rho(1-\cos\theta) \\ -\frac{1}{\rho}\sin\theta & \cos\theta & 0 & 0 & 0 & \sin\theta \\ 0 & 0 & 1 & \rho\theta & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\sin\theta & -\rho(1-\cos\theta) & 0 & 0 & 1 & -\rho(\theta-\sin\theta) \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

If we use a Hamiltonian using long. canonical coords. (t, w_t) rather than (z, δ)

$$\mathcal{H} = -\frac{qA_s}{p_0} - \sqrt{w_t^2 - \left(\frac{mc}{p_0}\right)^2 - w_x^2 - w_y^2}, \qquad \text{with} \qquad w_t = \frac{U}{p_0 c},$$

rather than (z, δ) then the M_{56} obtains an extra term $+l/\gamma^2$.