Chapter 5. Beam Delivery Systems*

William T. Chu
Lawrence Berkeley National Laboratory
University of California, Berkeley

We must fully exploit the physical and biological advantages of ion beams to realize their therapeutic potential. In this section, we discuss technologies to deliver optimum radiation dose distributions, i.e., accurately delivering a prescribed dose distribution to the treatment volume while minimizing the radiation dose delivered to surrounding sensitive, normal structures of the body.

5.1. Dose localization characteristics of ions

Energetic ion beams (protons and heavier ions), penetrating materials in the beam path, slow down primarily by losing their kinetic energy as they ionize atoms in the medium. The energy loss of a particle per unit length (usually expressed in keV/µm in tissue) may be expressed by the Bohr-Bethe formula:

\[
\frac{dE}{dx} = \frac{4k^2Z^2n_e}{m_e v^2} \left( \frac{2m_e v^2}{I} \right),
\]  

(5-1)

where \(Z\) is the effective charge and \(v\) is the velocity of the projectile particle, \(n_e\) is the electron density of the medium, and \(I\) is the mean ionization potential. Relativistic terms and low energy correction terms are not included in Eq. (5-1). It shows that the energy loss per unit path length (or stopping power of the medium), is inversely proportional to the square of the velocity of the projectile particles, which results in a rise to a sharp maximum in ionization near the end of the range where the projectile velocity approaches \(v=0\). This rise is known as the Bragg peak [Bragg and Kleeman, 1904].

When a pencil beam of energetic ions goes through materials in the beam path, including the patient body, the beam particles scatter. The scattering makes the lateral dose falloff becomes

* This work was supported by the Director, Office of Science of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.
less steep, and therefore the beam penumbra becomes wider. As the scattering is stochastic in nature, it causes straggling in the range of a monoenergetic particle beam broadening the Bragg peak and making the distal dose falloff less steep. In treating human cancer with radiation, it is necessary to deliver desired therapeutic dose to the treatment volume while sparing adjacent healthy critical tissue. Dose localization characteristics is an important property to be exploited in clinical applications as many tumors are immediately adjacent to critical organs, which must be spared of unwanted radiation as much as possible. Physical advantage of heavier charged particles is that they suffer significantly less multiple scattering than lighter ones. The heavier charged-particle beams exhibit sharper lateral as well as distal dose falloff as at the treatment field boundary. Based on these physical advantages of dose localization, Robert R. Wilson proposed as early as 1946 the rationale for using accelerated proton and heavier ion beams for radiotherapy of human cancer and other diseases [Wilson, 1946]. The clinical expectation of using the energetic ions is increased local control with a decrease in normal-tissue complications.

Fig. 5.1. The bottom graph shows depth dose curves for three types of radiation: photons (21 MeV), protons (148 MeV) and carbon ions (270 MeV/u). The particle beams show sharp Bragg peaks at the end of their range. The top graph shows the broadening of the pencil beam for the three types of radiation.

As shown in Fig. 5-1, when a beam of monoenergetic ions enters the patient body, the depth-dose distribution is characterized by a relatively low dose in the entrance region (plateau) near the skin and a sharply elevated dose at the end of the particle range (Bragg peak). The range can
be adjusted, so that the Bragg peak occurs in an extended treatment volume. The radiation dose abruptly decreases beyond the Bragg peak as all the primary projectile particles stop; the fragments give rise to a small exit dose (tail dose).

5.2. Ion Beam Transport and Treatment-Beam Delivery Systems
Adjusting the beam optics to transport a given particle beam to a desired location with desired parameters is called “tuning” a beam line. Beam tuning in clinical facility must be reliable, reproducible and efficient. The analyzing power of the transport system is also important for physically separating particles with different momenta and charges for radiation treatment.

In a typical treatment room, the beam particles, after passing through a vacuum window at the end of the beam transport system, travel through several kinds of devices and drift spaces before entering the patient. These devices change the beam range, modulate the range, spread the beam laterally, and shape its lateral profile. Controlling of the treatment beam delivery obviously requires dose measuring devices and beam monitoring devices [Chu, Renner and Ludewigt, 1993], which are not discussed here.

Materials and drift spaces in the beam path can potentially modify an ion beam because the material scatters the beam particles (multiple scattering), smears its energy (range straggling), and fragments some portion of it (nuclear fragmentation). Minimizing scattering will narrow the lateral dose falloff (penumbra), minimizing straggling will sharpen the distal dose falloff, and minimizing the fragmentation will maintain the beam quality. On the other hand, to cover the extended treatment volumes in clinical applications, we may intentionally broaden the pencil beams laterally and modulate the stopping range.

5.3. Bragg peak, Spread-Out Bragg Peak (SOBP), and distal dose falloff
The dimensions of a particle beam stopping inside the treatment volume in patient are primarily determined by the emittance (spot size [ divergence) and the energy spread of the incident beam, which are significantly modified by intervening materials in the beam path, including the patient body. The width of the Bragg peak is primarily determined by the initial energy spread
of the incident beam particles and the range straggling in intervening materials. For example, for a narrow pencil beam of monoenergetic carbon ions of a range of 20 cm in water, the variance of the longitudinal spread due to the straggling is $\sigma_z \approx 0.06$ g/cm$^2$, and that of the mean deflection due to multiple scattering is $\sigma_k = \sigma_c \approx 0.11$ g/cm$^2$ [Litton, Lyman and Tobias, 1968]. Therefore, unless the beam is modified, the concentrated energy of the Bragg peak would be deposited in a small volume given by these variances. To cover an extended treatment volume with the Bragg-peak dose, the range of the beam must be modulated to spread out the Bragg peak, and the beam profile must be transversely broadened to cover the cross-sectional area of the treatment volume.

