Transport Theory

Introduction

Phase space picture

Beam ellipse

Fermi-Eyges theory

Equivalent sources

Angular confusion

Summary

(See also **arXiv1204.4470v2** [physics.med-ph] 29 Apr 2012.)

Proton Transport in a Water Tank

Transverse scale is exaggerated!
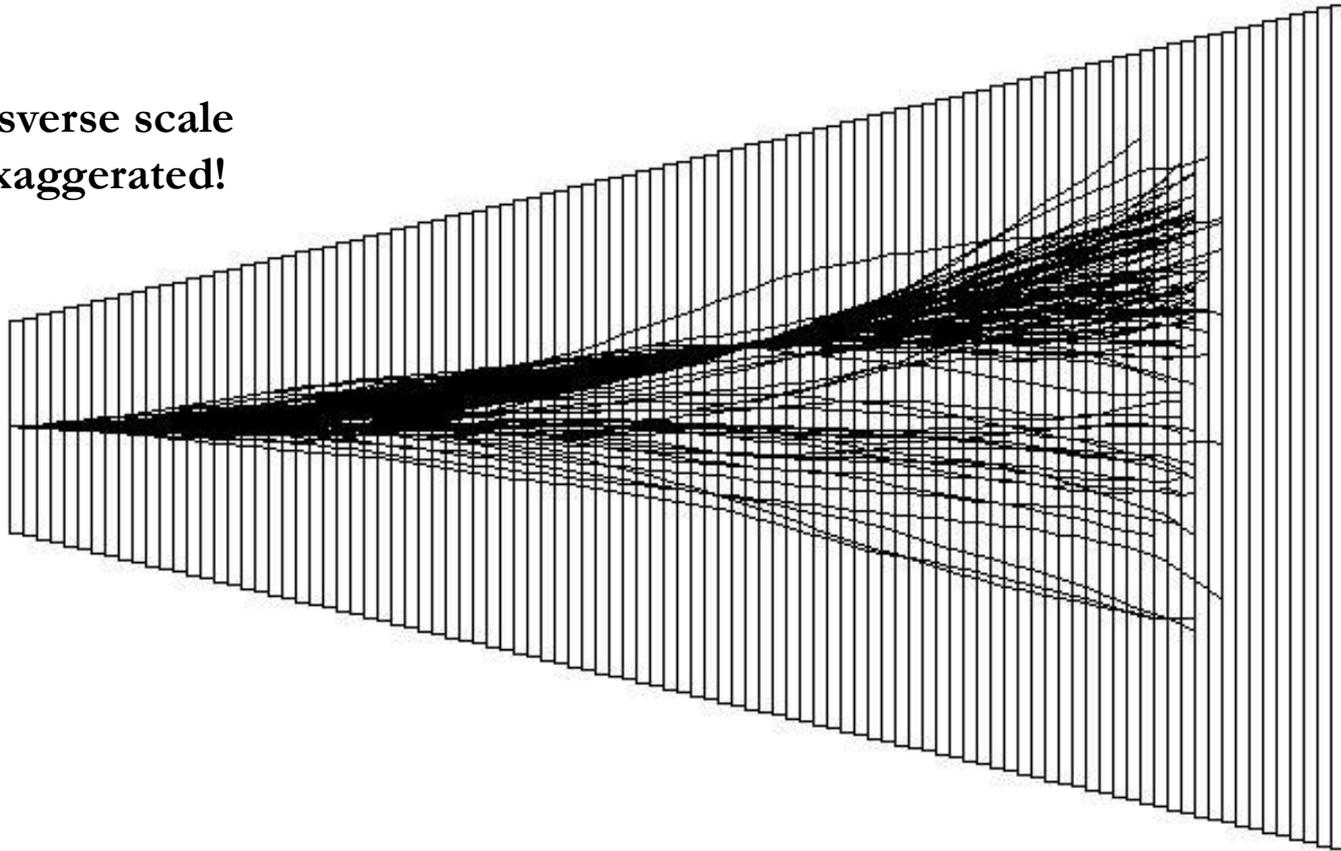
In this Monte Carlo simulation we let a perfect beam enter a water tank. The projection on the yz plane (side view) is shown here, with the transverse scale greatly exaggerated. We ran 20K tracks, plotting the first 30 and after that, only tracks that happened to pass through a narrow off-axis slit. The angular spread of those protons is less than the gross angular spread, and they appear to come from a virtual source somewhat downstream of the entrance face. Transport theory predicts these and related quantities.

What is Transport Theory ?

To generate the previous slide we simulated 20K proton histories. We divided the water tank into fairly thin slabs. We started each proton traveling along the axis. Given the proton's position, direction and energy ('phase space variables') entering a slab, we choose at random new values of position, direction and energy leaving the slab from probability distributions consistent with theories of stopping and multiple scattering. We repeat that until the proton stops. Then we do the next proton.

That method is called a 'condensed history' Monte Carlo. (It imitates nature slab by slab rather than at the atomic level, which would take far too long.) Monte Carlo programs can be used in very complicated geometries where analytic methods need inadmissible approximations. However, it takes a very long time to get accurate results.

Transport theory does not work proton by proton. Instead, it characterizes the entire cohort of protons by a few (typically 7) parameters and computes the evolution of those parameters through the slabs. This only works in simple situations, but when it does it is faster by many orders of magnitude.

Finding the transverse penumbra of a proposed beam spreading system is a typical problem that can be solved either with MC simulation or analytically. The solutions agree fairly well with each other, and with experiment. One advantage of the analytic approach is that it may yield insights that are useful in design work. That is true of the penumbra problem.

Model Beam for Phase Space Discussion

Because scattering and motion (drift) go on simultaneously in a water tank, let's look at a beam line which is simpler for pedagogical purposes: 3 scatterers separated by voids or drifts, then a collimator and a final void. In this case we have chosen to put a pinhole or slit on the beam axis.
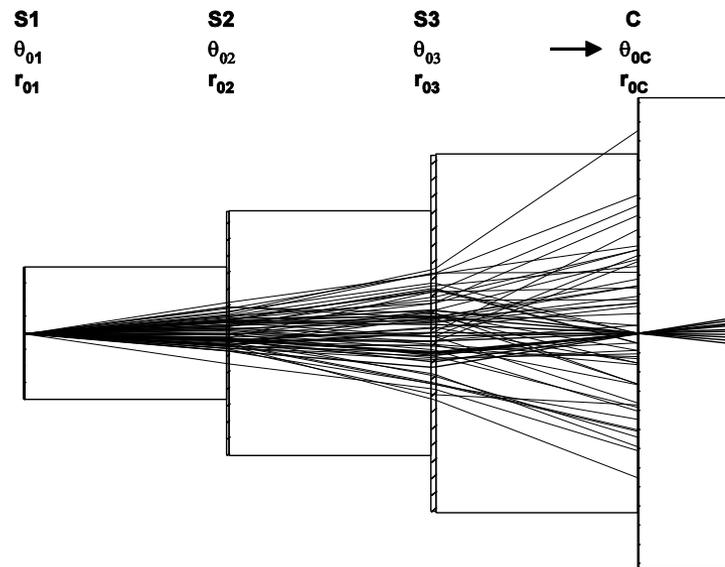
Next we'll define 'phase space variables', a 2D 'phase space diagram', and see how the diagram changes in a scatterer or a void. (In practice, air acts pretty much like a void. Only rarely do we have to account for scattering or energy loss in air.)

The Phase Space Picture

Our 'side view' slide gives a good feeling for what goes on in a thick slab but does not provide a good basis for quantitative calculations. That is provided by the phase space picture of beam transport. Phase space diagrams and ellipses are commonplace in accelerator design but much less so when discussing scattered beam lines, so we'll explain the idea in some detail.

The phase space variables of a proton at a particular level (z) in the beam are

$$\mathbf{x} \quad \mathbf{x'} \quad \mathbf{y} \quad \mathbf{y'} \quad \mathbf{T}$$

the positions, slopes and kinetic energy. (In small angle approximation the slopes equal the angles.) To transport the beam through a given beam element is to find the outgoing phase space variables given the incoming ones.

Often it is convenient to use a quantity other than $\mathbf{T}$, for instance momentum or residual range in water, as the 'longitudinal' variable. $\mathbf{T}$ does not appear explicitly in the phase space diagram but we need to remember that (whatever variable we use) it decreases as $\mathbf{z}$ increases until the proton stops. The scattering power of a given slab is greater if $\mathbf{T}$ or its proxy is smaller.

A 2D phase space diagram is a scatter plot of $\mathbf{x'}$ vs. $\mathbf{x}$. As long as only scatterers and drifts are involved, $\mathbf{y'}$ vs. $\mathbf{y}$ behaves the same way and need not be drawn.

The Effect of a Scatterer
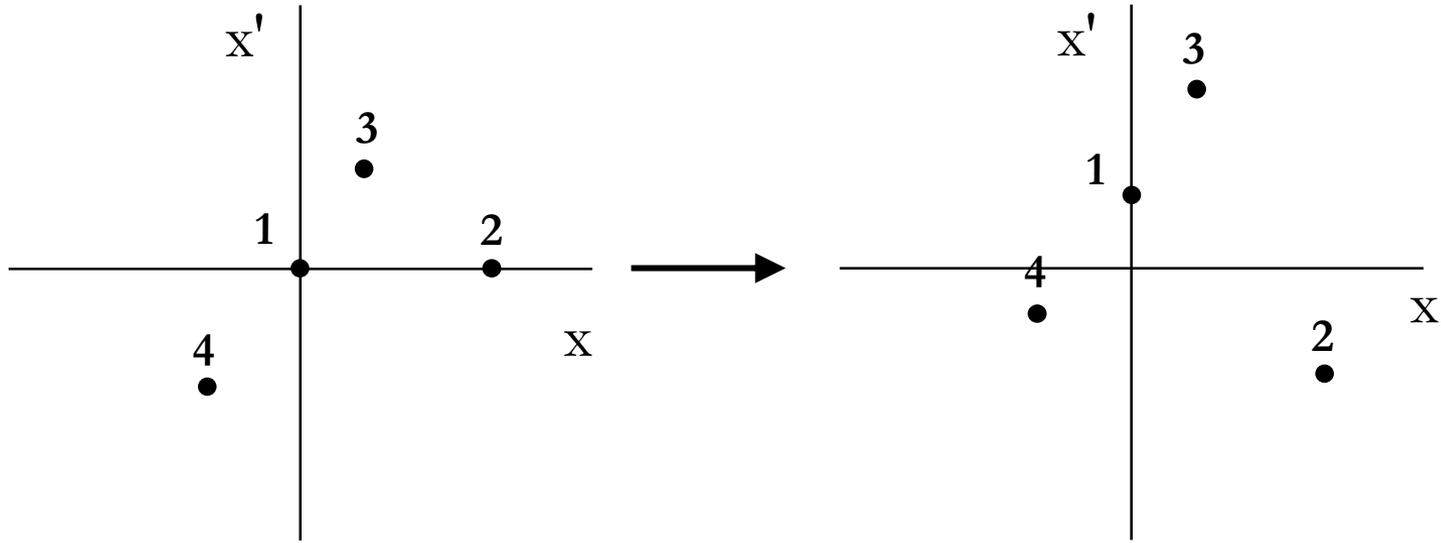
This phase space diagram shows only four protons. As protons move through the scatterer (left diagram to right) each protons receives a random vertical kick. The size of the kick obeys a Gaussian distribution whose characteristic width (θ_0) depends on the scatterer strength as given by Highland's formula or (slightly better) Hanson's formula.

