Inverse problem of linear combinations of Gaussian convolution kernels (deconvolution) and some applications to proton/photon dosimetry and image processing

W Ulmer

Department of Radiooncology, Klinikum Frankfurt/Oder and MPI of Biophysical Chemistry, Göttingen, Germany

E-mail: waldemar.ulmer@gmx.net

Received 17 October 2009, in final form 4 May 2010
Published 4 June 2010
Online at stacks.iop.org/IP/26/085002

Abstract

The deconvolution of a single Gaussian kernel is extended to a sum of Gaussian kernels with positive coefficients (case 1) and to a Mexican hat (case 2). In case 1 the normalization requires the sum of the normalized Gaussian kernels to be always 1, i.e. \( c_0 + c_1 + c_2 + \cdots = 1 \). Each coefficient satisfies \( c_k > 0 \).

In case 2 (Mexican hat) the properties \( c_0 + c_1 = 1 \) with \( c_0 > 1 \) and \( c_1 < 0 \) hold; \( c_1 = 1 - c_0 \) has to be accounted for the normalization. We discuss examples of the deconvolution of both cases. Case 1 is considered in an analysis of transverse profiles (protons and photons) and in the deconvolution of CT images to eliminate scatter. Case 2 is applied to a proton Bragg curve measured by an ionization chamber and a diode detector. In the domain of the Bragg peak there is a different physical behavior between both measurement methods.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

In many problems of physics and applications we have to deal with Gaussian convolutions. Examples are scattering problems in image processing, image-guided radiotherapy, energy straggling and lateral scatter of Bragg curves (particle therapy), influence of the finite detector size on profiles, etc. Besides the mentioned aspects emerging in radiation physics, Gaussian convolutions and the related inverse problems (deconvolutions) also occur in transport phenomena (diffusion, heat) of statistical physics (Saitoh 2004, Matsuura and Saitoh 2006) and quantum mechanical tasks, i.e. path-integrals, many-body problems and nuclear physics (Feynman and Hibbs 1965, Ulmer and Hartmann 1978, Ulmer 1980, Markov 1989, Primas...
In order to derive inverse formulas of the heat and diffusion transport Matsuura and Saitoh (2006) have used the exponential operator formulation that we have previously considered (see equation (15), Ulmer and Kaisssl 2003).

It is well known that a Gaussian convolution satisfies the integral transform (in one dimension):

$$
\varphi(z) = \int K(\sigma, u - z) \Phi(u) \, du
$$

Equation (1) provides a connection between the source function $\Phi$ (input) and the image function $\varphi$ (output). The extension to three dimensions is straightforward (see also section 2); the rms value $\sigma$ has to satisfy $\sigma \geq 0$ (the limit case $\sigma \rightarrow 0$ transforms equation (1) to a $\delta$-distribution). The source/image function according to equation (1) may be applicable in very different fields or problems. Thus, the connection between an idealistic, passive ‘point-detector $D_p$’ and an active ‘finite-size detector $D_m$’ has been studied by many authors in the past decade (Garcia-Vicente et al 1998, 2000, Bednarz et al 2002, Crop et al 2009, Das et al 2008, Palmans 2006, Ulmer and Kaisssl 2003). Then we have to associate $D_p$ with $\Phi$ and $D_m$ resulting from dose measurements with $\varphi$. Since detectors/ionization chambers do not only receive signals from a radiation field, but they also interact with it by scattering processes, a Gaussian convolution kernel and not simply a finite size kernel appears to be adequate (Garcia-Vicente et al 1998, 2000). If the rms value $\sigma$ is determined (or estimated) by a reliable method (e.g. Monte Carlo calculation), then the inverse relation of equation (1) is able to determine $D_p$ from the measurement value $D_m$ with the help of the kernel $K^{-1}(\sigma, u - z)$:

$$
\Phi(z) = \int \varphi(u) \cdot K^{-1}(\sigma, u - z) \, du.
$$

There are two possible representations of $K^{-1}$ (Ulmer and Kaisssl 2003):

$$
K^{-1}(\sigma, u - z) = \sum_{n=0}^{N} c_n(\sigma) \cdot H_{2n} \left( \frac{u - z}{\sigma} \right) \cdot K(\sigma, u - z)
$$

$$
K^{-1}(\sigma, u - z) = \delta(u - z) + \sum_{n=1}^{N} c_n(\sigma) \cdot H_{2n} \left( \frac{u - z}{\sigma} \right) \cdot K(\sigma, u - z)
$$

$$
N \rightarrow \infty; c_n = (-1)^n \sigma^{2n} / (2^n / n!)
$$

$$
N \rightarrow \infty; c_n = (-1)^n \sigma^{2n} \cdot (2^n - 1) \cdot (4^n \cdot n!)^{-1}
$$

In both representations, $H_{2n}$ denotes a Hermite polynomial of even order, i.e. the inverse kernel $K^{-1}$ can be regarded as a special generalized Gaussian convolution kernel, since it contains two-point Hermite polynomials $H_{2n}(u - z)/\sigma$ in a rigorously defined order. With regard to practical applications we have to restrict $N$ to a finite upper limit. The kernel $K(\sigma, u - z)$ and its inverse $K^{-1}(\sigma, u - z)$ satisfy the relation

$$
\int K(\sigma, u - z') \cdot K^{-1}(\sigma, u - z) \, du = \delta(z - z').
$$

Equation (5) is not restricted to Gaussian kernels. Without any completeness we mention a further application of formulas (1)–(4) to radiation physics. The transverse profile of proton beamlets is described by one Gaussian kernel, if the Highland approximation (Highland 1975, Gottschalk et al 1993, Hollmark et al 2004) is assumed. Based on problems of absolute

1989, Ulmer and Matsinos 2010).
dosimetry, objections have been put forward against this approximation (Kusano et al 2007, Pedroni et al 2005, Schaffner 2008, Tourovsky et al 2005, Ulmer and Matsinos 2010). Thus, a possible improvement of the description of the scatter behavior of protons would be obtained, at least, by two Gaussian kernels in order to yield a better approximation of the multiple scatter theory (Bethe 1953, Molière 1955).