### 5.3.1. Range Straggling

Range straggling is the dispersion of the path length of a particle beam due to statistical fluctuations in the energy-loss process [Bichsel and Yu, 1972]. The end result is to produce a smearing of the range of the stopping particle beam. For a particle traveling in a direction $z$, with a mean range $R$ (cm), the distribution of ranges, $s(z)$, is Gaussian,

$$s(z) = \frac{1}{\sqrt{2\pi\sigma_z}} e^{-\frac{(z-R)^2}{2\sigma_z^2}}$$  \hspace{1cm} (5-2)

Tobias et al. [1980] have shown that the variance $\sigma_z$ in the beam range is experimentally parameterized to be:

$$\sigma_z(cm) = 0.012 R^{0.951} A^{0.5}$$  \hspace{1cm} (5-3)

where $A$ is the projectile particle mass number. In the region where this formula is valid ($2 < R < 40$ cm), $\sigma_z$ is almost proportional to the beam range, $R$, and inversely proportional to the square root of the particle mass number, $A$ (see Fig. 5.2). For example, for protons with a range of 30 cm passing through 25 cm of water, the variance in the range becomes 2.6 mm, or the distal dose falloff distance from 90% to 10% dose level becomes 4.4 mm, which may be unacceptable in certain clinical applications. For a carbon-ion beam of the same range, the figures would be reduced by $1/\sqrt{12}$. When expressed as a percentage of the total range, the rms range straggling is $\approx 1.0\%$ for protons, $\approx 0.5\%$ for helium ions, $\approx 0.3\%$ for carbon ions and $\approx 0.25\%$ for neon ions.
for most ranges of clinical interest. Removing as much material from the beam line as possible may reduce range straggling and multiple scattering. For example, magnetic deflection can eliminate the needed material to spread the beam in a scattering system. Changing the accelerator energy can eliminate material degraders used to change the energy of the beam.

![Graphs showing variance of straggling and multiple scattering as a function of depth for various heavy charged particles.](image)

Fig. 5.2. (Top) The variance of straggling, $s_z$, and (Bottom) the variance of multiple scattering, $s_y$, as a function of depth for various heavy charged particles.

### 5.3.2. Range Extension and Spread Out Bragg Peak (SOBP)

#### 5.3.2.1. LET, OER and RBE

Failures in local control of tumors treated with conventional radiation are often due to its inability to completely eradicate anoxic (deprived of dissolved gaseous oxygen) tumor cells, which are resistant to conventional radiation. Beams of ions with higher atomic charges ($Z$) produce higher linear energy transfer (LET, often expressed in an energy loss per unit thickness...
of water-equivalent material, keV/µm) exhibit biological advantages of lower oxygen enhancement ratio (OER) and higher relative biological effectiveness (RBE). These result in a high likelihood of an enhanced therapeutic potential when compared with protons or conventional radiation. Tobias and Todd [1967] gave the scientific justifications for utilizing light and heavy ion beams to take advantages of the enhanced RBE and reduced OER at Bragg peak and reduced lateral penumbra.

Clinical trials treating cancer patients using ion (mostly neon) beams took place at the Bevalac from 1977 until 1992 prior to the closure of the accelerator in 1993 [Castro, 1995]. Clinical trials using carbon ions have been conducted at the Heavy Ion Medical Accelerator in Chiba (HIMAC) of the National Institute of Radiological Sciences (NIRS), Chiba, Japan, at the Hyogo Ion Beam Medical Center (HIBMC) in Hyogo, Japan, and at the Gesellshaft für Schwerionenforschung (GSI) in, Darmstadt, Germany.

5.3.2.2. Variable Range Shifters

An ion beam extracted from an accelerator and transported to the patient treatment area is usually prepared in a small diameter (rms radius <1 cm) and a small spread in energy. The beams from synchrotrons have a typical value energy spread $\Delta E/E \approx 10^{-4}$ for a given extraction pulse, and $\Delta E/E \approx 10^{-3}$ when averaged over many pulses. If a monoenergetic beam is introduced into an absorbing medium, such as soft tissue, it will stop in a narrow region at the end of its range, $R$, which is given by the integral over the stopping power, $S$, of the medium:

$$R = \frac{\int E \, dE}{S}$$  \hspace{1cm} (5-4)

The stopping power, $S$ (or $dE/dx$) measured in MeV/cm is given by:

$$S = 0.307 \frac{Z^2}{A} \frac{Z}{A} I(\bar{\rho}),$$  \hspace{1cm} (5-5)

where $Z$ is the charge number and $\bar{\rho}$ is the velocity of projectile particle. And, for the medium, $Z'$ is the nuclear charge, $A'$ the atomic weight, $\bar{\rho}$ the density, and $L$ the stopping number per unit mass [NAS/NRC, 1964]. The range of completely ionized particles traveling with the same velocity scales with $A/Z^2$, where $A$ is the mass number of the projectile. This explains the longer ranges of secondary particles created in a nuclear fragmentation since they have practically the
same velocity as the primary particle before it fragmented. The range, energy and energy loss for different ions in various absorbing media are tabulated in reference publications [see, for example, Ziegler, 1980].

5.3.2.3. Spread Out Bragg Peak (SOBP)

We can accelerate ion beams of therapeutic interest to reach a desired depth in the patient body. Because the treatment volumes of clinical interests are usually thicker than the width of the Bragg peaks, the energy of the incident beam has to be modulated so that the Bragg-peak dose is deposited throughout the treatment volume from its distal edge to the proximal edge. Clinical requirement is to deliver a biologically uniform dose distribution throughout the entire treatment volume. We may stack Bragg curves of different ranges to create a SOBP, by changing either absorber thickness or accelerator energy. In beam scanning, we may divide the treatment volume into many layers in depth, and scan layer by layer by changing the residual range of the beam. Because the width of the pristine Bragg peak is only a few mm and the target thickness may be even larger than 16 cm, we will need an inconveniently large number of layers to create an appropriate SOBP. In order to reduce the number of layers, the Bragg peak is spread out to a ‘mini-peak’ (also called ‘small SOBP’) of a 5–10 mm thickness, and a reduced number of mini-peaks is axially stacked by changing their range (Fig. 5.3). Obviously, we must design the slope of the mini-peak to obtain correct SOBP shape using the range stacking [Chu, Ludewigt, and Renner, 1993; Schaffner et al., 2000].
FIG. 5.3. Biological (solid line) and physical (dashed line) dose distribution of a 50 mm SOBP delivered by superposing small SOBPs produced by a ridge filter design following the Gaussian approach. The Gaussian width is 4.3 mm and the thickness of the range shifter plates equivalent to 5.0 mm water. The biological SOBP produced by the conventional broad ridge filter is shown for comparison (dotted line). From Schaffner et al. [2000].