N.B. the phase space diagram is emphatically **not** a picture or cross section of the beam as we might get on film ! That would be a scatter plot of **y** vs. **x**.

The Effect of a Drift
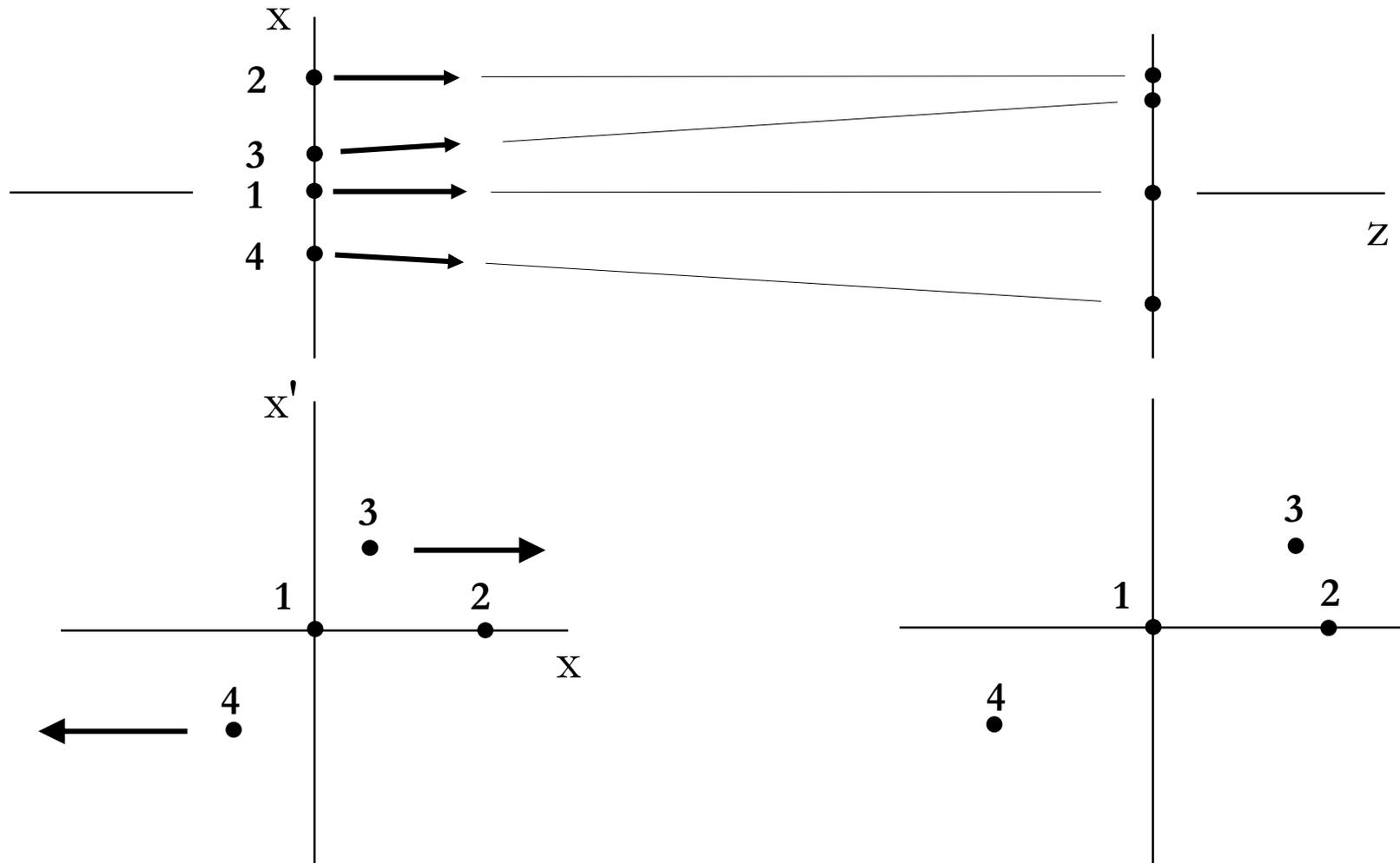
By contrast, in a pure drift the phase space distribution shears. Protons above the x axis move right, protons below move left, and protons on the axis stay where they are. Think of laminar flows in phase space.

The Beam Ellipse
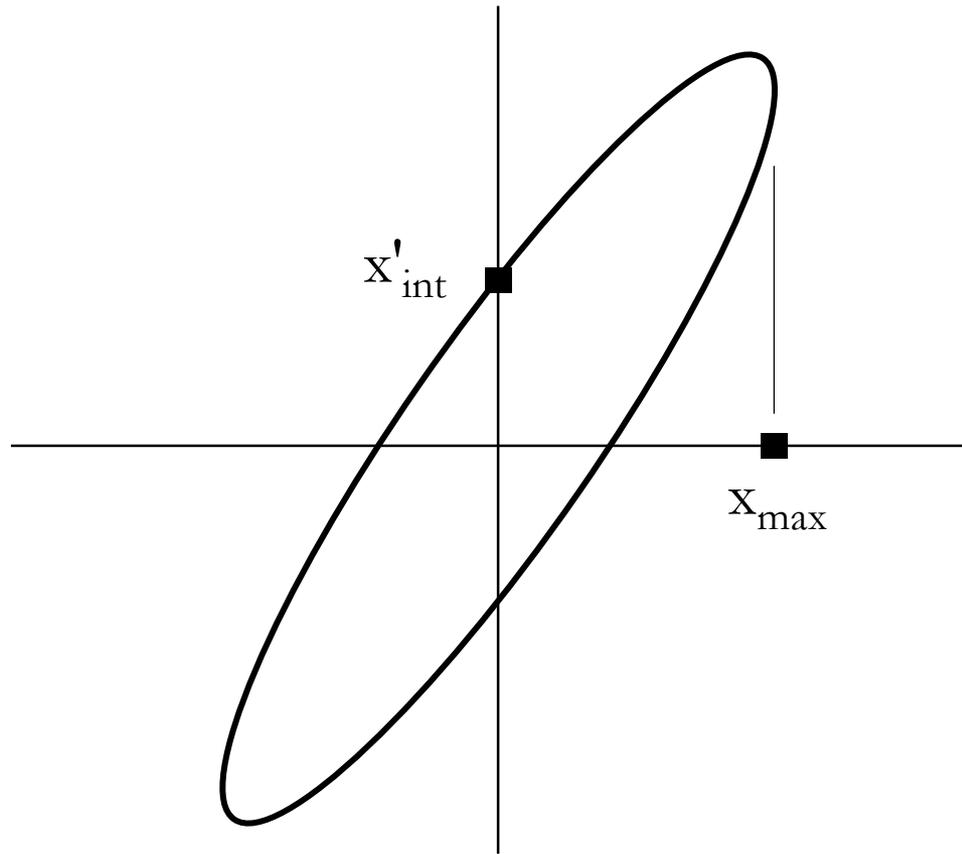
As we'll see shortly, the phase space distribution representing the beam can be bounded, at any given beam element, by an ellipse. The three ellipse parameters (for instance, semi-axes and tilt angle) can be computed at any z by means of transport theory. The area enclosed by the ellipse is $\pi\, x'_{int}\, x_{max}\ (= \pi\, x_{int}\, x'_{max})$ and that (or sometimes area/π) is called the the beam emittance at that **z**.

Liouville's Theorem
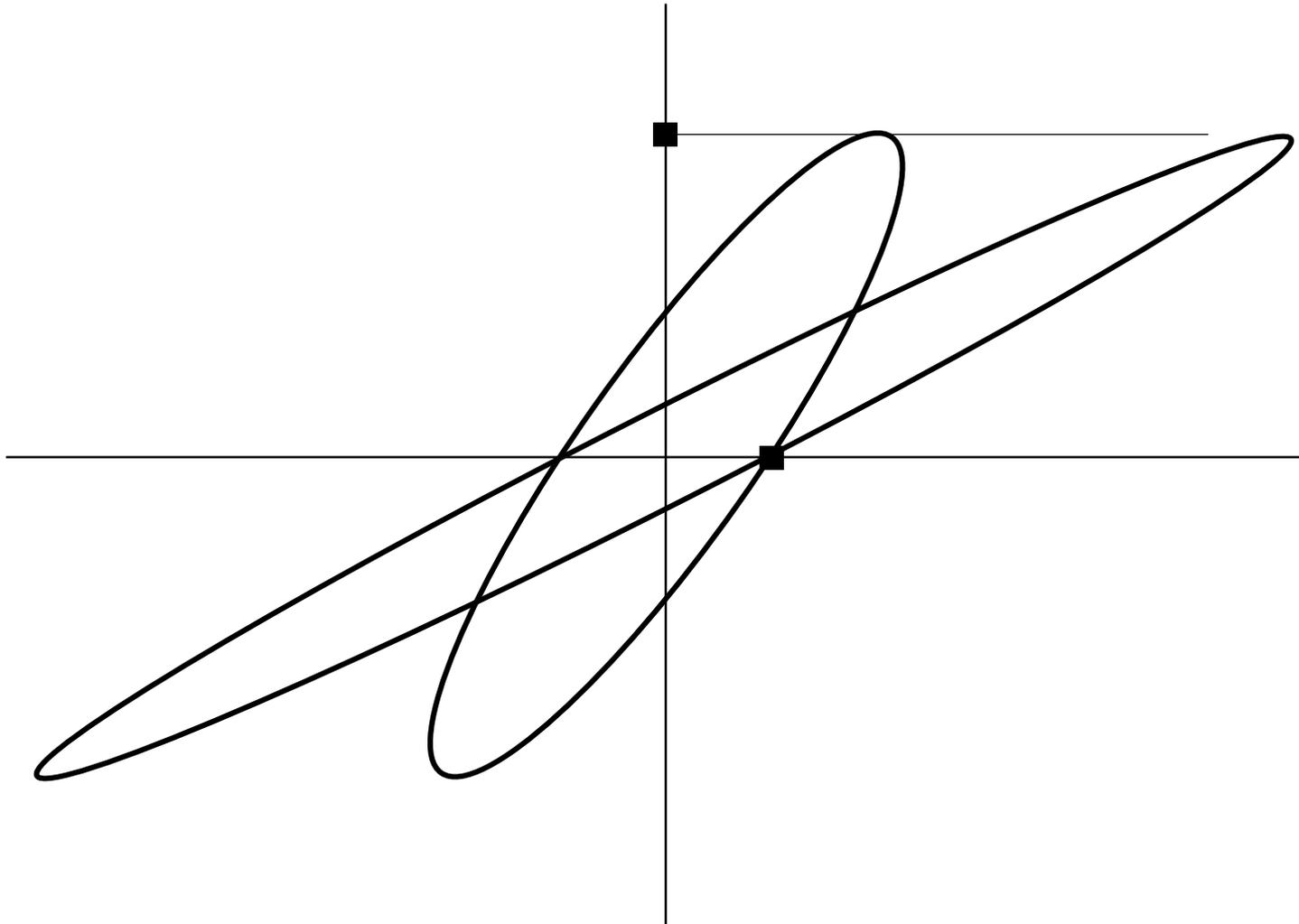
Phase space area, or beam emittance, is conserved in a drift. That follows from the previous formula for the area, plus the fact that y_{max} and x_{int} do not change in a drift. This is a specific instance of ***Liouville's Theorem***.