It is rather evident that the Compton scatter of photon beams (MV domain) cannot be described by a single Gaussian kernel, since long-range lateral scatter of $\gamma$-quanta induced either by large-angle or multiple small-angle processes (i.e. scattered photons may suffer repeatedly further scattering to yield finally large-angle scatter contributions) leads to tails with different ranges. Therefore, the resulting overall scatter kernel $K_s$ can be represented by a sum of three kernels, which differ by the $\text{rms}$ values and related weights:

$$K_s = c_0 \cdot K_0(\sigma_0, u - x, v - y) + c_1 \cdot K_1(\sigma_1, u - x, v - y) + c_2 \cdot K_2(\sigma_2, u - x, v - y). \quad (6)$$

This kernel has been studied by Ulmer et al (2005) to develop a superposition/convolution algorithm for photon beams. It should be mentioned that the following properties have to be satisfied: $\sigma_0(z) < \sigma_1(z) < \sigma_2(z)$ and $c_0 + c_1 + c_2 = 1$. Since the photons of the MV domain and protons show a preference of small-angle scattering, the relation $c_0 > c_1 > c_2$ usually holds. All three kinds of $\sigma$ values depend on the depth $z$; equation (6) has been extended to two dimensions. The inverse kernel $K_s^{-1}$ of the kernel $K_s$ will be developed in the next section; it represents a possible tool in image-guided radiotherapy (IGRT). Although we mainly deal here with inverse problems of equation (6), which have not yet been elaborated at the present state-of-art, we should point out that in many problems of deconvolutions, fast Fourier transform (FFT) together with Wiener filters is applied. A very concise paper on the Fourier-based deconvolution and filter functions has been given by Ming Fang et al (1994).

However, some critical aspects result from Fourier-based deconvolution in connection with step functions (see e.g. the applications given in tin section 3.1). These well-known problems of FFT have been previously discussed (Garcia-Vicente et al 2000, Liu et al 2008, Ulmer and Kaissl 2003). We should also point out that there exist convolution/deconvolution problems in image processing with regard to the kinetics of perfusions (e.g. PET), where the presented methods with combinations of Gaussian kernels are hardly applicable, and diffusion methods (Saitoh 2001, 2004) appear to be too complicated. For this field the EM algorithm (Figueiredo and Nowak 2003, McLachlan and Krishnan 1997, Tuminonen 2004) has proven to be valuable.

2. Theoretical methods: convolutions and deconvolutions of Gaussian kernels

2.1. Mathematical and quantum statistical aspects of convolution kernels and inverse problems (deconvolutions)

The kernel $K$ according to equation (1) may either be established by nonrelativistic transport theory (Boltzmann equation) or, as we prefer here, by a quantum statistical derivation, where a relationship with the path-integral formulation of Neumann’s density matrix will be obtained (Feynman and Hibbs 1965). It is a noteworthy result (Ulmer 2007) that a quantum-stochastical partition function leads to a Gaussian kernel as a Green’s function, which results from a Boltzmann distribution function and a nonrelativistic exchange Hamiltonian $H$. An operator formulation of a canonical ensemble is obtained by the following way: let $\phi$ be a distribution (or output/image) function and $\Phi$ a source function, which are mutually connected by the operator $F_H$ according to equation (7). The exchange Hamiltonian $H$ in equation (7) incorporates the
interaction of the source field Φ (e.g. proton or photon fluence) with the environment via 
F_H(H/E_{ex}) to yield the image field ϕ:

\[
\begin{align*}
\varphi &= F_H \Phi = \exp(-H/E_{ex}) \Phi \\
H &= -\frac{\hbar^2}{2m} d^2/dz^2
\end{align*}
\]  
\tag{7}

Since \exp(-H/E_{ex}) according to equation (7) has to be dimensionless, we can identify E_{ex} with the exchange or transfer energy between proton/photon fluence Φ and environmental medium via electrons to transform Φ to the image fluence ϕ (m: mass of an electron, \hbar = h/2π, h: Planck’s constant):

\[
\begin{align*}
\varphi &= F_H \Phi = \exp(0.25 \sigma^2 d^2/dz^2) \Phi \\
\sigma^2 &= 2\hbar^2/mE_{ex}
\end{align*}
\]  
\tag{8}

Only in the thermodynamical limit (equilibrium) we can write  \( E_{ex} = k_B T \), where \( k_B \) is the Boltzmann constant and \( T \) is the temperature. It must be noted that the operator equation (8) was formally introduced (see Ulmer and Kaissl 2003) to obtain a Gaussian convolution as a Green’s function and to derive the inverse convolution. \( F_H \) may formally be expanded (Lie series) in the same fashion as the usual exponential function \( \exp(\xi) \); \( \xi \) may either be a real or complex number. This expansion is referred to as a Lie series of an operator function. In quantum mechanics such operator functions can be defined by the Hamiltonian \( H \), which has to be a continuous (\( H \) may either be bounded or unbounded) operator (von Neumann 1955). The Hamiltonian \( H \) of free particles (equation (7)) satisfies this precondition, and plane waves represent solution functions of \( H \) as well as \( F_H \).

According to spectral theorem of functional analysis we have to consider the eigenvalue problem:

\[
\begin{align*}
F_H \Phi &= \lambda \Phi \\
\Phi_k &= \exp(-ikz)/(\sqrt{2\pi})
\end{align*}
\]  
\tag{9}

With the help of equations (7)–(9) we are able to determine the spectrum \( \lambda(k) \):

\[
\begin{align*}
F_H \Phi_k &= \lambda(k) \Phi_k = \exp(-ikz)/(\sqrt{2\pi}) \\
\lambda(k) &= \exp(-\sigma^2 k^2/4)
\end{align*}
\]  
\tag{10}

The integral operator kernel according to the spectral theorem is obtained by

\[
\begin{align*}
K(\sigma, u - z) &= \int \Phi_k^*(z) \Phi_k(u) \lambda(k) dk = \frac{1}{2\pi} \int \exp(-\sigma^2 k^2/4) \exp(ik(u - z)) dk \\
K(\sigma, u - z) &= \frac{1}{\sigma \sqrt{\pi}} \exp(-(u - z)^2/\sigma^2)
\end{align*}
\]  
\tag{11}