As illustrated in Fig. 5-3, the distal part of the SOBP is almost entirely made up of the Bragg peak; whereas, as one moves toward the proximal peak, the more and more of dose comes from the plateau dose while the contribution of the Bragg peak diminishes. The value of RBE is a function of the penetration depth, and rises very rapidly to a maximum at the Bragg peak (Fig. 5.4). For example, along a carbon ion track, LET values vary from approximately 10 keV/µm in plateau to a few hundred keV/µm in the stopping region. To achieve a biologically uniform dose across the SOBP, the physical dose distribution must be adjusted to slope down as one moves from the proximal to the distal part of the SOBP (Fig. 5.4). We may design an appropriate SOBP using algorithms developed to calculate the biological effects of mixed LET radiation [see, for example, Zaider and Rossi, 1980; Petti, Lyman and Castro, 1991].
We must take into consideration other factors to design an appropriate physical dose distribution in order to obtain a biologically uniform dose throughout the SOBP. For the ion beams, the primary nuclei may fragment in the absorbing material and turn into lighter nuclei, which have longer ranges and lower RBE. Fragments create a tail dose beyond the distal fall-off of the SOBP. As depicted in Fig. 5.5, the fragment dose is not confined to only the tail region; the projectile particles fragment throughout the medium, and therefore the fragment dose is also embedded throughout the SOBP [Llacer et al., 1984].
5.3.2.4. Ridge filters

A simple ridge filter performs the same function of range stacking by using a passive absorbing material of various thicknesses. Fig. 5.6 shows an example of a bar ridge filter, which was developed at LBNL for spreading the Bragg peak of neon-ion beams. The cross-section of a ridge is made up of several steps of absorbing material. Similar to the beam stacking method described above, this type of ridge filter also stacks different amplitudes of Bragg peak at several residual ranges. The “ruffles,” shown at left of Fig. 5.6, spreads the Bragg peak into a mini-peak (small SOBP), and a biologically uniform and smooth SOBP results when the small SOBP is spread out by the bar ridge filter. In designing the filter, we must take into consideration the effects of scattering, RBE and fragmentation as functions of the penetration depth, so that SOBP exhibits biologically uniform dose distribution throughout the treatment volume. The idea to intercept a monoenergetic beam with variable thickness of absorbing materials was incorporated in designing spiral ridge filters and rotating propellers for proton therapy [for a review, see Chu, Ludewigt, and Renner, 1993].
5. Beam Delivery

Fig. 5.6. A bar ridge filter, shown at right, modulates neon-ion beams to form a SOBP. Each ridge is made up by several steps. The “ruffles,” shown at left, spreads the Bragg peak into a mini-peak (small SOBP). Biologically uniform SOBP results when the small SOBP is spread out by the bar ridge filter.

For heavier particles, such as carbon and neon ions, ridge filters are typically made from a metal, such as brass or copper, because heavier particles suffer less multiple scattering than the light ones. Other reasons to use metal are: for a same absorbing power, higher-Z materials fragment the ions less than the low-Z materials, and metal filters are compact and easier to be machined.

5.4. Lateral broadening of ion beams

Accelerated ion beams extracted from an accelerators are usually transported in a well-focused narrow pencil beam, as the small size of the beam makes the beam transport magnets and beam monitors compact. Upon entering the treatment room, the beam has to be modified, i.e., the particle range modulated and the pencil beam laterally broadened to cover extended treatment volumes.

Clinical requirements call for large uniform radiation fields, often as large as 30 cm x 30 cm, and occasionally even larger. This does not imply that clinicians routinely use such large areas; the large-area capacity allows accommodation of treatment fields of various irregular shapes and sizes in different orientations. To obtain such large fields, we must spread out the beam laterally,
that is, in the direction perpendicular to its central ray. The aim is to produce a large field that covers any treatment field with a uniform dose with a variation of less than ±2.5%. While laterally spreading the beam, we must take into consideration of optimization of such beam characteristics as sharpness of the lateral dose falloff, beam utilization efficiency, dose rate, neutron production, beam fragmentation, the ease of beam tuning, repeatability and stability of the delivered dose distributions, and patient safety.

The simplest beam broadening method uses a single scattering foil or plate. When an ion beam traverses a medium, each particle is deflected by many small-angle elastic scatterings mainly due to elastic Coulomb scattering from the orbital electrons and nuclei within the medium. The number of particles $N$ in solid angle $d\Omega = \sin\theta d\theta d\phi$ is broadened to a 2-dimensional Gaussian-like distribution for small scattering angles, $\theta$ [Segre, 1953]:

$$\frac{d^2N}{N} = \frac{1}{2\pi q_0^2 e^{\frac{1}{2}q_0^2}} \sin\theta d\theta d\phi, \quad (5.6)$$

and the angular deflection distribution is-

$$\frac{dN}{N} = \frac{1}{\sqrt{2\pi q_0^2}} e^{-\frac{q_0^2}{2}} d\theta_{plane} \quad (5.7)$$

where $\theta$ is the polar angle that is measured from the incident direction in 3-dimensional space, and its standard deviation, $q_0$, is defined as the half-width of the Gaussian distribution where the magnitude is $1/\sqrt{e} = 0.61$ of the maximum. For the derivation of the value of $q_0$, see Eq. (5-10) below.

In analyzing multiple Coulomb scattering, we must distinguish the non-projected (polar angle $\theta$) and projected ($\theta_{plane}$) angular distributions, which are related approximately,

$$\theta^2 = \theta_{plane,x}^2 + \theta_{plane,y}^2, \quad (5.8)$$

where the $x$ and $y$ axis are orthogonal to $z$ direction of the incident projectile. The angular deflections into $\theta_{plane,x}$ and $\theta_{plane,y}$ are independent and identically distributed.
For hadronic projectiles, the strong interactions also contribute to multiple scattering. The theory of Molière accurately represents the Coulomb scattering distribution [Molière, 1955; Bethe, 1953]. At larger angles (greater than a few $\theta_b$) the distribution behaves like Rutherford scattering, having larger tails than does a Gaussian distribution. For applications in radiation therapy, it is sufficient to use a Gaussian approximation for the central 98% of the projected angular distribution. We may safely ignore the larger angle scattering, and therefore we will use Gaussian approximations in the following discussions.