Phase Space Summary

Phase space coordinates at z : x x' y y' T

Phase space diagram: scatter plot, at some z, of x' vs. x

Thin scatterer : random $\pm$ vertical kicks, Gaussian, rms θ_0.

Pure drift : phase space distribution shears.

(Thick scatterer : some of each.)

Phase space ellipse: encloses a collection of points in the phase space diagram. In accelerator physics the bounding ellipse is an approximation. In scattering systems it is exact, to the extent that the Gaussian approximation applies.

Emittance $\equiv$ ellipse area $= \pi\, x'_{int}\, y_{max} = \pi\, y'_{int}\, x_{max}$. Conserved in a pure drift, and obviously not conserved in a scatterer.

The evolution of the beam ellipse (its 3 parameters) with z can be computed using transport theory: Fermi-Eyges or an equivalent.

Phase Space Evolution of the Model Beam

At last we're able to trace the evolution of phase space in our beam consisting of three scatterers S separated by voids V followed by a collimator C :
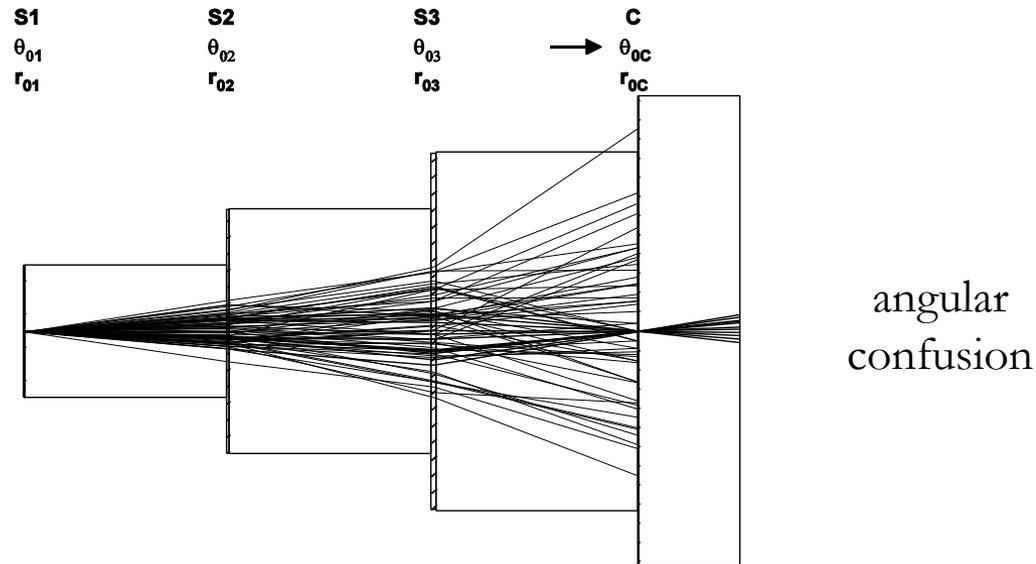
The next slide shows the phase space cloud at the **entrance** to each element. We've given the incident beam a little size and angular spread to make it more visible. S1 spreads it vertically. V1 shears it without changing the area. But S2 changes the area significantly. Continuing in that vein, a bounding ellipse is shown at the entrance to C. The pinhole selects a strip whose half-height is proportional to the angular spread on axis. These diagrams demonstrate that the increase in emittance at a scatterer is proportional to the beam size entering that scatterer, a key consideration in minimizing lateral penumbra.
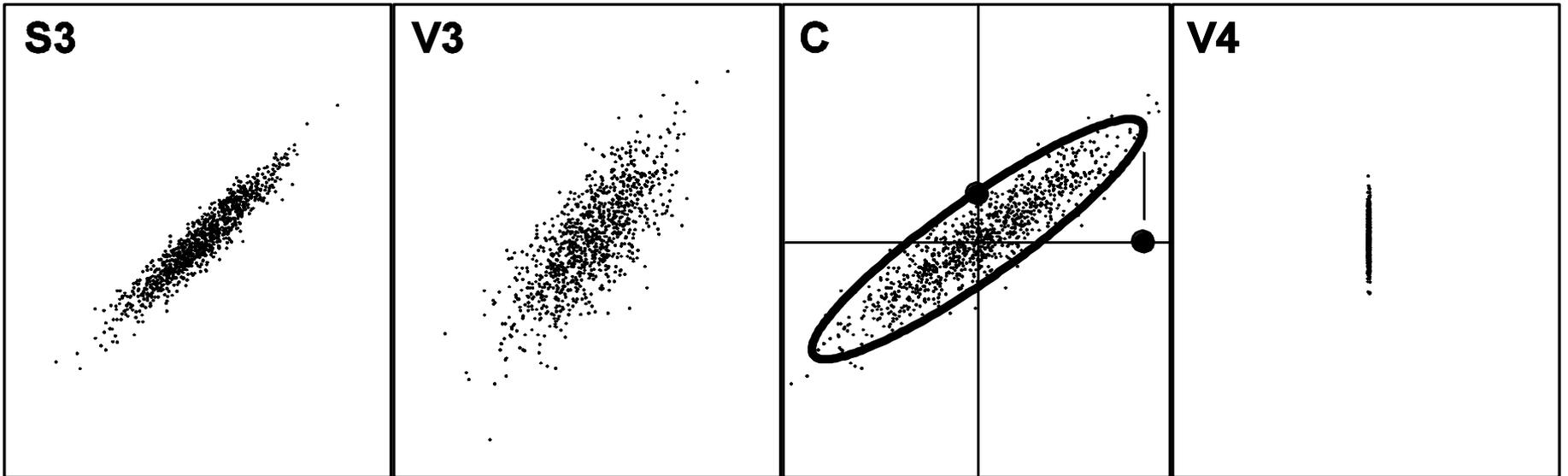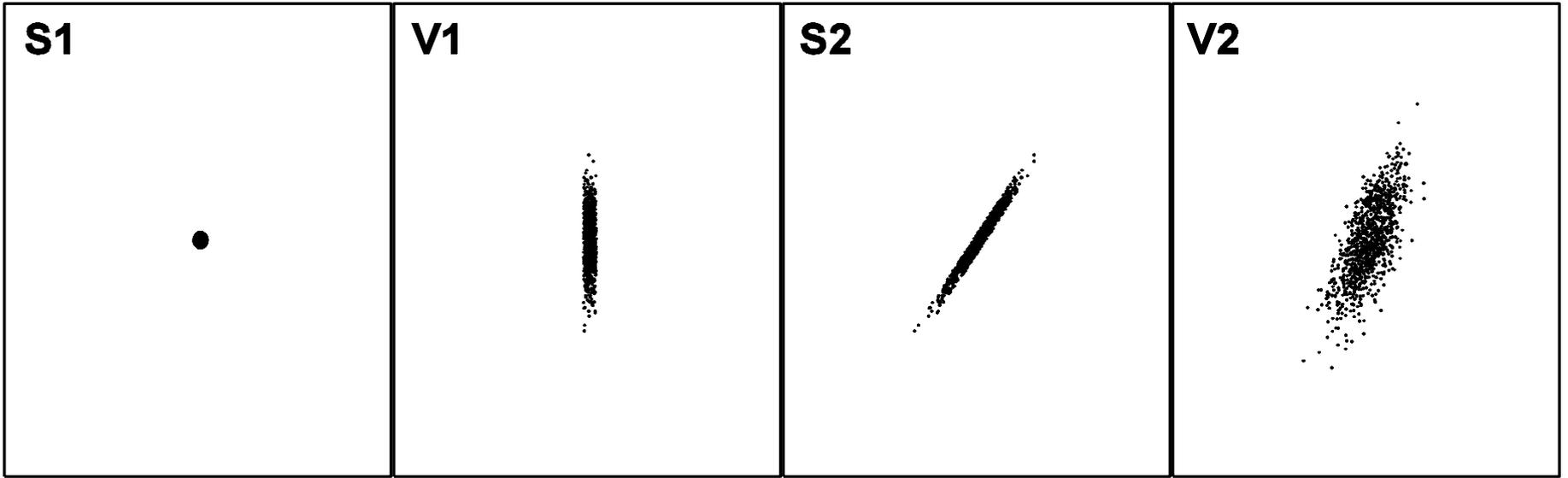
# A More Interesting Beam Line