The final result, namely equation (11) and the related integration procedure, i.e. a two-point Gaussian convolution kernel (1) as a Green’s function of the operator \( F_H = O^{-1} \), has been previously derived (Ulmer and Kaissl 2003). \( O \) and \( O^{-1} \) represent Lie series, which we shall now specify in this section (equations (14) and (15)). The only reason for the distinction between \( F_H \) and \( O^{-1} \) is that the origin of \( F_H \) is of quantum mechanical nature, i.e. the identification of \( \sigma \) according to equation (8), whereas \( O \) and \( O^{-1} \) are free of this restriction/assumption. The density-matrix formulation of (nonrelativistic) quantum mechanics is based on principles, resulting from the Schrödinger equation (in one space coordinate):

\[
H \psi(k,z) = E(k) \psi(k,z),
\]  
\tag{12}
where $\Psi(k, z)$ is the eigenfunction, $E(k)$ is an eigenvalue of equation (12) and $k$ represents a quantum number. The density matrix $\rho$ is constructed by the following equation (13):

$$\rho(z', z) = \sum_{k=0}^{\infty} \exp(-E(k)/k_B T) \psi \ast (k, z') \psi(k, z).$$  \hspace{1cm} (13)

For the statistical motion of free particles with $E_k = \hbar^2 k^2 / 2m$, we obtain $\rho(z', z) = K(\sigma, z' - z)$, if $E_{\text{ex}} = k_B T$. In the presence of a further potential $V(z)$ in the Hamiltonian $H$ according to equation (8), i.e. $H \rightarrow H = V - (\hbar^2 / 2m) \cdot d^2 / dz^2$, only equation (13) with $\rho(z', z) \neq K(\sigma, z' - z)$ holds and cannot be reduced to a convolution. In the continuous case, the summation over $k$ has to be replaced by an integral. With the help of equation (13), various properties (e.g., the partition function) can be calculated. According to Feynman and Hibbs (1965), we equate $\rho(z' - z)$ to the path-integral kernel $K(\hbar / k_B T, z' - z)$ of a free particle, useful for the calculation of perturbation problems in statistical physics (i.e. for $V(z) \neq 0$), if the formal substitution $i \hbar \rightarrow I / k_B T$ is carried out. Thus, the operator-function formalism according to relations (7)–(11) can be regarded as an operator calculus of path-integral kernels. Some essential physical aspects of the exchange energy $E_{\text{ex}}$ yielding energy/range straggling of protons on the quantum statistical basis as given above have been analyzed in other communications (Ulmer 2007, Ulmer and Matsinos 2010).

The basic formulas of all subsequent procedures are the two operator functions:

$$O = \exp \left( -\frac{1}{4} \sigma^2 \frac{d^2}{dz^2} \right)$$ \hspace{1cm} (14)

$$O^{-1} = \exp \left( \frac{1}{4} \sigma^2 \frac{d^2}{dz^2} \right).$$ \hspace{1cm} (15)

$O$ and $O^{-1}$ obey the following relation:

$$O \cdot O^{-1} = O^{-1} \cdot O = \exp \left( -\frac{1}{4} \sigma^2 \frac{d^2}{dz^2} \right) \cdot \exp \left( \frac{1}{4} \sigma^2 \frac{d^2}{dz^2} \right) = 1 \text{ (unit operator)}.$$ \hspace{1cm} (16)

Both operator functions and their actions are formally defined by Taylor expansions of the exponential function, which represent Lie series of operator functions (Ulmer and Kaissl 2003):

$$O^{-1} = 1 + \sum_{n=1}^{\infty} \frac{1}{n! \cdot 4^n} \cdot \sigma^{2n} \cdot \frac{d^{2n}}{dz^{2n}}$$ \hspace{1cm} (17)

$$O = 1 + \sum_{n=1}^{\infty} \frac{1}{n! \cdot 4^n} \cdot (-1)^n \cdot \sigma^{2n} \cdot \frac{d^{2n}}{dz^{2n}}$$ \hspace{1cm} (18)

A comprehensive review of Lie series has been given by Gröbner and Knapp (1967). Let $\Phi(z)$ be a source and $\psi(z)$ an image function, and the following relationships are obtained:

$$\varphi(z) = O^{-1} \cdot \Phi(z)$$

$$\Phi(z) = O \cdot \varphi(z)$$ \hspace{1cm} (18)
From relations (17) and (18) follows that $\Phi(z)$ and $\varphi(z)$ have to belong to the function space $C^\infty$ (Banach space), i.e. to the set of functions with derivatives of infinite order. The integral operator notation (Green’s function) of $O^{-1}$ (Ulmer 1980) and $O$ (Ulmer and Kaissl 2003) are

$$\varphi(z) = O^{-1} \cdot \Phi(z) = \int K(\sigma, u - z) \cdot \Phi(u) \, du$$

$$\Phi(z) = O \cdot \varphi(z) = \int K^{-1}(\sigma, u - z) \cdot \varphi(u) \, du$$

(19)

The integral operator kernel $K$ is the normalized Gaussian kernel according to equation (1). The integral operator $K^{-1}$, representing the inverse task of a Gaussian convolution kernel, has already been stated in equations (3) and (4). The integral operator correspondence to the relation $O \cdot O^{-1} = 1$ is given by equation (5).

The principal difference between the differential and integral operator formulation is the class of permitted functions. In the case of the operators $O$ and $O^{-1}$ acting on $\varphi$ and $\Phi$ we have to consider the Banach space of the class $C^\infty$ of functions (inclusive polynomial functions of finite order as a special case), whereas with regard to $K$ and $K^{-1}$ we only require the Banach space $L_1$ of Lebesgue integrable functions. This fact has some important practical implications, i.e. summations in finite intervals are permitted (i.e. voxel integrations). In spite of this advantage, there are many applications where the differential operator calculus is easier to handle (Ulmer and Kaissl 2003), e.g., if we regard the class of smooth functions. A further noteworthy aspect is the derivation of some fundamental properties. At this place, we only mention iterated kernels: the repeated application of $O^{-1}$ and $O$ ($n$ times) provides:

$$O^{-1} \cdot \ldots \cdot O^{-1} (n \text{ times}) = O^{-n} = \exp(0.25 \cdot n \cdot \sigma^2 \cdot d^2/dz^2) \Rightarrow$$

$$K = \frac{1}{\sqrt{n} \cdot \sigma^2 \cdot \pi} \cdot \exp(-(u - z)^2/(n \cdot \sigma^2))$$

$$O \cdot \ldots \cdot O (n \text{ times}) = O^n = \exp(-0.25 \cdot n \cdot \sigma^2 \cdot d^2/dz^2) \Rightarrow$$

$$K^{-1} = K^{-1}(\sigma_n, u - z) \quad (\sigma^2 \Rightarrow \sigma^2 = n \cdot \sigma^2)$$

(21)

With respect to the half-width parameter $\sigma$, we have not made any assumption. According to equation (8), we may identify $O^{-1}$ with $F_{\text{Hf}}$ then we obtain

$$\sigma = \sqrt{2} \cdot h/\sqrt{m \cdot E_{\text{cs}}}$$

(22)

However, in many problems involving statistics, we do not have to deal with statistical quantum mechanics and $\sigma$ might not satisfy relation (22); consequently, it may be regarded as a free parameter with the dimension of a length (e.g., in statistical mechanics/thermodynamics, etc). It is also possible that in statistical applications beyond the scope of physics, $\sigma$ and the quantity $u - z$ may have a completely different meaning.