In Fig. 5.6, a particle, incident on a scattering material of thickness $z$, emerges with a projected angle $\theta_{plane}$ with projected spatial displacement $y_{plane}$. They are related to $\theta_b$ as following:

$$ y_{plane}^{rms} = z \theta_{plane}^{rms} = \frac{1}{\sqrt{3}} z \theta_b $$

(5-9)

where $\theta_b$ is given as-

$$ \theta_b (\text{radian}) = \frac{14.1 (MeV)}{\sqrt{k p}} Z \frac{L}{L_0} + \frac{1}{9} \log_{10} \frac{L}{L_0} $$

(5-10)

Here $p$, $\sqrt{k}$, and $Z$ are the momentum, velocity, and charge number of the incident particle, and $L/L_0$ is the thickness of the scattering medium in radiation lengths (g/cm²). Often an improved Highland formula of Eq. (5-10) is mentioned in literature [Highland, 1975; Lynch and Dahl, 1991], but the improvement really does not add anything in applications in radiation therapy.
physics. Gottschalk et al. [1993] extensively reviewed different approaches to multiple scattering as applied to radiation therapy. The radiation lengths of particles in various materials are tabulated in standard reference publications [see for example, Hagiwara et al., 2002].

Eq. (5-10) describes scattering from a single material, while the usual problem involves the multiple scattering of particles traversing many different layers and mixtures. Since it is from a fit to a Molière distribution, it is incorrect to add the individual contributions in quadrature; the result turns out to be systematically too small. It is much more accurate to apply Eq. (5.10) once, after finding L and L₀ for the combined scatterer. Lynch and Dahl [1991] have extended this phenomenological approach, fitting Gaussian distributions to a variable fraction of the Molière distribution for arbitrary scatterers, and achieve accuracies of 2% or better.

The magnitude of multiple scattering is often expressed in a projected deflection distribution in a z-y plane. At the range R, the projected deflection y of the projectile is given by:

$$s(y) = \frac{1}{\sqrt{2\pi} \sigma_y} e^{-\frac{y^2}{2\sigma_y^2}}, \quad (5-11)$$

and the variance σ_y in water is experimentally parameterized as:

$$\sigma_y (water) = \frac{0.0294 R^{0.896}}{Z^{0.207} A^{0.396}}, \quad (5-12)$$

where Z is the projectile charge, and σ_y and R are expressed in units of cm [Tobias et al., 1980]. Litton, Lyman and Tobias [1968] have presented a detailed analysis of multiple scattering of heavy charged particles pertinent to therapy applications. For ions of clinical interest, Z=A/2, and therefore Eq. (5-12) shows that σ_y is inversely proportional to A^{0.6}; as in the case of range dispersion, lateral dispersions from multiple scattering of heavier ions are less than for lighter species (Fig. 5.2).
FIG. 5.8(A). Typical values of the width of the lateral dose falloff as function of depth in water for a neon-ion treatment field. The increase of the dose falloff with depth arises mainly from multiple scattering of the beam in the patient tissue. The lines are drawn to guide the eye.

Fig. 5.8(B). Relative increase in the penumbra due to multiple scattering in water and copper for proton, helium-ion, and neon-ion beams. The variance ($\sigma$) of the lateral dose falloff is expressed in terms of the variance at the entrance of the absorber ($\sigma_0$).

Clinical applications in the past decades have resorted to several different methods for lateral spreading of the beam. They may be divided into static (or passive) and dynamic (or active) beam delivery systems.

5.4.1. Passive Beam Delivery Systems

5.4.1.1. Double-scattering beam delivery method

The single-foil scattering method provides dose distributions of acceptable variations ($\pm 2.5\%$) only in relatively narrow fields. If we try to widen the circle of utilization by using a larger
portion of the scattered beam, an excess dose would result around the central ray. To obtain a broader uniform dose distribution at the isocenter, some of these excess particles near the central ray must be removed, either by blocking or scattering. A method to achieve this goal is the double-scattering method, which was developed at Berkeley Lab to spread helium-ion beams at the 184-Inch Synchrocyclotron [Crowe et al., 1975], and for proton beams at the Harvard Cyclotron Laboratory (HCL) [Koehler, Schneider and Sisterson, 1977].

A double-scattering system depicted in Fig. 5-9 employs an occluding post of sufficient thickness to block the beam particles in the central portion of the Gaussian distribution. Immediately downstream of the occluder, the transmitted beam intensity distribution is shaped as an annulus with a null in the middle. Its projected dose profile exhibits two peaks as shown in the figure. The second scatterer, of an appropriate thickness and placed strategically, diffuses the particles in these two peaks filling the dose void in the middle, and produces at the isocenter a larger flat-dose area.

In one such system developed at HCL for the 180-MeV proton beam, the choice of two scatterers with $R_1 = 1.7 \tilde{R}$ and $R_2 = 1.3 \tilde{R}$ produced a flat proton field of a radius out to $1.5 \tilde{R}$ with a ±2.5% dose deviation. The notations used are: $R_1$ and $R_2$ are the rms radius due to multiple scattering in the first and second scatterer, respectively, and $\tilde{R}$ is the projected radius of the occluding post at isocenter. [Koehler, Schneider and Sisterson, 1977].

![Fig. 5.9. Schematic representation of a double scattering with an occluding post. The first scatterer scatters incident beam into a Gaussian (gray color shows the transmitted beam). The occluding post blocks the particles near the central ray. When scattered by the second scatterer, the “hole” in the middle is filled in to produce a flat field (gray area) at isocenter.](image-url)
scatterer scatters incident beam into a Gaussian (gray color shows the transmitted beam). The occluding post blocks the particles near the central ray. When scattered by the second scatterer, the “hole” in the middle is filled in to produce a flat field (gray area) at isocenter.

For a broader beam, we may use an annular (ring-shaped) occluder, which would transmit beam particles through the central opening and outside the occluder. The dose profile would exhibit three peaks with an annular void. When scattered by the second scatterer, the broadened three peaks fill the annular dose void and produce a large flat-dose area at the isocenter. For example, such a double-scattering system designed at HCL for 250 MeV proton beams, with a distance of 3 m from the first scatterer to the isocenter, produced ±2.5% flatness in a circular treatment area of ≈25 cm useful radius with ≈23 cm water-equivalent residual range [Gottschalk, 1986]. We could obtain a flat dose field of even larger area by using an annulus-and post occluder set, as shown in Fig. 5.10. Such a system was used at LBNL to broaden a neon-ion beam of energy per nucleon of 670 MeV to a flat field of a diameter of 20 cm.