The next two slides anticipate the lecture on lateral penumbra. The beam line (uniform scatterer S1, contoured scatterer S2, collimator, water tank) models the NPTC gantry.
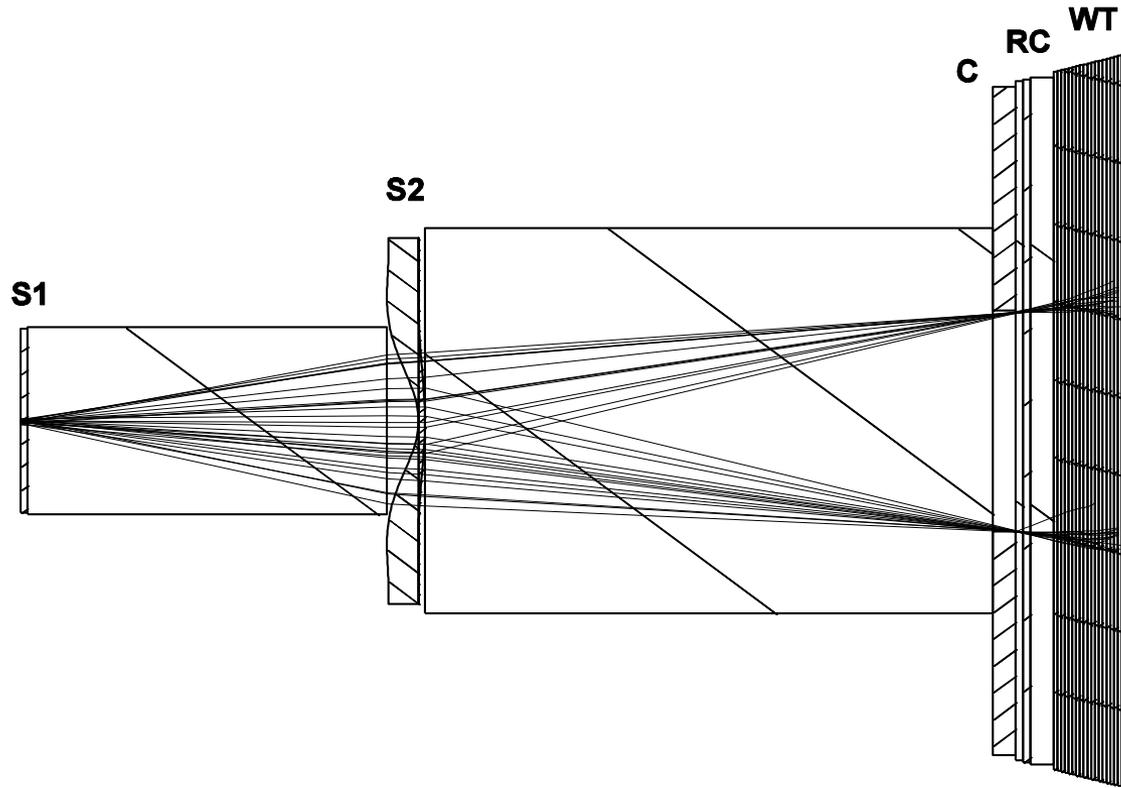
From 1M events we selected ≈ 30 that passed through 2 mm pinholes at the upper and lower collimator edges. The angular confusion is obviously related to the size of the beam at S2. Also, the ray bundles appear to point back to a well defined virtual source with a finite size. Computation shows it is about halfway between S1 and S2
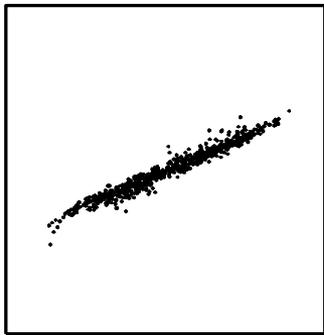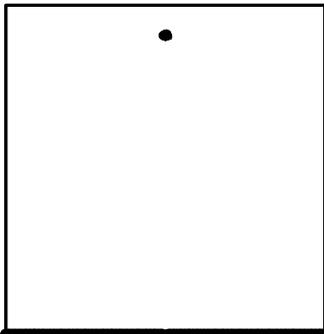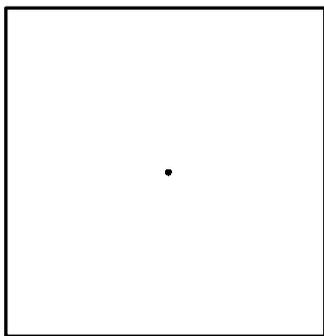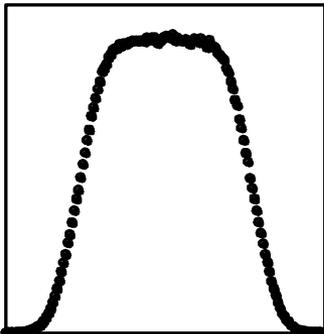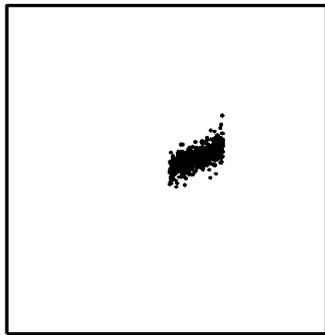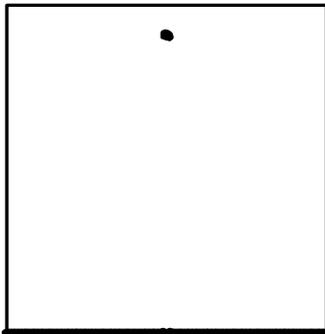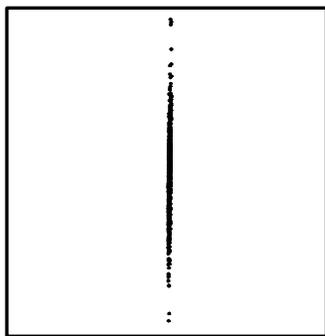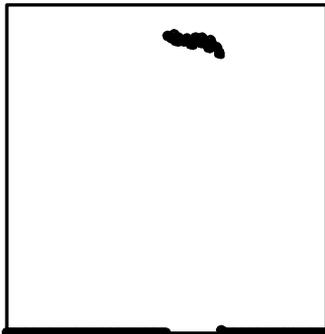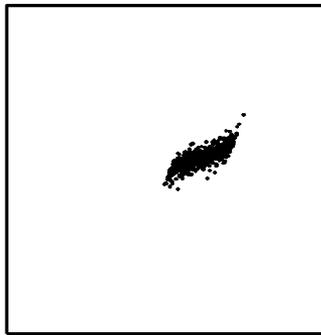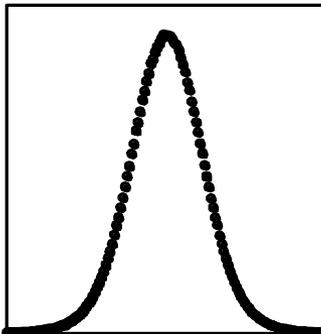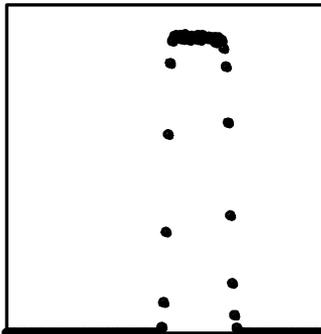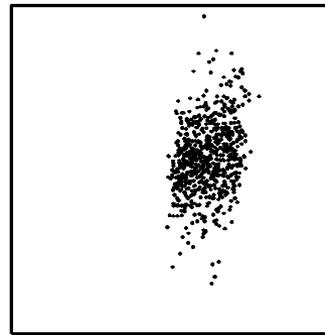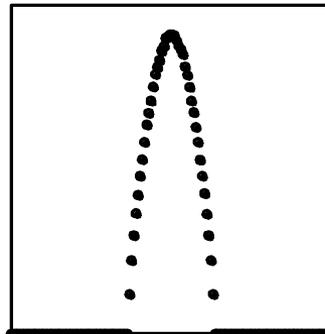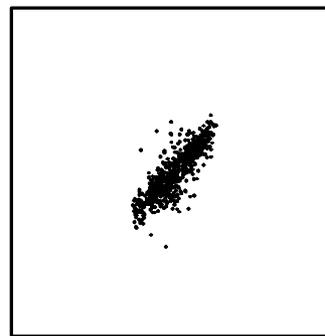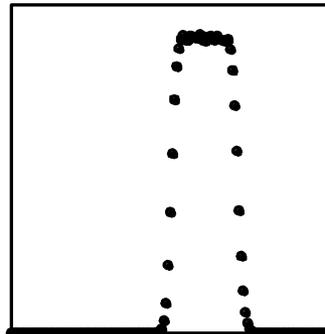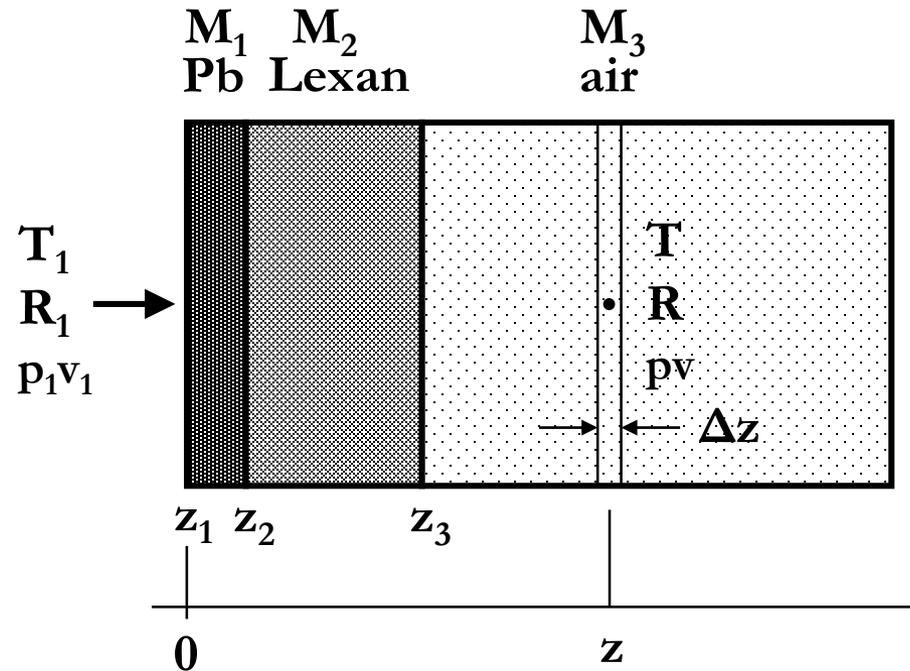
Fermi-Eyges Theory

Fermi-Eyges theory propagates the parameters of the beam ellipse through matter. Originally (L. Eyges, Phys. Rev. 74 (1948) 1534-1535) it assumed an ideal beam into a single homogeneous slab, and it used a particular formula (our T_{FR}) for the scattering power. Later it was generalized to any Gaussian incoming beam, a 'stack' of uniform slabs, and other scattering powers. However, transverse heterogeneity is not allowed or equivalently, the unknown beam at z must still be Gaussian. That means the theory does not apply to most beam lines, and even less to dose reconstruction in the patient. However, it is still worth learning because it is used as a building block in more complicated computations.

Scattering Power

T is defined as the rate of increase of mean squared projected MCS angle:

$$T \equiv \frac{d < \theta_x^2 >}{dz}$$

It is the topic of the next lecture. There are presently 6 different formulas, including T_{dM} suggested recently by BG (Med. Phys. (2010), in press). Any accurate formula consists of three factors: material dependent, energy dependent, and a *single scattering correction*.

$$T_{dM} = f_{dM}(pv, p_1 v_1) \times \left(\frac{E_s}{pv}\right)^2 \frac{1}{X_S}$$

$$\frac{1}{\rho X_S} \equiv \alpha N r_e^2 \frac{Z^2}{A} \left\{ 2 \log(33219 \, (AZ)^{-1/3}) - 1 \right\}$$

Compounds or mixtures: $$\frac{1}{\rho X_S} = \sum_i w_i \left(\frac{1}{\rho X_S}\right)_i$$

$$f_{dM} \equiv 0.5244 + 0.1975 \lg(1 - (pv/p_1 v_1)^2) + 0.2320 \lg(pv) - 0.0098 \lg(pv) \lg(1 - (pv/p_1 v_1)^2)$$

Scattering length X_S is a property of the current material. $\lg \equiv \log_{10}$ .Other symbols are defined in the next lecture. The main thing is that if you know where you are in the stack (z) and what the incident energy was, you can compute the number T.