An additional result, which is easily derived from the operators $O$ and $O^{-1}$, is that $\sigma$ does not have to be a constant parameter; it may also depend on the variable $z$. Such a situation occurs in proton dosimetry, where the energy straggling parameter $\tau_{\text{stragg}}$ is an increasing function of $z$, i.e. $\tau_{\text{stragg}}$ can be identified with a specific $\sigma(z)$. Based on this generalization, $O^{-1} \cdot O = 1$ is still satisfied:

$$\exp(0.25 \cdot \sigma(z)^2 \cdot d^2/dz^2) \cdot \exp(-0.25 \cdot \sigma(z)^2 \cdot d^2/dz^2) = \exp(0.25 \cdot \sigma(z)^2 \cdot d^2/dz^2 - 0.25 \cdot \sigma(z)^2 \cdot d^2/dz^2) = 1$$

(23)

Thus, the integral operator kernels $K$ and $K^{-1}$ remain unchanged, if the corresponding substitution $\sigma \rightarrow \sigma(z)$ is carried out. The quantity $\sigma(z)$ has to be bounded and positive definite. If we put $\sigma(z) = |\sin(k \cdot z)|$, then the kernel $K$ assumes the shape of a periodic
δ-kernel, if the sine becomes 0 and \( z \) satisfies \( z = n \cdot \pi / k \) (\( n = 0, 1, 2, \ldots, \infty \)). Therefore, we consider two types of Fourier series:

\[
\begin{align*}
\sigma (z) &= \sum_{n=0}^{\infty} A_n \cdot \sin(n \cdot k \cdot z) + C \tag{24} \\
\sigma (z) &= \sum_{n=0}^{\infty} A_n \cdot \sin(n \cdot k \cdot z) + B_n \cdot \cos(n \cdot k \cdot z). \tag{25}
\end{align*}
\]

In equation (24) \( C \) may be a constant to satisfy \( \sigma (z) \geq 0 \). The whole procedure is also applicable to the deconvolution kernel \( K^{-1}(\sigma (z), u-z) \), which removes the blurring of a source function by a Gaussian convolution. Applications to image processing (CT/CBCT) are challenging; we discuss a model example of a phantom in section 3.4.

For the sake of completeness, we have to give the 3D extension. For both the convolution and deconvolution procedures, we have to modify \( O^{-1} \) and \( O \) by the substitution \( d^2/dz^2 \rightarrow \Delta \) (Laplace operator). The convolution kernel \( K \) is now simply written in three dimensions, to yield

\[
K = (1/(\sqrt{\pi} \cdot \sigma (x, y, z)))^3 \cdot \exp\left[-((x-u)^2 + (y-v)^2 + (z-w)^2)/\sigma (x, y, z)^2\right]. \tag{26}
\]

Integrations have to be carried out over \( u, v \) and \( w \). The deconvolution procedure is similar; for this purpose, we write the corresponding Hermite-polynomial expansion in each dimension according to equation (3) by \( F_1(\sigma, x-u), F_2(\sigma, y-v) \) and \( F_3(\sigma, z-w) \), which have to be multiplied with the Gaussian kernel (26):

\[
\begin{align*}
F_1 &= \sum_{n=0}^{N} c_n(\sigma) \cdot H_{2n}\left(\frac{u-x}{\sigma}\right), \quad F_2 = \sum_{n=0}^{N} c_n(\sigma) \cdot H_{2n}\left(\frac{v-y}{\sigma}\right), \quad F_3 = \sum_{n=0}^{N} c_n(\sigma) \cdot H_{2n}\left(\frac{w-z}{\sigma}\right) \\
N \rightarrow \infty, \sigma &= \sigma (x, y, z)
\end{align*}
\]

\[
K^{-1} = \prod_{k=1}^{3} F_k \cdot K(\sigma, u-x, v-y, w-z). \tag{28}
\]

2.2. Extension to linear combinations with different Gaussian kernels

As already pointed out, there are many applications in which the restriction to one Gaussian convolution kernel is insufficient (e.g., the lateral scatter of protons). A further example is the scatter kernel of photons previously developed (Ulmer et al 2005). Since the whole scatter behavior can be described by small-angle scattering and, additionally, by middle- and long-range tails, it is not possible to account for the Compton scatter of photon beams in the MV region by one single Gaussian kernel. In the following, we restrict ourselves to a sum of three Gaussian kernels and one dimension; the extension to additional dimensions is straightforward (see equation (6)):

\[
K_s = \sum_{i=0}^{2} c_i \cdot K_i(\sigma_i, z-u). \tag{29}
\]