Fig. 5.10. Schematic illustration of a double scattering system employing a ring-and-post occluder set. The first scatterer scatters incident beam into a Gaussian (gray color shows the transmitted beam). The beam transmitted by the occluder set shows absence of particles behind the post and the ring. The transmitted beam (shown in gray color) is scattered by the second scatterer, and the resulting Gaussian-like distributions add up to make a large uniform dose field at isocenter (shown in gray color).

Successively larger-area dose fields could be obtained by increasing the number of annular rings of increasing radii. A practical limit is reached when the beam utilization efficiency drops too low to perform a treatment in a reasonable time, i.e., several minutes. The flat dose area obtained by the double scattering system is circular. The beam particles remaining in this flat
circular area is shown to be at best 37%, which sets the upper limit to the beam utilization efficiency of double scattering systems. For an actual patient treatment, the circular flat field would be further collimated, resulting in an even lower efficiency [Chu, Ludewigt, and Renner, 1993].

Given the ion species, the beam energy, the beam spot size, and the beam divergence, we could design a suitable double-scattering system by specifying thicknesses and throws of the two scatterers (drift distances to the isocenter), and the physical dimensions of the occluder assembly. To ensure the proper workings of a double-scattering system, the beam profile broadened by the first scatterer must have a correct [ ] at the occluder. We must also accurately align the central axis of the beam and that of the double-scattering system to ensure the dose uniformity as the proper functioning of the system critically depends on the cylindrical symmetry. In the Berkeley Lab system, a misalignment of the beam by 1 mm at the occluder assembly produced ±7% tilt in the dose distribution at the isocenter.

Scattering characteristics of a given scatterer depend on the incident particle species (charge, Z, and atomic mass, A) and the beam energy. Therefore a double-scattering system developed for a certain beam must be modified if any of these beam parameters are changed. Providing a different double-scattering system for each time when particle species and energy are varied would be costly. Furthermore, changing the systems each time the beam is changed is not practical in clinical operations. Gottschalk proposed a method to resolve this problem, which varies the distance from the first scatterer of the occluder assembly and the second scatterer. When an absorber modulates the energy of the beam, and consequently the value of the beam width is made larger, the occluder assembly may be moved upstream nearer to the first scatterer so that the projected radii at the isocenter are proportionally increased. This process compensates the parameters in such a way that the resulting dose distribution again exhibits an acceptable deviation from the average [Gottschalk, 1986]. This method has not been utilized at any proton therapy facilities because moving the scatterers in a beam line is inconvenient, especially automatic shifting of the position of scatterers on a gantry beam line is not very practical.
Another practical point to note is that the double scattering method requires thick scattering foils, which produces secondary particles for beam particles with Z>1, which lowers the peak-to-plateau ratio and raises the tail dose. These fragments also lower the RBE and raise the OER values, thereby lowering the biological advantage. As discussed above, the beam utilization efficiency is low, i.e., below 37% and typically 20%. The low efficiency implies that a large portion of radiation is absorbed in the occluder, as well as in collimators and scatterers, resulting in increased background radiation in the treatment room. This becomes a serious problem when a double-scattering system must be placed near the patient, e.g., on a rotating gantry. Shielding needed to block unwanted radiation may become unacceptably heavy.

5.4.1.2. Contoured bi-material filter
The different scattering characteristics of ions of different atomic-mass may be exploited in the preparation of large uniform dose fields. As discussed, for a given effective thickness (g/cm²) of scatterer, a high atomic-mass material scatters less, and a comparable low atomic-mass material scatters more (Fig. 5.2). A pencil beam is laterally spread out to a Gaussian-like beam spot and is made to impinge upon the second scatterer. In order to flatten the Gaussian-like field, the beam particles near the central ray must be scattered out more than the particles further away from it. This differential scattering must be achieved while keeping the range modulation of the beam constant at all radial distances of the scatterer system. Gottschalk proposed an elegant solution: a bi-material (e.g., beryllium and lead, or copper and plastic) scatterer (Fig. 5-11). This type of “contoured” filters have successfully been produced and used at HCL and elsewhere [Gottschalk, 1987; Gottschalk and Wagner, 1989].
In a contoured filter, depending on the radial distance from the axis, particles suffer different scattering while the range is modulated by the same amount. Compare it with a ridge filters described above (Fig. 5.5), which employs successively increasing thickness of a given material. A beam going through the thicker part of the filter would suffer more scattering than those through thinner parts. The result is that the beam spot at isocenter would be a function of the residual range. Certain clinical applications call for keeping the same spot size for varying residual ranges, and we can design a bi-material filter to modulate the range of the particle beam by different amount while keeping its scattering characteristics constant at all residual ranges. Such a “compensated modulator, has been fabricated to achieve this goal for modulating the range of a 156-MeV proton beams at HCL [Gottschalk, A. M. Koehler and M. S. Wagner, 1989]. Because this type of modulator can be placed further upstream from the patient than a simple propeller the lateral dose falloff can be significantly improved.

5.4.2. Dynamic Beam Delivery Systems
A dynamic beam delivery system produces a desired radiation dose distribution, usually a large flat dose field, when a controlled extraction of the beam particles from an accelerator is strictly coupled with prescribed patterns of the motion of the beam spot inside the treatment volume. For example, two dipole magnets, placed in tandem so that their magnetic fields and the incident beam form three orthogonal directions, can direct the beam spot away from isocenter in a predetermined way to produce a desired dose distribution. We could devise a magnet with
compound coils, or a rotating set of permanent magnets, to accomplish similar effects. These dynamic systems have a distinct advantage over the scattering systems in minimizing the material in the beam path, reducing the beam range straggling, reducing fragmentation of the incident beam particles, and decreasing the background radiation for the patients.

One of the earliest dynamic beam delivery systems was the Lissajous pattern maker, which was developed at Uppsala for proton beams [Larsson et al., 1959]. Two dipole magnets, placed in tandem with their magnetic field directions orthogonal to one another and to the beam direction, were energized sinusoidally with two different frequencies that are related to each other in a ratio of integers. A beam going through the system would draw a rectilinear Lissajous pattern in the plane at isocenter. If the beam spill level is held constant, a limited region in the middle part of the Lissajous pattern would exhibit a dose distribution of an acceptable uniformity since the speed of the beam spot is approximately constant there.