Moments of T

$$A_0(z) \equiv \int_0^z T(u)\,du$$

$$A_1(z) \equiv \int_0^z (z - u)\,T(u)\,du$$

$$A_2(z) \equiv \int_0^z (z - u)^2\,T(u)\,du$$

$$B(z) \equiv A_0 A_2 - A_1^2$$

The moments of the scattering power A_n, as well as a special combination B, play a central role in Fermi-Eyges theory. u is a z-like variable representing distance along the slab or stack of slabs. In T(u) it is a sort of bookkeeping tool. T depends upon a number of things such as kinetic energy and scattering material, and these in turn depend on where we are in the stack. We will need some subroutine which knows the construction of the stack and the incoming energy, and can therefore return the needed quantities when called with a specific value of u: WhatsHere(u, ...) .

Probability Density in Phase Space

The main result of Fermi-Eyges theory (L. Eyges, loc. cit.) is the probability, at any $z > 0$, of finding a proton at x in dx and θ in $d\theta$ if it enters a single slab along the z axis at $z = 0$:

$$P(x, \theta)\, dx\, d\theta \;=\; \frac{1}{2\pi\sqrt{B}}\, e^{-\frac{1}{2}\frac{A_0 x^2 - 2A_1 x\theta + A_2 \theta^2}{B}}\, dx\, d\theta$$

P itself is the 2D probability *density* in phase space. Everything flows from this equation. If we multiply by N, the number of protons in the beam, we obtain the number density in phase space. By generalizing the A's we can apply the equation to a non-ideal (but known) incident beam and mixed slabs.

If we set the second fraction in the exponent equal to one we obtain the iso-probability contour at 61% of the central intensity:

$$F(x, \theta) \;\equiv\; A_0 x^2 - 2A_1 x\theta + A_2 \theta^2 \;=\; B$$

which, according to the analytic geometry of conic sections, is an ellipse since it can be shown that B > 0. This gives quantitative meaning to the beam ellipse.

Physical Interpretation of the A's

One can relate the A's to mean values of the phase space variables at z. We give the formulas here without proof. The square roots of A_0 and A_2 respectively define the bounding box of the ellipse in θ and x , or the rms values of θ and x. A_1 is more subtle. It is the *covariance* of x and θ, related to the tilt of the ellipse. If $A_1 > 0$ (θ generally increases with x) the beam is divergent.

$$A_0 = \;<\theta^2>$$

$$A_1 = \;<x\theta>$$

$$A_2 = \;<x^2>$$

$$\pi\sqrt{B} = \;\text{area enclosed by ellipse}$$

It can be shown that $B \equiv A_0 A_1 - A_2{}^2 \geq 0$. Therefore its square root is real, and one can show that $\pi B^{1/2}$ is the area enclosed by the beam ellipse: the beam's *emittance* at z. A_0, A_2 and B form an alternative set of ellipse parameters. A_1 carries a bit more information than B, namely which way the ellipse is tilted, but in scatter/drift systems the ellipse always tends towards quadrants 1 and 3 : the beam ultimately diverges. Only with magnets can we make converging beams.

Beam Ellipse in A_n Notation

θ (rad)

$A_1/A_0^{1/2}$

$A_0^{1/2}$

Q

R $\quad A_1/A_2^{1/2}$

$(B/A_2)^{1/2} \rightarrow$ P

x (cm)

S

O $\quad A_2^{1/2}$

$(B/A_0)^{1/2}$

area enclosed $= \pi\, B^{1/2} = \pi\, (A_0\, A_2 - A_1^2)^{1/2}$

OR = most probable θ given x, $\theta = (A_1/A_2)\, x$

OQ = most probable x given θ, $x = (A_1/A_0)\, \theta$

Beam Ellipse in x, θ Notation
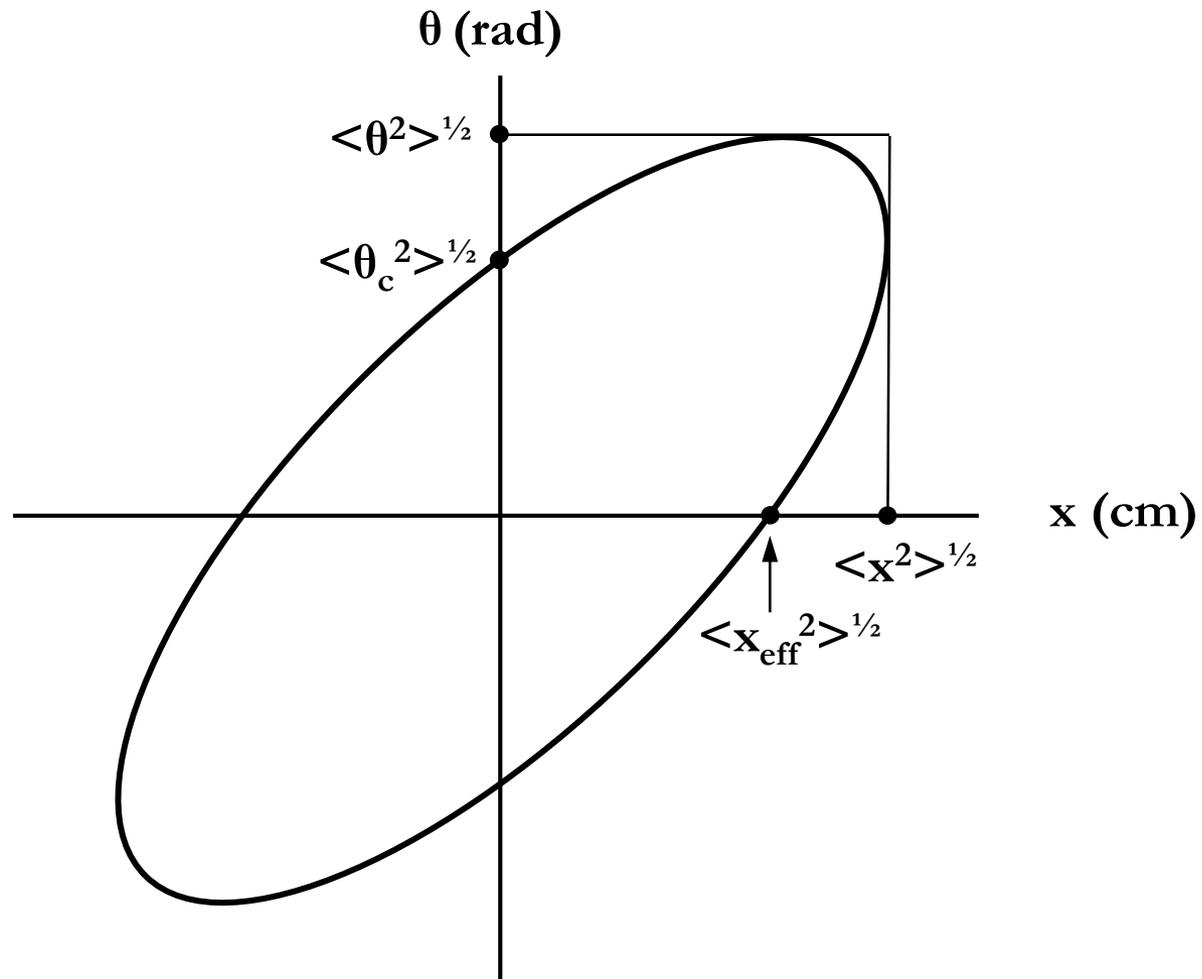
area enclosed = $\pi \, (\langle\theta^2\rangle\langle x^2\rangle - \langle x\theta\rangle^2)^{1/2}$

Propagation of A's in Mixed Slabs

The computation of the scattering power moments A can be generalized to the case where the entering values are not 0 (imperfect beam). Assume the slab extends from 0 to z , the unprimed A's are initial values and the primed A's are final values. Then

$$A_0' = A_0 + \int_0^z T(u)\, du$$

$$A_1' = A_1 + A_0\, z + \int_0^z (z - u)\, T(u)\, du$$

$$A_2' = A_2 + 2\, A_1\, z + A_0\, z^2 + \int_0^z (z - u)^2\, T(u)\, du$$

This process can be repeated as often as necessary, so we have a formalism for propagating the beam ellipse through a mixed finite slab geometry.

If the initial A's are 0, these equations are identical to the original theory by inspection. If T is zero (no scattering in $0 - z$, pure drift) it is easy to show that the equations lead to $\Delta B = 0$: no emittance change in a drift.

Emittance Increase in a Scatterer

Suppose a beam line consists of thin scatterers separated by drifts. The propagation of A's through a given scatterer can then be approximated by

$$
\begin{aligned}
A_0' &= A_0 + T\,\Delta z \\
A_1' &= A_1 + A_0\,\Delta z + T\,(\Delta z)^2/2 \\
A_2' &= A_2 + 2A_1\,\Delta z + A_0\,(\Delta z)^2 + T\,(\Delta z)^3/3
\end{aligned}
$$

To lowest order in z

$$
B' = A_0'A_2' - A_1'^2 \simeq A_0 A_2 - A_1^2 + T\,\Delta z\, A_2
$$

or

$$
\Delta B = \Delta(\epsilon^2) = A_2\,(T\,\Delta z) = <x^2> \times \Delta <\theta^2>
$$

In other words, Fermi-Eyges theory is consistent with our guess that the emittance increase in a scatterer equals the incident rms beam size times the change in rms scattering angle. At the end of a beam line having N scatterers

$$
\epsilon^2 = \epsilon_{\text{beam}}^2 + \sum_{i=1}^{N} (\overline{x^2})_i\,(\Delta\overline{\theta^2})_i
$$

which leads to a formula for the angular confusion θ_C , which in turn we need to compute transverse penumbra.

Stepwise Propagation of A's

Sometimes (as when reconstructing dose in a patient) we discretize the inhomogeneities and do Fermi-Eyges theory step by step. The following compact form (N. Kanematsu, Nucl. Instr. Meth. B **266** (2008) 5056-5062) is easily derived from the integral version by assuming T is constant within Δz :

$$\Delta A_0 \;=\; \tilde{T}\Delta z$$

$$\Delta A_1 \;=\; (A_0 + \frac{\tilde{T}}{2}\Delta z)\Delta z$$

$$\Delta A_2 \;=\; (2A_1 + (A_0 + \frac{\tilde{T}}{3}\Delta z)\Delta z)\Delta z$$

$\tilde{T}$ denotes T at the midpoint of the step. These equations are equivalent to integration by the midpoint rule and only work well if Δz is small.