The subscripts of the Gaussian kernels \( K_i \) refer to the corresponding \( \sigma_i \) values, the assumption being that the three \( \sigma \) values are different. Without loss of generality, we may write: \( \sigma_0 < \sigma_1 < \sigma_2 \). Since each Gaussian kernel is normalized to 1, and \( K_i \) has also to be
normalized, the coefficients (weights) of the Gaussian kernels have to obey
\[
\sum_{i=0}^{2} c_i = 1. \tag{30}
\]
The kernel \( K \) results from a Green’s function of the operator \( O^{-1} \):
\[
O^{-1} = \sum_{j=0}^{2} c_j \cdot \exp(0.25 \cdot \sigma_j^2 \cdot d^2 / dz^2). \tag{31}
\]
The principal question arises how to construct \( K^{-1} \) from the given \( K \). We point out that \( K^{-1} \) is not the sum of the three inverse kernels \( K_0^{-1}, K_1^{-1} \) and \( K_2^{-1} \), as one might assume. Therefore, the inequality holds
\[
K^{-1} \neq \sum_{i=0}^{2} c_i \cdot K_i^{-1}. \tag{32}
\]
This may be easily seen with the help of the operator \( O \), which must satisfy \( O \cdot O^{-1} = 1 \). A further verification of relation (32) is obtained via equation (5), if the kernels \( K \) and \( K^{-1} \) are substituted by \( K \) and \( K^{-1} \). The integral operator relation (5) yielding the \( \delta \) function must hold for every integral operator. With the help of \( O \) and \( O^{-1} \) we obtain
\[
O \cdot O^{-1} = 1 = O \cdot \left( c_0 \cdot O_0^{-1} + c_1 \cdot O_1^{-1} + c_2 \cdot O_2^{-1} \right) \tag{33}
\]
It has to be noted that we have used these equations for \( O^{-1} \), which represent special versions of an operator calculus (Feynman 1962), in a modified form in a former investigation (Ulmer 2007). Let \( A \) and \( B \) be (in general, noncommutative) operators with \( AB - BA \neq 0 \); then this operator calculus provides an evaluation of the inverse operator \([A+B]^{-1}\):
\[
[A + B]^{-1} = A^{-1} - A^{-1} \cdot (B \cdot A^{-1}) + A^{-1} \cdot (B \cdot A^{-1}) \cdot (B \cdot A^{-1}) - A^{-1} \cdot (B \cdot A^{-1}) \cdot (B \cdot A^{-1}) + \cdots + (-1)^n \cdot A^{-1} \cdot (B \cdot A^{-1})^n \ (n = 0, 1, 2, \ldots, \infty). \tag{34}
\]
Since all operators under these considerations represent pure functions of differential operators (Lie series), which mutually commute, we can simplify the above expression:
\[
[A + B]^{-1} = \sum_{n=0}^{\infty} (-1)^n A^{-n-1} B^n. \tag{35}
\]
In the following, we perform the identifications
\[
\begin{align*}
A &= c_0 \cdot O_0^{-1} \\
B &= c_1 \cdot O_1^{-1} + c_2 \cdot O_2^{-1} \\
O_s &= [A + B]^{-1}
\end{align*} \tag{36}
\]
With regard to \( B \) we have to use the binominal theorem to evaluate the powers \( B^n \):
\[
B^n = (c_1 \cdot O_1^{-1} + c_2 \cdot O_2^{-1})^n = \sum_{j=0}^{n} \binom{n}{j} \cdot c_1^{n-j} \cdot c_2^j \cdot O_1^{-(n-j)} \cdot O_2^{-j}. \tag{37}
\]
Inserting this expression into equations (33) and (36) we obtain
\[
O_s = [A + B]^{-1} = \sum_{n=0}^{\infty} \frac{1}{c_0^{n+1}} \cdot (-1)^n \cdot \sum_{j=0}^{n} \binom{n}{j} \cdot c_1^{n-j} \cdot c_2^j \cdot O_1^{-(n-j)} \cdot O_2^{-j} \cdot O_0^{n+1}. \tag{38}
\]
Inverse Problems

In the standard case of a linear combination of two Gaussian kernels \[
O_s = [A + B]^{-1} = \sum_{n=0}^{\infty} \frac{1}{c_{0}^{n+1}} \cdot (-1)^{n} \cdot \sum_{j=0}^{n} \left( \frac{n}{j} \right) \cdot c_{1}^{n-j} \cdot c_{2}^{j} \cdot \exp \left\{ \frac{1}{4} \cdot \sigma_{j,n}^{2} \cdot d^{2}/dz^{2} \right\},
\]
where \[
\sigma_{j,n}^{2} = (n-j) \cdot \sigma_{1}^{2} + j \cdot \sigma_{2}^{2} - (n+1) \cdot \sigma_{0}^{2}.
\]
The translation of equations (39) and (40) to an expansion by kernels is not as easy as it might appear.

At first, we consider the case \( n = 0 \). This term assumes the shape \((1/c_{0}) \cdot \exp(-0.25 \cdot \sigma_{0}^{2} \cdot d^{2}/dz^{2})\), i.e. we obtain in every case the deconvolution kernel of the type (3) or (4): \((1/c_{0}) \cdot K_{0}^{-1}(\sigma_{0}, u - z)\). The succeeding term \( n = 1 \) is of principal interest. Recall that \( \sigma_{0} < \sigma_{1} < \sigma_{2} \). Then for \( j = 0 \), equations (39) and (40) yield the exponential operator \( \exp(0.25 \cdot (\sigma_{1}^{2} - 2 \cdot \sigma_{0}^{2}) d^{2}/dz^{2}) \). If \( \sigma_{1}^{2} - 2 \cdot \sigma_{0}^{2} > 0 \), the resulting kernel is a convolution kernel. The special case \( \sigma_{1}^{2} - 2 \cdot \sigma_{0}^{2} = 0 \) would imply a \( \delta \)-kernel and \( \sigma_{1}^{2} - 2 \cdot \sigma_{0}^{2} < 0 \) again a deconvolution kernel (equations (3) and (4)). The substitution \( \sigma \to \sigma_{0,1} \) provides the corresponding deconvolution parameter. It is also possible that for \( n > 1 \) and \( j = 0 \), we either obtain a deconvolution or a \( \delta \)-kernel. In general, we may assume that for \( n \leq n_{0} \) the parameter \( \sigma_{j,n}^{2} \) is negative and equation (39) leads to the type \( \exp[-0.25 \cdot \sigma_{j,n}^{2} \cdot d^{2}/dz^{2}] \) and, by that, to deconvolution kernels. For \( n > n_{0} \) with \( \sigma_{j,n}^{2} > 0 \), we obtain convolution kernels according to the differential operator function \( \exp[0.25 \cdot \sigma_{j,n}^{2} \cdot d^{2}/dz^{2}] \). We denote these kinds of convolution/deconvolution kernel expansions by the subscripts \( j, n \) and brackets, i.e. \( K_{(j,n)} \) and \( K_{(j,n)}^{-1} \):

\[
K_{(j,n)}^{-1} = \frac{1}{c_{0}^{j}} \cdot K_{0}^{-1} + \sum_{n=1}^{M} (-1)^{n} \cdot c_{0}^{n-1} \cdot \sum_{j=0}^{n} \left( \frac{n}{j} \right) \cdot c_{1}^{n-j} \cdot c_{2}^{j} \cdot K_{(j,n)}(\sigma_{j,n}, u - z),
\]

\[
M \to \infty
\]
if \( \sigma_{j,n}^{2} < 0 \); \( K_{(j,n)}'(\sigma_{j,n}, u - z) = K_{(j,n)}^{-1}(\sigma_{j,n}, u - z) \)
if \( \sigma_{j,n}^{2} > 0 \); \( K_{(j,n)}'(\sigma_{j,n}, u - z) = K_{(j,n)}(\sigma_{j,n}, u - z) \)
if \( \sigma_{j,n}^{2} = 0 \); \( K_{(j,n)}'(\sigma_{j,n}, u - z) = \delta(u - z) \)