### 5.4.2.1. Wobbler systems using dipole magnets

The simplest form of wobbler beam delivery systems uses one rotating dipole. A beam incident along the central axis of a dipole is deflected in a certain angle, and if the dipole is rotated around the central axis of the incident beam, the exiting beam will wobble around the initial direction and produce an annular-shaped dose distribution in a plane at isocenter. The resulting dose field is similar to that produced by the double scattering method using a post occluder as shown in Fig. 5.9. Actually, Koehler, Schneider and Sisterson [1977] considered such a wobbler before putting the double-scattering system in clinical use at HCL.

A modern wobbler system consists of two dipole magnets placed in tandem with their magnetic field directions and the beam direction are orthogonal to each other (Fig. 5.13(A)). When the magnets are energized sinusoidally with the same frequency but with a 90 degree phase shift between them, and their magnetic field strengths are properly controlled, a beam entering the system along the common axis of the wobbler magnets would emerge from the wobbler with the beam direction wobbling around the central beam axis, and “paints” an annular-shaped fluence distribution as shown in Fig. 5.13(B). We can change the diameter of the fluence annulus by adjusting the amplitudes of the magnet currents. A large area of uniform dose is obtained by
painting the treatment area in several concentric annuli with different diameters, each with a certain predetermined fluence fraction, as schematically depicted in Fig. 5.13(B).

<table>
<thead>
<tr>
<th>Fig. 5.13 (A). A wobbler beam delivery system is schematically shown. The beam entering the axes of the wobbler magnets wobble abound the incident direction, and “paint” an annular shape at a plane at isocenter.</th>
<th>Fig. 5.13 (B). This diagram illustrates making a large field by summing wobbled beams of two different radii and one non-wobbled beam.</th>
</tr>
</thead>
</table>

NIRS in Chiba, Japan developed a wobbler system for its 70 MeV proton beams, and the Riken Ring Cyclotron Facility in Tokyo used a similar system for carbon-ion beams with energy per nucleon of 135 MeV [Ohara et al., 1991]. A bigger wobbler system was developed at LBNL to spread out the neon-ion beams of energy per nucleon up to 580 MeV for clinical use [Renner and Chu, 1987]. More recently the Heavy Ion Medical Accelerator at Chiba (HIMAC) at the NIRS has used wobblers in developing the “broad-beam three-dimensional irradiation (BB3-DI) system,” which employs wobblers to produce large fields of carbon ion beams [Futami et al., 1999], and the “layer-stacking irradiation system” for carbon ion beams to deliver three-dimensional conformal therapy, which laterally spreads the carbon beams using a wobbler system and produces SOBP using layer-stacking method [Kanematsu et al., 2002].
A wobbler system is considered to be the simplest form of dynamic beam delivery systems, mainly because of its simple power supply. Wobbler magnets are in general powered by a simple motor generator, which could be free running, i.e., independent of the accelerator extraction and with little computer control. A scanning beam delivery system employs also two dipole magnets, but their operation is strictly controlled by a computer. The scanner power supplies must have very fast transition time when moving the beam spot from one pixel to the next. The scanner may hold a beam spot dwelling at a pixel while the system delivers the prescribed fluence. These imply that the beam spills must be controlled together with the scanning magnets. One may say that wobbling is a form of scanning, and a scanner may be readily used as a wobbler. However, doing so would not be cost effective as fabricating a wobbler costs much less than making a scanner.

To provide correct dose distributions, designing a simple powering of wobbler magnets implies several important conditions that have to be met. If we maintain constant both the beam extraction level and the angular speed of the wobbled beam spot, an annular dose distribution produced by the wobbler would have no dependence on the azimuthal angle, provided that the beam-on and beam-off points are exactly overlapped to obtain the same number of wobbles at all azimuthal angles for a given beam pulse. If the overlap is uneven, for an average of \( n \) wobbles per beam pulse, there will be a region painted over either \( n+1 \) or \( n-1 \) times, which would end up with a fractional dose deviation of \( \pm 1/n \) from an average. We may reduce the effect of this dose deviation to an acceptable level in several ways. First, increasing the number of wobbles per beam spill, \( n \), will decrease the dose deviation per spill. Secondly, by making the phases of the wobbling and the beam extraction proceed asynchronously, and by painting a given annulus over many beam pulses, the effect may be statistically diluted on the overall dose distribution since the regions of the uneven overlap will happen randomly in azimuthal angles. Next is the requirement of a large dynamic range in beam extraction levels. We may sum over several wobbles of different radii to produce a large area of flat dose. For example, if five radii are used, the largest wobble covers about 100 times the area of the smallest wobble. To complete all the different wobbles on comparable numbers of beam spills, a control of the spilled beam intensity over a dynamic range of \( \sim 100 \) will be necessary, which may increase treatment time [Renner and Chu, 1987].
The pertinent parameters to obtain flat fields employing a wobbler are: constant spill intensity as the time structures in the extracted beams will translate into a spatial variation in the dose distribution, a precise wobble radius and a correct particle-number fraction at each of the wobble radii. There are tradeoffs for the wobbler beam delivery system compared to the double-scattering system. First, thinner or no scatterers are needed in the beam path, and therefore we could obtain a certain beam range at a lower accelerator energy than that required for scattered beams. Less or no beam is lost in scatterers and occluders. As the size of the flat field is readily varied according to the port size, less beam is lost in collimation, and therefore, the beam utilization is higher than that for a scattered system. The accuracy in beam alignment is not as critical as in the case of the scattering system; however, the high stability of the beam entering the wobbler magnets throughout a treatment time is critical to provide concentric annular dose distributions. Since the effective source size, i.e., the beam spot size is small, the wobbler produces sharper lateral dose falloffs than those attainable with the scattering system. The neutron production in absorbers and collimators is curtailed, and, therefore, the shielding requirement is lower than that for the double-scattering method. This becomes an important consideration when one tries to place a beam delivery system on a rotating gantry. As a dynamic mode of beam delivery, the wobbler system requires on-line monitoring of the wobbler magnetic fields.