Example: Stopping-Length Water Slab
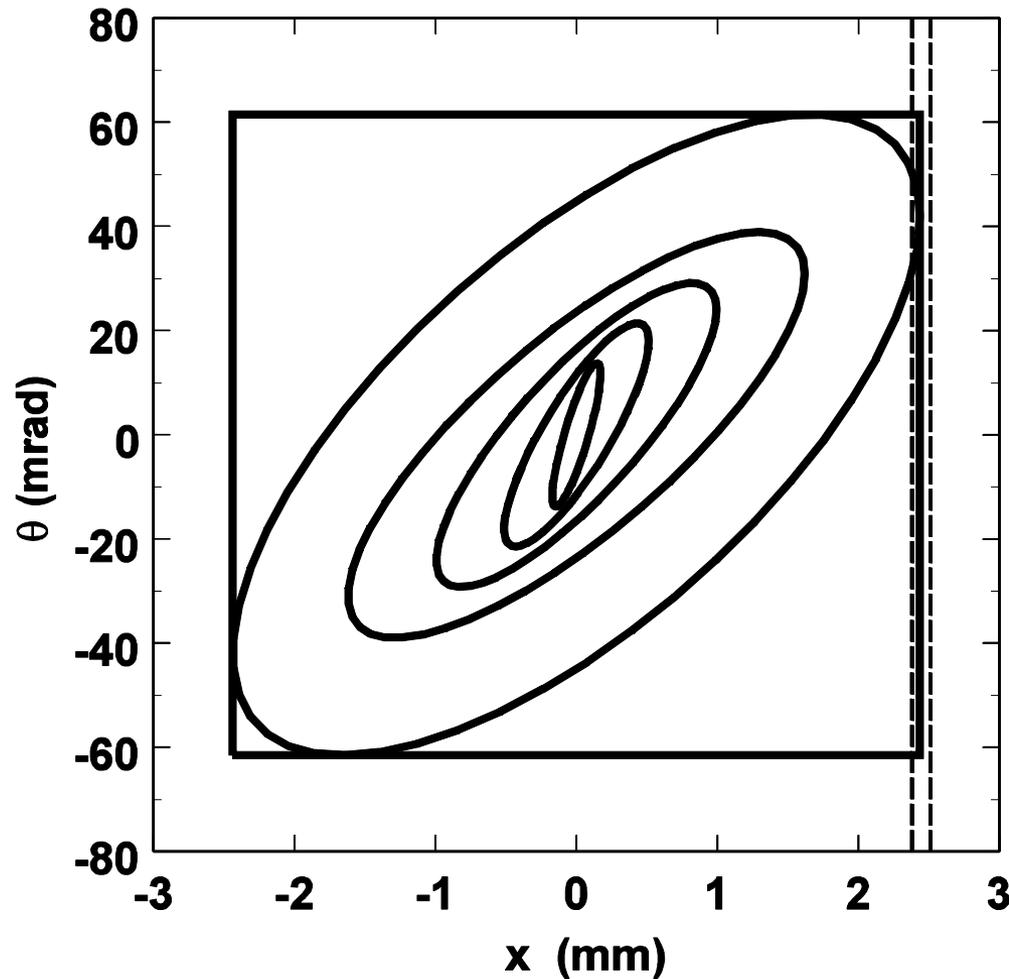
A 127 MeV proton pencil beam enters a near stopping length (z = 0.97 R1) slab of water divided into five equal slabs. The bounding box of the final ellipse is shown. It agrees with a measurement (dashed lines) of the rms projected transverse spread by Preston and Koehler (unpublished MS (1968) HCL).

Example: Pb/Lexan/air stack
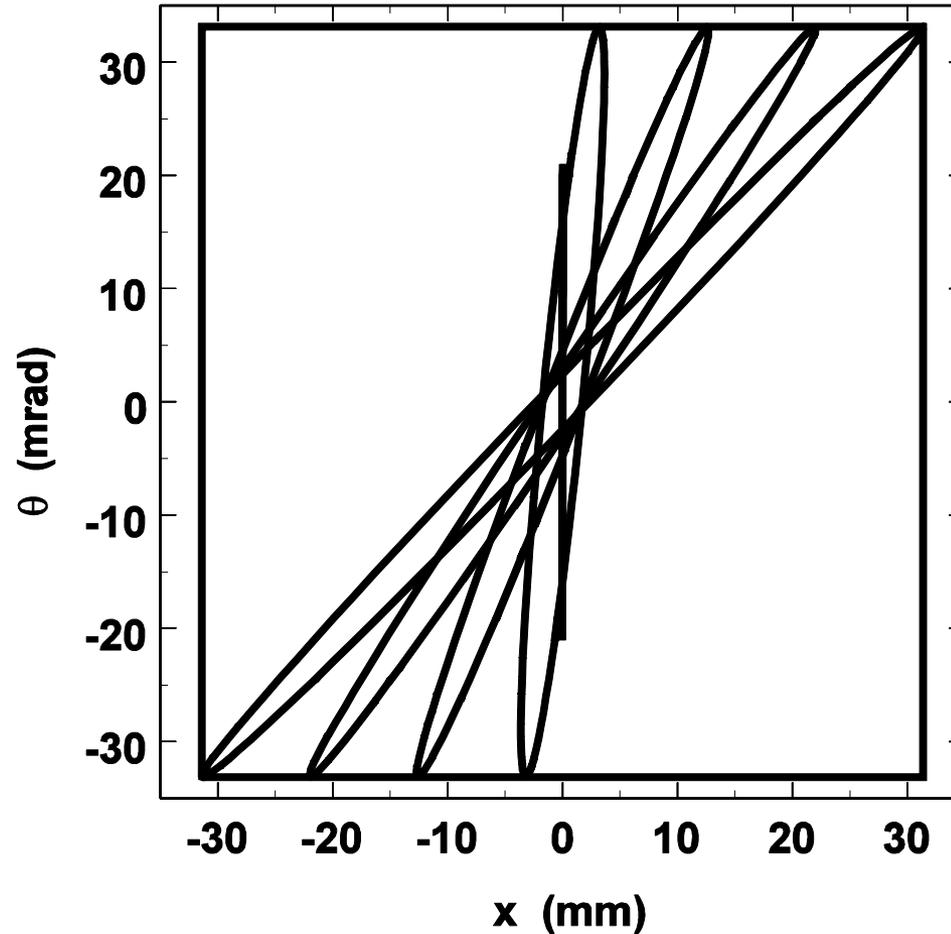
A 230 Mev proton beam enters a Pb/Lexan/air stack typical of one position of the range modulator (and drift space to the second scatterer) of the IBA radiotherapy nozzle. The Pb is an almost pure scatterer. The Lexan shows some scatter and some drift. The air (here divided into 3 equal slabs) is an almost pure drift.

Equivalent Sources

Frequently, particularly in beam line design, we need to replace a subsystem (a stack of mixed slabs) by an 'equivalent source'. For instance, each position of a range modulator looks (to the second scatterer S2) like a given Pb/Lexan/air stack. After taking multiple scattering in the stack into account, where exactly do the protons seem to be coming from? This is particularly important if the modulator takes up most of the space upstream of S2. Three ways of computing an equivalent source and its distance from the arbitrary measuring plane (MP) are found in the literature:

The 'effective extended source' is that erect upstream ellipse which would drift into the observed ellipse at z_{MP} (ICRU Report 38, 1984).

The 'virtual point source' is that erect degenerate upstream ellipse which would yield, at z_{MP}, the same most probable direction as a function of x (ibid.).

The 'effective scattering point' is that point from which, if z_{MP} is far downstream, protons at xrms appear to be coming (Gottschalk et al., Nucl. Instr. Meth. **B74** (1993) 467-490).

Exactly which one to use in a particular case is a bit of a mystery even to us; we have generally used the last one. However it is easy to show that the differences between the three source *distances* are usually small in practice.

Effective Extended Source
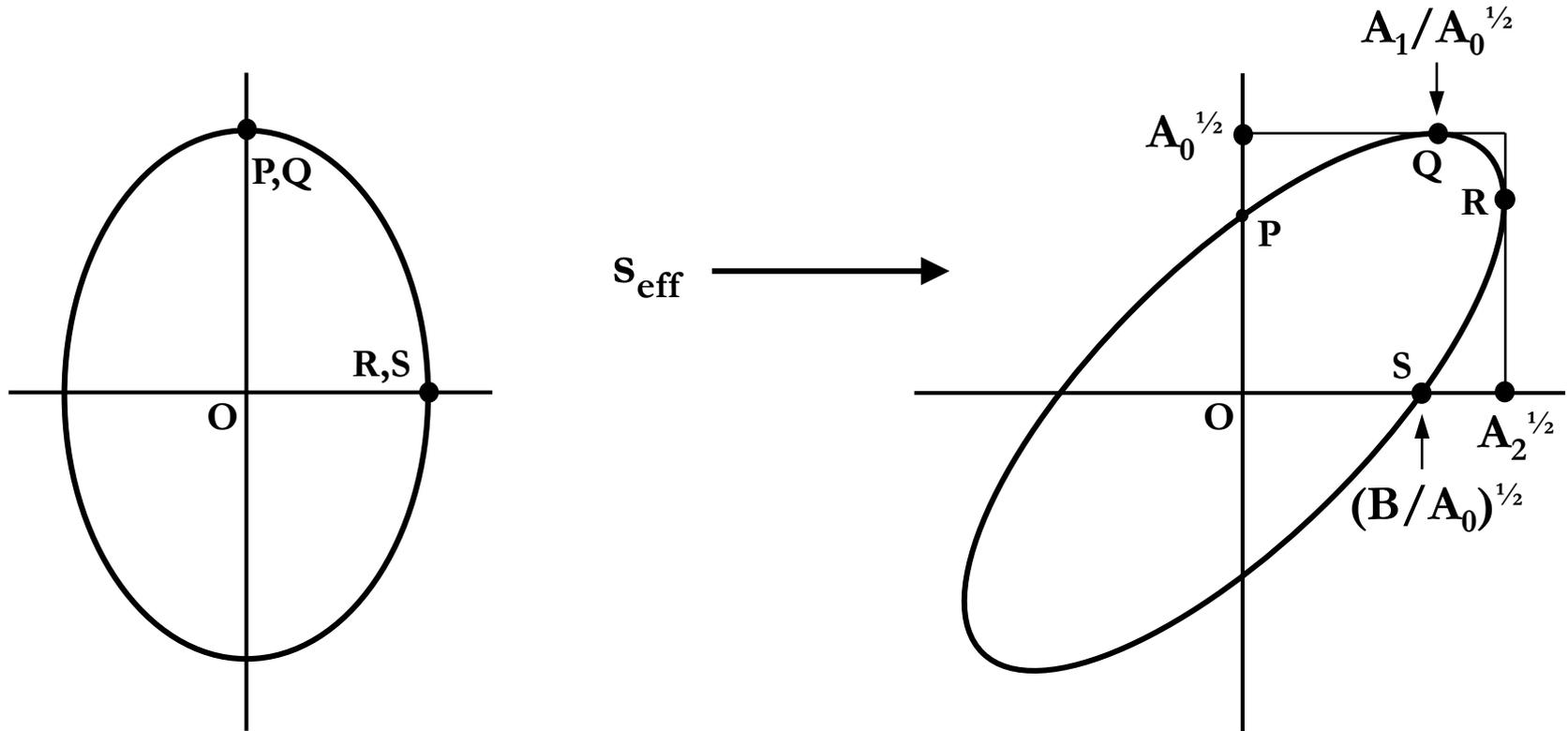
The 'effective extended source' is that ellipse which would drift into the given ellipse at the MP. Since there are infinitely many, we also require that it be erect. The drift distance of a general phase space point (x,θ) from an initial $(0,\theta)$ is $S = x/\theta$. Applying that to point Q we find

$$S_{eff} = A_1/A_0 \quad \text{or} \quad z_{eff} = z_{MP} - A_1/A_0 .$$

Point S does not move in a drift so the source size is $x_{rms,\ eff} = (B/A_0)^{1/2}$.