Equation (41) is rather complicated; therefore, some comments are noteworthy. The condition \( \sigma_{2} > \sigma_{1} > 2\sigma_{0} \) can always be satisfied in our applications (multiple Compton scatter of photons and multiple Molière scatter of protons), we call this condition the ‘standard case’. By that, the deconvolution kernel \( K_{s}^{-1} \) can be simplified:

\[
K_{s}^{-1} = \frac{1}{c_{0}^{j}} \cdot K_{0}^{-1}(\sigma_{0}, u - z) + \sum_{n=1}^{M} (-1)^{n} \cdot c_{0}^{n-1} \cdot \sum_{j=0}^{n} \left( \frac{n}{j} \right) \cdot c_{1}^{n-j} \cdot c_{2}^{j} \cdot K_{n}(\sigma_{1}, u - z).
\]

In the standard case of a linear combination of two Gaussian kernels \( K_{s} = c_{0}^{j} K_{0} + c_{1}^{j} K_{1} \), the deconvolution kernel \( K_{s}^{-1} \) assumes the shape:

\[
K_{s}^{-1} = \frac{1}{c_{0}^{j}} K_{0}^{-1} + \frac{1}{c_{0}^{j}} \sum_{n=1}^{M} \left( \frac{c_{1}^{j}}{c_{0}^{j}} \right)^{n} \cdot (-1)^{n} \cdot K(n \cdot \sigma_{1}^{2} - (n+1) \cdot \sigma_{0}^{2}, u - z).
\]

For \( j = 0 \), formula (43) results from formula (42). In practical applications, we have also to keep \( M \) finite (besides \( N \) according to equations (3) and (4)). If \( c_{0} \gg 0.5(c_{1} + c_{2} \ll 0.5) \), then
Figure 1. Three types of a ‘Mexican hat’. The area of the three types is normalized to 1. Type 1: deep valley; type 2: broad valley with moderate depth; type 3: flat valley.

Table 1. Parameters for the three cases of ‘Mexican hats’.

<table>
<thead>
<tr>
<th>Type</th>
<th>$c_0$</th>
<th>$c_1$</th>
<th>$\sigma_0$</th>
<th>$\sigma_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.24</td>
<td>-0.24</td>
<td>0.70</td>
<td>0.20</td>
</tr>
<tr>
<td>2</td>
<td>1.9</td>
<td>-0.90</td>
<td>1.5</td>
<td>0.8</td>
</tr>
<tr>
<td>3</td>
<td>1.5</td>
<td>-0.50</td>
<td>1.4</td>
<td>0.9</td>
</tr>
</tbody>
</table>

we obtain rapid convergence, and it is sufficient to restrict the calculation of the inverse kernel up to order $2 \leq M \leq 4$.

A further task we shall consider is the so-called Mexican hat. For this purpose, we need two Gaussian convolution kernels satisfying the conditions: $c_0 > 1$, $c_1 < 0$, $c_0 + c_1 = 1$ (normalization of the total kernel, $\sigma_0 > \sigma_1$):

$$K_{\text{MH}} = c_0 \cdot K_0(\sigma_0, u - z) + c_1 \cdot K_1(\sigma_1, u - z)$$

$$O^{-1}_{\text{MH}} = c_0 \cdot O^{-1}_0 + c_1 \cdot O^{-1}_1. \quad (44)$$

Figure 1 shows three examples; the parameters are displayed in table 1.

The parameters of the types 2 and 3 should be regarded; the comparison shows that small changes in the coefficients and the $\sigma$ values may lead to rather large changes in the shape of the valleys.

In order to obtain positive-definite values at the zero point of $K_{\text{MH}}$ (figure 1), the following condition has to be satisfied:

$$c_0/\sigma_0 + c_1/\sigma_1 > 0. \quad (45)$$

The deconvolution kernel $K_{\text{MH}}^{-1}$ of the ‘Mexican hat’ is readily determined via equation (41). We have to account for the property that $\sigma_0 > \sigma_1$ yielding $\sigma_{0,n}^2 = \sigma_n^2 = n \cdot \sigma_1^2 - (n+1) \cdot \sigma_0^2 < 0$ and, therefore, deconvolution kernels in every order of $n$:

$$K_{\text{MH}}^{-1} = \sum_{n=0}^{M} K^{-1}_{(n)}(\sigma_n, u - z)$$

$$\sigma_n^2 = (n + 1) \cdot \sigma_0^2 - n \cdot \sigma_1^2; \quad M \to \infty. \quad (46)$$
This equation implies that, for each \( n \), we have to determine the inverse kernel with the parameter \( \sigma_n \), i.e. we obtain a sequence of inverse kernels \( K^{-1}_{(n)} \) according to equation (46). For practical applications, it is sufficient to keep the order of terms up to \( M = 4 \), if \( c_0 \gg c_1 \). The convergence would be slow and some higher order terms would be required, i.e. if (for instance) \( c_0 = 5 \) and \( c_1 = -4 \) \( (c_0 + c_1 = 1 \) must hold). With regard to specific applications to proton dosimetry, we also may consider a shift operator and its removal. Let us assume that \( K_{\text{MH}} \) is unshifted and \( K^{\text{-1}}_{\text{MH}} \) is a shifted inverse kernel. Then we have to use the translation operator to create the shift:

\[
\exp(\pm \xi \cdot \frac{d}{dz}) \cdot K^{\text{-1}}_{\text{MH}}(u - z) = K^{\text{-1}}_{\text{MH}}(u - (z \pm \xi)).
\] (47)

The shift \( \xi \) is removed by the inverse operation, i.e.

\[
\exp(\mp \xi \cdot \frac{d}{dz}) \cdot K^{\text{-1}}_{\text{MH}}(u - (z \pm \xi)) = K^{\text{-1}}_{\text{MH}}(u - z).
\] (48)

With regard to shifts of these two kernels we should note that \( \sigma_0 \) and \( \sigma_1 \) have to be kept constant; a \( z \)-dependence of them is not possible to maintain equation (47) or equation (48).

We shall consider the application of the approach to specific issues (the calculation of pristine Bragg curves, comparison with measurements obtained by ionization chambers and diode detectors, etc) in proton therapy. It must be emphasized that the application of the approach to proton therapy should not be seen as a restriction on its generality.