5.4.2.2. Specifications for scanning
In this section, we describe two scanning methods for producing large irregularly shaped fields of uniform dose distributions. We classify scanning beam delivery systems according to the ways in which the beam spot is moved: namely, continuous scanning (raster scanning) and discrete scanning (pixel scanning). Raster scanning employs a continuous but variable-speed motion of the beam spot while controlling the levels of the beam extraction. In a pixel scanning method the spot is moved to a pre-determined position to deposit a prescribed number of particles, then move the spot to the next position and the process is repeated. The pixel scanning is sometimes called the voxel (volume pixel) scanning or spot scanning.
The purpose of developing a beam scanning technology is to realize three-dimensional conformational therapy, which will enable us to utilize the full potential of ion beams. The result will be applicable to proton therapy beam delivery as well.

![Schematic drawing of a scanning system](image)

Fig. 5.14. Schematic drawing of a scanning system. An ion beam of a given energy is scanned in to produce a desired fluence distribution in a selected layer in the treatment volume. The entire treatment volume is scanned by changing energy of the beam and scan all the layers.

Scanning a beam across a treatment area may produce a large field of a specified dose distribution. In principle, scan speed and beam intensity can be varied as a function of the spot location in the treatment volume to generate the desired dose distribution. A schematic diagram of scanning is shown in Fig. 5.14. In general, a scanner consists of two dipole magnets, one for the fast scan in the $x$ direction and the other for the slow scan in the $y$ direction. (Here the $x$ and $y$ represent arbitrary orthogonal directions.) Range modulation (either by changing the accelerator energy or inserting a variable thickness absorber in the beam path) moves the stopping region of the beam spot, i.e., the distance from skin to the Bragg peak, in the $z$ direction.

Spatial characteristics of scanning have been analyzed by Leemann et al. in terms of its scanning speed and power requirement [Leemann et al., 1977]. In a homogeneous medium, if a scanning pattern is described by a distribution function of relative beam density function $F(x, y, z)$, and the contribution to the dose at $(x, y, z)$ from a pencil beam with its Bragg peak centroid at $(x', y', z')$
by a pencil beam function \( p(x-x', y-y', z-z') \), the dose \( D(x, y, z) \) is then expressed as a Fredholm integral equation of the first kind:

\[
D(x,y,z) = \int \int \int F(x',y',z') p(x-x',y-y',z-z') \, dx' \, dy' \, dz'.
\] (5-13)

**Fig. 5.15.** (Left diagram) The treatment volume is schematically shown as a rectangular box. A pencil beam has lateral \( p_x \) and \( p_y \) distributions given by scattering, and \( p_z \) is a Bragg curve. (Right diagram) Upon completion of the scanning in a layer at a certain depth, the beam range is changed and the scanning resumes in the next layer.

If the desired dose distribution \( D \) is specified at all points \((x, y, z)\) inside a treatment volume and outside, and the pencil-beam function \( p \) is known, then the desired density function \( F \) can be obtained. The three coordinate functions of \( p \) are practically independent, thus \( p \) may be written as

\[
p(x-x',y-y',z-z') = p_x(x-x') p_y(y-y') p_z(z-z').
\] (5-14)

The lateral beam profile functions, \( p_x \) and \( p_y \), are transverse dose cross section (approximately Gaussian), which are determined by the beam divergence and scattering, whereas \( p_z \) is a Bragg-peak depth dose profile of a pencil beam for a given depth.

The dose distribution \( D \) is a three-dimensional convolution of a density function \( F \) with the pencil-beam distribution \( p \), and Eq. (5-13) may be symbolically written as:

\[
D(x,y,z) = F \ast p.
\] (5-15)
When $D$ and $p$ are given, we may find an optimal $F$ that would best fill the treatment volume using pencil beams [Staples and Ludewigt, 1993]. The procedure consists of finding the density function $F$, which delivers a dose that best comply the required dose $D$ to the treatment volume, while minimizing the lateral and distal falloffs. In practice, restrictions of scanning system and beam extraction control (changing beam flux during a scan) must be taken into consideration. The optimization procedure follows closely those developed by Brahme [Brahme, Kallman and Lind, 1990] and by Lind [Lind, 1990]. The optimal density function $F$ is determined through an iteration process. We will assume the first guess of $F$ to be

$$F_0 = D_0$$

(5-16)

and the next iteration is

$$F_{n+1} = C[F_n + a(D_0 - F_n p)]$$

(5-17)

where $D_0$ is the desired dose distribution, $C$ is a constraint operator ensuring non-negative density function amplitude, and $a$ is a convergence speed parameter. This method of determining the density function $F$ satisfies that $F$ is non-negative, the resulting $D$ is never smaller than the desired dose $D_0$ throughout the treatment volume (no under-dosing), the dose outside the treatment volume is minimized, and the widths of lateral and distal dose falloffs are minimized.

As an illustration, Fig. 5.6 shows the optimization result for a one-dimensional example, where a Gaussian beam irradiates a line segment. In this example, the desired dose distribution ($D_0$) is a step-and-slope function, shown in gray area, and the density function $F$ is labeled as scan density function (white line under $D_0$). This density function $F$ ensures that the optimized dose (solid line above $D_0$) delivers full dose in the treatment volume, and the width of the dose falloff is minimized.
Fig. 5.16. One-dimensional example where a Gaussian pencil beam irradiates a line segment.

The desired dose distribution ($D_0$) is a step-and-slope function, shown in gray area, the scan density function $F$ in white lines under $D_0$, and the optimized dose distribution $D$ in solid lines above $D_0$. Based on the article of Staples and Ludewigt [1993].

The density function $F$ is the irradiation of Bragg peak density defined throughout the treatment volume and describes the amount of beam deposited in the volume with the center of the Bragg peak at a particular location. $F$ can also be viewed as a scan pattern function in raster scanning and a beam occupation distribution that can be directly used to control a pixel scanning system.