Virtual Point Source
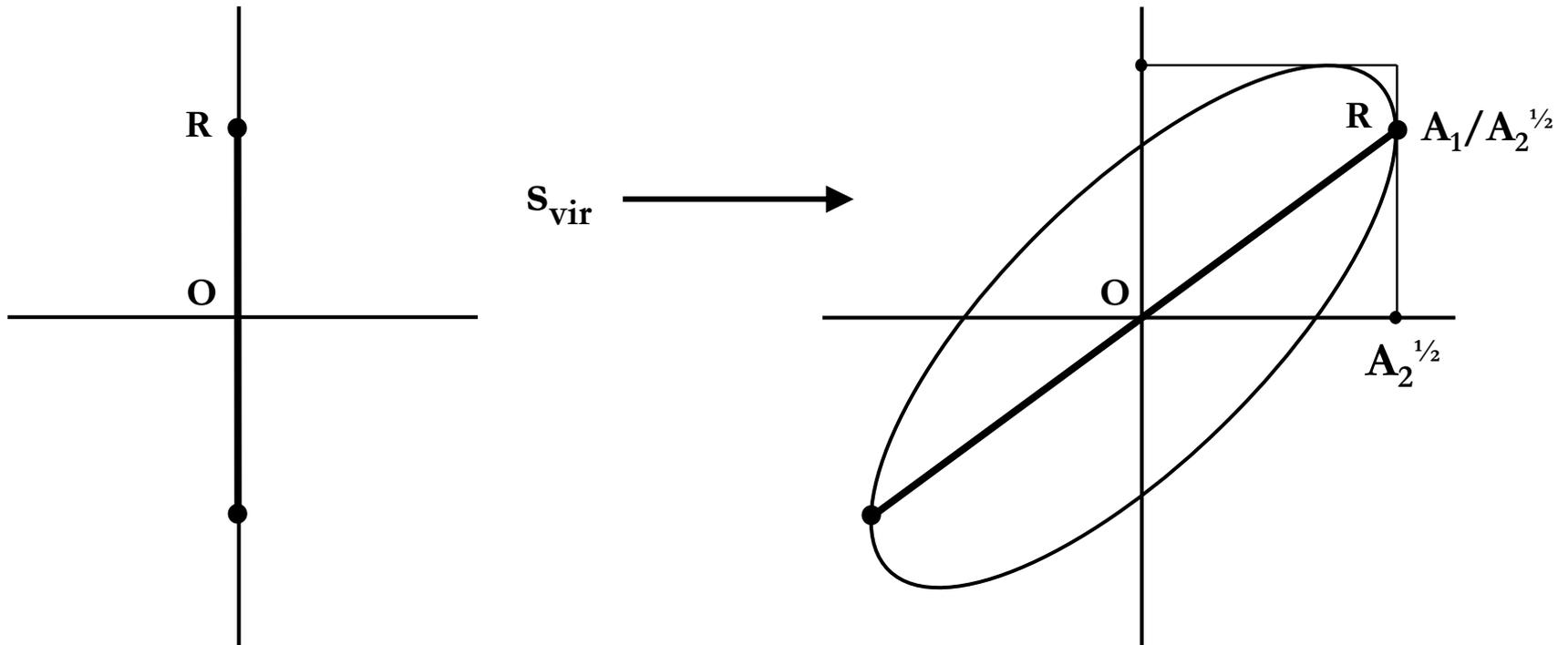
The 'virtual point source' is that ellipse which drifts into **OR,** the locus of the most probable proton direction given x. Reasoning as before

$$S_{vir} = A_2/A_1 \quad \text{or} \quad z_{vir} = z_{MP} - A_2/A_1$$

The VPS is the point from which protons appear to emanate: $x_{rms,vir} = 0$

Effective Scattering Point
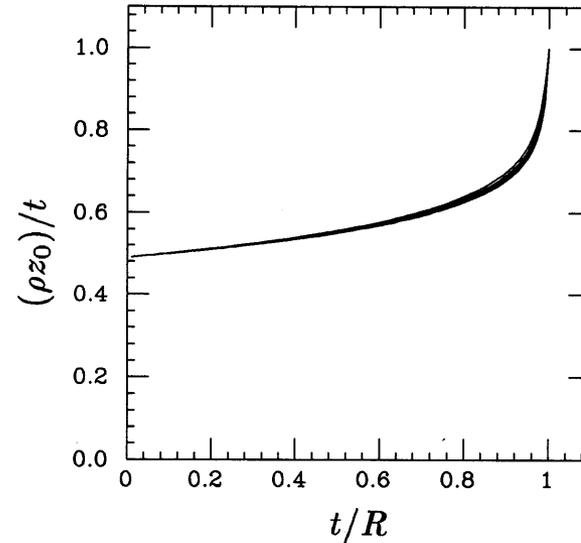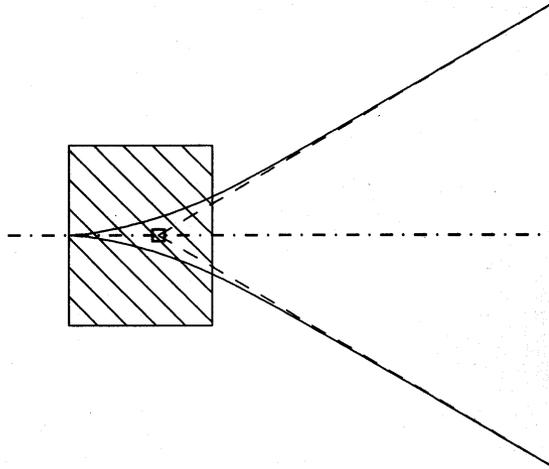
Gottschalk et al., Nucl. Instr. Meth. **B74** (1993) 467-490 defined an equivalent source distance using the back projected asymptote of $x_{rms}(z)$, which Kanematsu (Phys. Med. Biol. **54** (2009) N67-N73) calls the 'effective scattering point'. It can be identified with the upper right hand corner of the bounding box after a long drift, whence

$$\mathbf{S_{esp} = (A_2/A_0)^{1/2}} \quad \text{or} \quad \mathbf{z_{esp} = z_{MP} - (A_2/A_0)^{1/2}}$$

Recalling $\mathbf{S = x/\theta}$ it is obvious from the general ellipse diagram that

$$\mathbf{S_{vir} > S_{esp} > S_{eff}} \quad \text{or} \quad \mathbf{z_{vir} < z_{esp} < z_{eff}}$$

In most practical cases the three are very close together and any can be used.

Effective Scattering Point (Mixed Slabs)
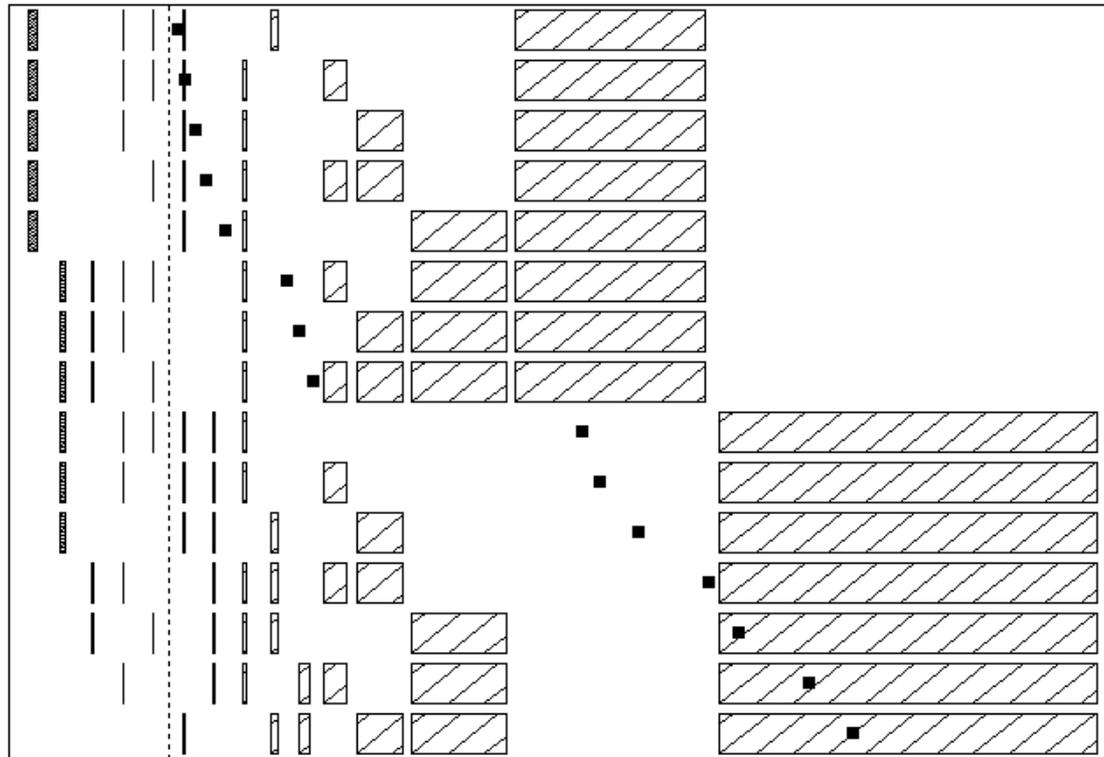
30JUL04   09:58:52.40

The STAR neurosurgery beam at the Burr Center uses 'lamination', a sequence of lead or plastic scatterers chosen from binary sets, to create the SOBP. The z of each scatterer is fixed so z_{esp} (full squares) jumps around, by as much as a few cm at the major transition, as we go through the sequence. Left uncorrected, that would cause a $1/r^2$ bump in the SOBP. The observed SOBP is flat, showing this effect is handled adequately by the computer program. The long throw (4.5m) helps a lot.