2.3. Monte Carlo calculations with the code EGSnrc

The EGSnrc code (Kawrakow and Rogers 2000) has been most widely applied to various problems of medical radiation physics. We have performed Monte Carlo calculations using EGSnrc with regard to two problems:

(1) \( 18 \) MV photon profiles in section 3.3. Absorption curve (field size: \( 20 \times 20 \) cm\(^2\)) and attenuation curves of the field sizes \( 1 \times 1, 2 \times 2 \) and \( 3 \times 3 \) cm\(^2\) and profiles (parallel beams). Previous results have been used (figure 2 in the publication of Ulmer et al 2005) with regard to the energy spectrum of 18 MV (central ray).
Inverse Problems 26 (2010) 085002

W Ulmer

(2) Image processing treated in section 3.4. Determination of the energy spectrum of 100 kV and 125 kV photons of CT/CBCT and their absorption/scatter behavior in some media of relevance (e.g. water-equivalent material and phantoms with different material densities (lung, bone)). The main purpose was the connection between Hounsfield values and the scatter parameters required for the kernel $K_s$ via Monte Carlo:

$$K_s = c_0 \cdot K_0(\sigma_0, u - x, v - y) + c_1 \cdot K_1(\sigma_1, u - x, v - y).$$

In general, the scatter parameters $\sigma_0$ and $\sigma_1$ depend on $z$, and this is the way to treat the depth-dependent scatter of a pencil beam. There is a correspondence between the Hounsfield value and electron density $\rho$ of a medium; the $\rho$-scaling of $\sigma_0$ and $\sigma_1$ can be carried out, if the scatter parameters are known for water, i.e. $\sigma_k(\text{medium}) = \sigma_k(\text{water}) \cdot \rho_{\text{medium}} / \rho_{\text{water}} (k = 0, 1)$. A rather circumstantial way for the elimination of scatter in CT/CBCT images would be the convolution/deconvolution of photon pencil beams in a similar fashion as usual in radiation therapy planning (Ulmer et al. 2005). The extension of the kernel $K_s$ to three dimensions and the inclusion of complete Fourier expansions according to equations (24)–(28) yield a 3D deconvolution of the complete volume based on the local electron density. This means that both $\sigma_0$ and $\sigma_1$ depend on $x, y, z$; they are calibrated with the reference value for water. The results of this procedure will be discussed in section 3.4.

3. Results

3.1. Convolution/deconvolution problems of two model cases with possible pitfalls

In order to test the derived deconvolution formulas of linear combinations of Gaussian kernels, we have carried out corresponding convolutions of some model cases and back calculations (deconvolution). Since the formulas for deconvolutions represent order-by-order calculations, a main purpose of the checks was to elucidate the required order/precision to obtain the source (origin) in a satisfactory way. With regard to the used order $N$ of Hermite polynomials in a figure we shall also state $N$ besides $c_0$ and $\sigma_0$, and in connection with $c_1, c_2, \sigma_1$ and $\sigma_2$ the order of the iterated kernel (equations (41)–(43) and (46)) will be given by $M$. It is evident that these assumptions do not provide a rigid scheme, since the convergence and, by that, the precision of the inverse solution depend on the magnitude of the used parameters, e.g. the ratio $c_1 / c_0$.

The principal problem of deconvolutions and their possible pitfalls becomes apparent by figures 2 and 3, which show that rather different sources (three boxes adjacent or not) with different $\sigma$ values lead to similar image functions. We can conclude that the behavior of the $\sigma$ values has to be known in a rather accurate fashion to prevent artifacts by the deconvolution procedures. Only due to the accurate knowledge of the subjected convolution parameters it is possible that the inverse procedure reliably works with high precision. By appropriate rescaling of the geometry of figures 2 and 3, the results have some importance in IMRT tasks, where the application of deconvolutions is increasing (see also the quoted references, parameters see table 2).

We have also carried out various calculations with modified properties of the presented examples, but there is no principal difference in the results. The main motivation to test the deconvolution formulas could be confirmed. Further examples will be presented in appendix A.1.

3.2. Applications to problems of proton dosimetry

In this section, we present an analysis of a pristine Bragg peak measured by different ways and a transverse profile of a broad beam. In the latter case, the different scatter behavior of
primary protons and secondary protons emerging from nuclear reactions imply that the use of one single Gaussian for the scatter treatment of both primary and secondary protons (Highland approximation) is rather a crude approximation.

3.2.1. Measurements of a pristine proton Bragg peak by an ionization chamber and a diode detector and deconvolution by a ‘Mexican hat’. A dose measurement of a Bragg curve by either (a) an ionization chamber or (b) a diode detector should yield identical values for the dose in the domain of the initial plateau, if both curves have been appropriately calibrated. This is not a trivial situation, since different processes are involved in the measurements of cases (a) and (b) mentioned above: in case (a), ionization processes are recorded, whereas case (b) involves the creation of excited electron states in the diode and their transitions. For case (b), in the domain of the Bragg peak and, in particular at the distal end of a Bragg curve, shell transitions and the Barkas effect play a significant role; this does not affect the measurements conducted with an ionization chamber. In the domain from the Bragg peak down to the distal end, protons of very low energy may be absorbed in the wall of the ionization chamber, avoiding the production of a signal. Due to the 1 MeV cutoff, the Monte Carlo calculations lead to reduced ranges in the steep-decrease branch, at the distal end. Therefore, one may pose the question, in which way the real pristine Bragg peak including the distal end can be measured. Since the ionization chamber can only record ionization processes within the (air) gap between the walls, the appropriate calibration of the detector in one domain (e.g., in the plateau of the Bragg curve) does not necessarily imply the correctness of the calibration in another domain (e.g., in the Bragg peak). The ionization chamber is also characterized by a larger volume and reduced transfer energy $E_{\text{transfer}}$ to the medium of the gap (usually air), which leads to a ‘larger’ Gaussian $\sigma$. In contrast to this, the diode detector has a smaller volume
and $E_{\text{transfer}}$ is increased, which implies a ‘smaller’ Gaussian $\sigma$. The difference between these two Gaussian kernels is a ‘Mexican-hat’ kernel. The absorption of a proton in the wall of a chamber, escaping detection, leads to an additional shift. Since the analytical calculations with a convolution/deconvolution of a Mexican hat (formulas (42)–(49)) are easy to perform with model M2 (Ulmer 2007, Ulmer and Matsinos 2010), we have only considered this case. Figure 4 shows a pristine Bragg curve (Harvard cyclotron: 160 MeV; the nozzle leads to a shift of 1.41 cm), obtained by an ionization chamber, and its analytical fit ($\tau_{\text{in}} = 0.204$ cm). The related polychromatic pristine Bragg curve $D_{\text{poly}}$ results from a Gaussian convolution subjected to a monoenergetic curve $D_{\text{mono}}$ with the rms value $\tau$:

$$D_{\text{poly}}(z) = \int D_{\text{mono}}(u) \cdot K(\tau, u - z) \, du,$$

$$\tau^2 = \tau_{\text{straggle}}^2 + \tau_{\text{in}}^2. \tag{50}$$

The analytical curve in figure 4 is the base for a deconvolution by a shifted ‘Mexican-hat’ kernel. The result is readily compared with the measurements obtained via a diode detector in figure 5.