The finite beam spot size and the Bragg peak width of the beam spot obviously limit the attainable spatial resolution, which determines the sharpness of field edges, or the smallest size of inhomogeneities that can be compensated for. Conversely, a discrete scanning pattern $F(x, y, z)$, which is a collection of pencil beams, yields smooth dose distributions if the step size is kept under one half of the spot size. The LBNL raster scanner canned continuously in the $x$ direction in order to obtain a dose uniformity along the raster lines in the $x$ direction, and the $y$ scan was specified in such a way that the adjacent scan lines overlap at least one-half of the spot size. Lower limits for beam spot sizes are given by multiple scattering and are of the order of 2 ~ 3 mm for carbon or neon ions and 5 ~ 7 mm for protons for ranges in tissue between 20 cm to 30 cm. The spot sizes used in actual beam scanning are generally bigger.
5.4.2.3. Raster scanning

In a raster-scanning beam delivery system has been developed at LBNL in the late ‘80s to broaden neon-ion beams into large flat radiation fields. It was designed to scan an entire port area at a given depth during one synchrotron spill, which was spread out to a constant extraction level (flat top) of approximately 1 second long and repeated every 4 seconds [Chu, Renner and Ludewigt, 1998]. The magnets were designed to deflect a beam with a magnetic rigidity of 8.0 T-m up to ±20 cm in the horizontal and vertical directions at isocenter, which was ~6 m from the raster magnets. The fast-scan magnet was placed upstream of the slow-scan magnet. In the fast-scan magnet the beam was deflected parallel to its pole faces, and the gap of the fast-scan magnet had to be barely wide enough to pass through the beam spot. The slow-scan magnet had wider gap. The gaps of the fast- and the slow-scan magnets were 5.7 cm and 15.2 cm, respectively. Laminated iron cores of the magnets minimize the induced eddy currents during their operation.

The slow scan swept the vertical extent of the field, in slightly shorter than the synchrotron spill time. The fast scan painted the raster lines close enough with a 2-cm FWHM beam spot to ensure a dose distribution without peaks and valleys between the adjacent scan lines. An analysis showed that the accuracy of this edge matching depends on an accurate control of the magnetic fields (typically ±5×10⁻⁴) and the gradual slope of the edges of the beam spots (at least ~4 mm edge fall-off width). The fast x-scan speed was set at a constant 2400 cm/sec in the scan plane at the isocenter, which is equivalent to a frequency of 30 Hz over a ±20 cm field. The fast-scan speed was maintained constant for all sizes of fields; therefore, the scan frequency was higher for smaller fields. To allow the beam to dwell at a certain vertical position before the y-scan starts, the slow-scan power supply ran with a d.c. current offset. In order to achieve a dose uniformity of ±2.5% and minimize the introduction of non-uniformity by the magnet power supplies, the regulation of the magnet currents (which determine the rates of change in the magnetic field) was controlled to approximately ±0.25%. The required dose uniformity sets the specifications of the circuit control parameters: a closed loop d.c. feedback gain for both systems of at least 60 db and a unity gain bandwidth of 10 kHz for the fast-scan system and 1 kHz for the slow-scan system.
A raster scanner producing clinically acceptable compliance to the prescribed dose distribution places strict requirements on the accelerator performance. The unwanted temporal structure in the beam spills, i.e., the beam flux as a function of time, translates into spatial irregularities in the scanned radiation field. The extraction and transport optics must not change while the beam is extracted at different flux levels.

Fig. 5.15. (A) Schematic diagram depicting the beam scanning. Scanning the layer M with the Bragg peak introduces plateau dose into a part of the upstream layer N. Consequently, when the layer N is scanned with the Bragg peak, different doses must be deposited in various parts of the layer. Also, note that scanning the layer N deposits tail dose into all downstream layers. (B) Schematic diagram showing the stacking of mini-peaks to scan the entire treatment volume.

The raster scanning system has several advantages over the scattering system as well as the wobbler system. The raster scanning system maintains high beam quality of ion beams as it uses no absorbing material in the beam path. The beam utilization efficiency is higher as rectangular fields of various aspect ratios conform better with many irregularly-shaped ports than circular fields produced by a wobbler or a double scattering system. Going beyond the simple rectangular scans, by varying the extents of each fast scan, a raster scanner can produce irregularly shaped fields, and provide better conformations of the radiation fields to irregular treatment volumes.

When scanning a given layer by the Bragg peak, other layers upstream would be also irradiated by its plateau dose, while the layers downstream would receive its tail dose. Therefore,
delivering a uniform dose in each layer of an irregularly-shaped volume will not add up to a uniform dose distribution. In discussing above the density function (scanning pattern function) $F$, we used the term, “to scan a layer at a given depth,” to mean to scan with a pencil beam with its Bragg peak centroid at that depth.

As indicated in Fig. 5.15(A), to deliver a uniform dose into an irregularly-shaped volume, certain predetermined pattern of non-uniform dose must be delivered in each layer. Advanced therapy planning could provide the required prescription of dose distribution ($D_0$) in each layer. The raster scanning control system could integrate the variations of scan speed and the beam flux to deliver the planned distribution ($D$). The presence of tissue inhomogeneities in the beam paths further complicates the situation; however, their influence can readily be handled by an advanced raster-scanning algorithm.

Simple analyses show that, to achieve the compliance of a delivered dose distribution ($D$) to the specified dose distribution ($D_0$), raster scanning requires a variation of a factor of \textasciitilde20 in the density function $F$. A more detailed analysis shows that a factor of \textasciitilde7 variation in $F$ could deliver clinically acceptable dose distribution in irregularly-shaped treatment volumes. Obviously, to be clinically useful, a raster scanning system should be developed to integrate the variations in scan speed and beam flux. Probably the control system for such an “intensity-controlled raster scanning” method would be more complex than a pixel scanning system [Haberer et al., 1993]. Pixel scanning system and its use to achieve three-dimensional conformal therapy delivery through intensity modulated particle therapy (IMPT)# will be discussed in the next chapter. We note that, to overcome many complexities in stacked layer scanning methods, GSI has proposed a “depth scanning” algorithm, which is a fast intensity-controlled longitudinal scan in the beam direction. In this method, the scanner positions a beam of sufficient energy to reach the distal surface of the treatment volume, and the range is rapidly modulated to delivery a biologically uniform dose upto the proximal surface of the volume. [Weber, Becher, and Kraft, 2000].

# Actually the variable that is modulated in intensity modulate particle therapy is the number of particles, and is not intensity, of the beam. The term IMPT derives from intensity modulated radation therapy (IMRT), which varies the intensity of the radiation.
REFERENCES