Effective Scattering Point (Non-Uniform Slabs)
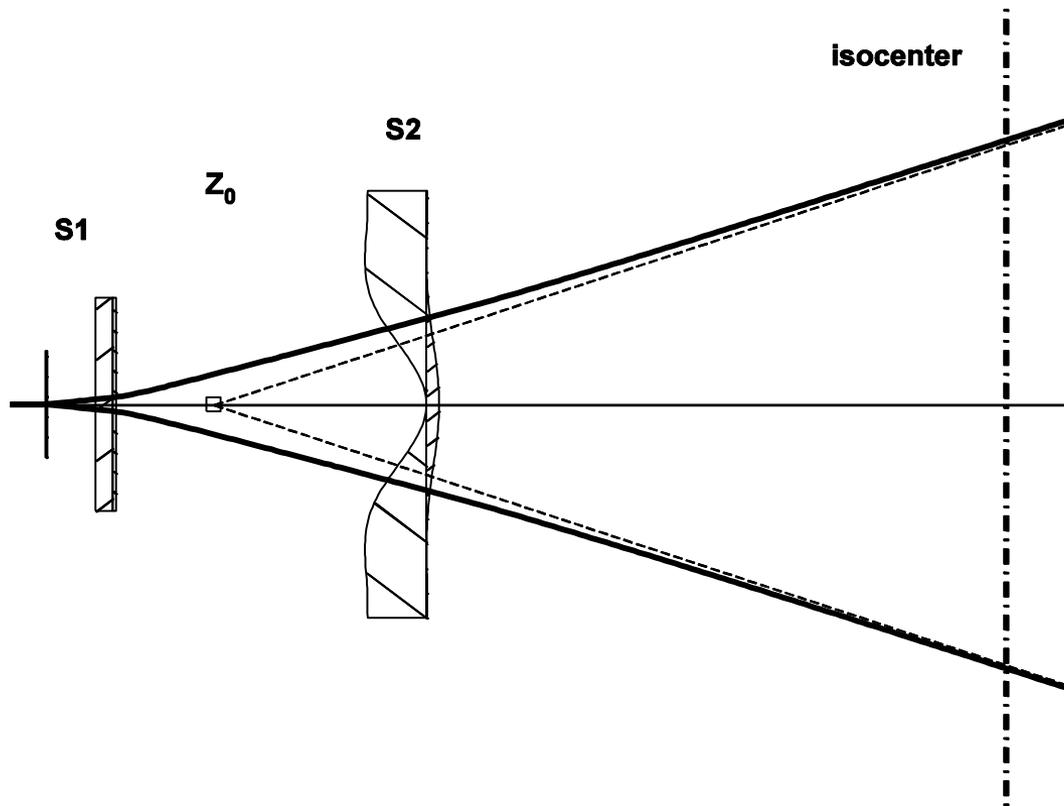
The equivalent source idea also applies to non-uniform slabs but we must choose an equivalent thickness for the non-uniform slab. For a compensated contoured scatterer, its thickness at x_{rms} (from the elements upstream) is a reasonable choice and seems to give the right answer (lecture on transverse penumbra).

The virtual source is computed by functions `Zeff1` (single scatterer), `Zeff6` (S1 in NEU geometry) and `Zeffn` (general case), all located in module `TOUT.FOR` .

Alternative Derivation of $x_{rms}(z)$
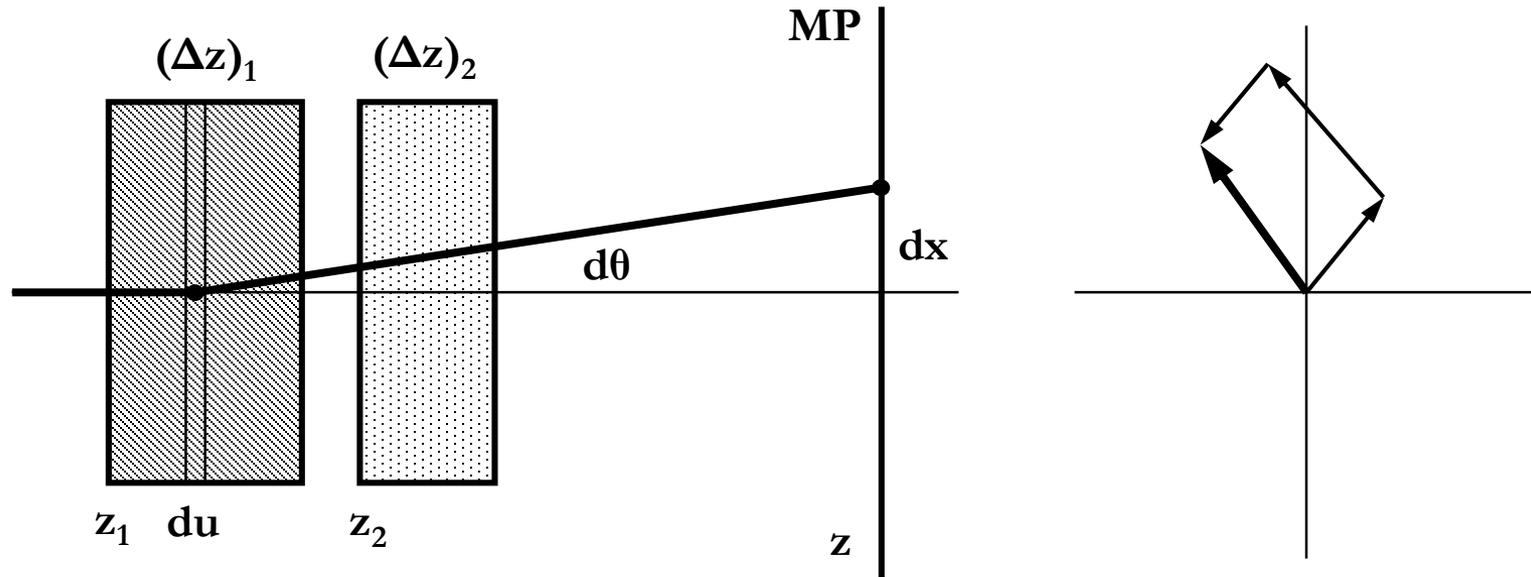
Preston and Koehler (1968) used this construction. Scatters from each element du, projected onto an arbitrary measuring plane, add in quadrature. If we use Highland's formula with the log correction for each slab taken out of the integral we find

$$x_{rms}^2(z) = \sum_{j=1}^{N} \left[\left(1 + \frac{1}{9} \log_{10} \frac{\rho_j (\Delta z)_j}{X_{0j}} \right)^2 \int_{z_j}^{z_j + (\Delta z)_j} \left(\frac{14.1(z-u)}{pv} \right)^2 \frac{\rho_j}{X_{0j}} \, du \right]$$

which is very similar in form to the Fermi-Eyges moment A_2. This formula is the basis for our subroutine Rch ('characteristic radius') in module TOUT.FOR. .

Experimental $x_{rms}(z)$
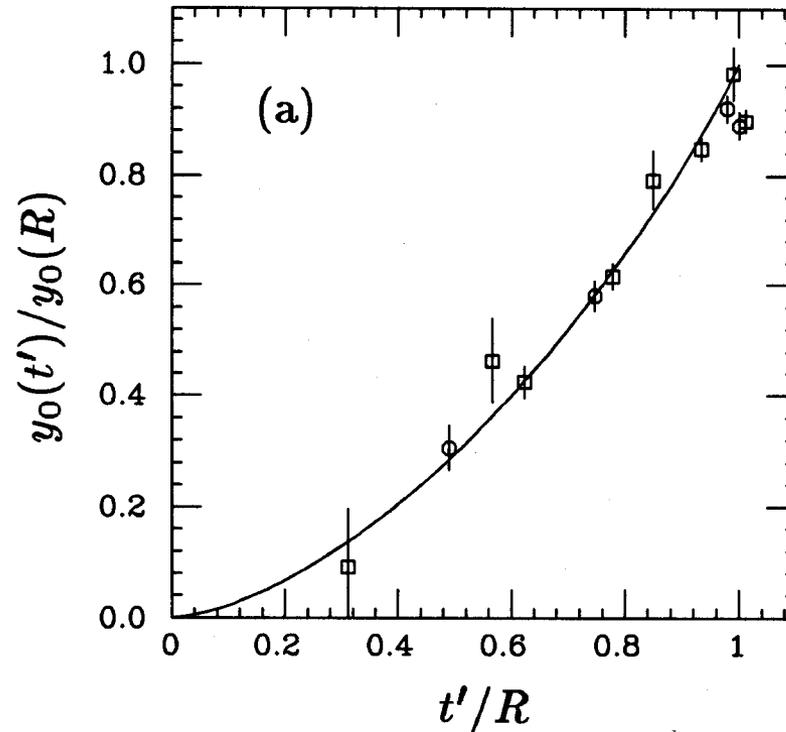
Gottschalk et al., NIM **B74** (1993) 467-490 Figure 5a: spreading of a pencil beam in a thick homogeneous target: (rms radius)/(max rms radius) v. (depth in target)/(initial mean range). Data points are measurements of Preston and Koehler for 112 and 158 MeV protons in aluminum and 127 and 134 MeV protons in water. The line represents the formula of the previous slide. It is indistinguishable from the analytic approximation of Preston and Koehler based on the 'ICRU35' scattering power.

Summary

The phase space picture is a useful way to visualize beam propagation in a system of scatterers and drifts.

Fermi-Eyges theory is a quantitative description of beam propagation in such a system. It has been generalized to mixed slab geometry, in which each slab is homogeneous and infinite in the transverse coordinates x and y. In that case the 1/e probability density contour in x , x' space (phase space) is an ellipse described by three parameters A_0, A_1 and A_2. These may be computed: they are simply the moments of scattering power namely $T(z) \equiv d<\theta^2>/dz$.

Scattering power is a complicated concept since, even when energy loss is negligible, $\Delta<\theta^2>$ is not exactly proportional to scatterer thickness (next lecture).

Though Fermi-Eyges theory is derived by writing down and solving a differential equation for the probability density in phase space, many of its results can be found more directly. We have described these, and used them to write down an alternative formula for the angular confusion in transverse penumbra problems.

The fundamental limitation of FE theory is that it deals only with Gaussian fluence distributions and the fluence in any but the most trivial proton radiotherapy problems is not Gaussian, due to transverse inhomogeneities or collimators. In practice, FE theory or its equivalent is used to predict the properties of pencil beams which are then combined to describe the actual situation.