The deconvolution kernel is constructed on the basis according to parameters in table 3 with $\xi = 0.03$ cm. Figure 5 reveals a potential pitfall in proton dosimetry. If both measurement curves are normalized to 1 at the position of the Bragg peak, the diode–detector curve would lead to a lower initial plateau.

<table>
<thead>
<tr>
<th>Figure</th>
<th>$c_0$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$\sigma_0$ (cm)</th>
<th>$\sigma_1$ (cm)</th>
<th>$\sigma_2$ (cm)</th>
<th>$N$</th>
<th>$M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/5</td>
<td>1.06</td>
<td>-0.06</td>
<td>-</td>
<td>0.104</td>
<td>0.06</td>
<td>-</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6/7 Highland</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>0.506</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>6/7 Molière $S_{pp}$</td>
<td>0.959</td>
<td>0.041</td>
<td>-</td>
<td>0.407</td>
<td>0.824</td>
<td>-</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>6/7 $S_{sp}$</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>0.728</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>-</td>
</tr>
</tbody>
</table>

**Figure 4.** The pristine Bragg peak of 160-MeV protons (Harvard cyclotron). Solid curve: analytical fit, discrete points: measured data (ionization chamber).
3.2.2. **Analysis of a transverse profile of a broad proton beam (Highland approximation and Molière’s multiple scatter theory).** The treatment of transverse proton profiles (beamlets and broad beams) is connected with a serious problem. Thus, it is usual and convenient to restrict the whole problem to the Highland approximation. This approximation is a compromise and works with one single Gaussian, which is slightly broader than the inner part of a Gaussian resulting from Molière’s theory in order to partially account for the tail. In particular, Gottschalk et al. (1993) have pointed out that the Highland approximation is too crude with regard to the required precision in proton radiotherapy. A further aspect is the behavior of secondary protons resulting from nuclear reactions, which start with a larger scatter angle, and the $\beta^+$ decay of the heavy recoils. This decay is connected with the isotropic emission of long-range $\gamma$-quanta; they even amplify the large angle tail of scatter contributions neglected by Highland (Ulmer and Matsinos 2010). Figure 6 represents the transverse profile based on the Highland approximation (dashes, $\sigma = 0.506$ cm) and on a linear combination of three Gaussians in order to include the Molière theory (two Gaussians, see also appendix A.2) and one Gaussian for the reaction protons and the long-range tail emerging from $\beta^+$ decays (the tail can be verified for $x > 4$ cm). The measurement data (crosses x) clearly indicate the crudeness of the Highland approximation. Figure 7 shows the profiles after deconvolution. With regard to the calculation procedure the cutoff for Highland has been taken at $x = 4.5$ cm, whereas for the sophisticated case a cutoff at $x = 6$ cm has been assumed. The parameters of the three Gaussian cases are valid at position $z = 15.1$ cm. The coefficient $c_2 = 1$ appears to be astonishing. However, we have also to account for the statistical weight referring to the stopping power $S_{pp}(z)$ of primary proton and to the stopping power $S_{sp}(z)$ of secondary protons. Thus, for an analysis of an experimentally determined profile we have weight $c_0$ and $c_1$ with $S_{pp}$ and $c_2$ with $S_{sp}$, i.e. $c_0$ and $c_1$ have to be multiplied with $S_{pp}$ and $c_2$ has to be multiplied with $S_{sp}$ at an arbitrary $z$ (in the present case at $z = 15.1$ cm): $(c_0, c_1) \rightarrow (c_0, c_1) \cdot S_{pp}$, $c_2 \rightarrow c_2 \cdot S_{sp}$. The stopping power functions $S_{pp}$ and $S_{sp}$ are stated elsewhere (Ulmer and Matsinos 2010). A deconvolution procedure appears to be useful either for checking scatter models against measured transverse profiles or for the determination of scatter parameters by a seeking procedure.

The long-range tails resulting from the photons due to the $\beta^+$ decay of heavy recoils prevent to reach the ideal box, and a further long-range Gaussian kernel would have to be
Figure 6. Proton profile ($6 \times 6 \text{ cm}^2$ field, 230 MeV protons, IBA cyclotron, $z = 15.1 \text{ cm}$, Highland (dashes) and Molière (dots)). Crosses refer to measurement data.

Figure 7. Deconvolution of profiles in figure 6; one Gaussian (Highland).

added similar to the cases in section 3.3. A consequence of figure 7 is that the total area resulting from the deconvolution of the three Gaussian kernels is only about 1% smaller than that obtained by the Highland approximation. Thus, the error in absolute dosimetry is not very significant in the case of a broad beam. This situation is changed in scanning techniques of proton beamlets, where the error is additive with regard to the superposition of all beamlets (Kusano et al 2007, Pedroni et al 2005, Schaffner 2008, Tourovsky et al 2005).

3.3. Application of deconvolutions to small transverse photon (18 MV) profiles (removal of scatter)

Similar to protons, small field sizes of photons are also very sensitive for the scatter behavior. Since a large field size—we have used a $20 \times 20 \text{ cm}^2$ field—provides the absorption depth
dose and the decrease of the energy fluence can be regarded as a surrogate of an infinite photon beam propagating in an infinite water box, it appears interesting in many problems of modern radiotherapy (e.g. IMRT, stereotaxy) to consider very small photon beams (field size $\leq 3 \times 3$ cm$^2$). Linear accelerators are characterized by a phase space, where the energy spectrum decreases with increasing radius from the central and the determination of the energy spectrum can be better determined by the attenuation (inclusion of lateral scatter) of small beams, i.e. we need the absorption curve by removal of scatter via a deconvolution. The $z$-dependent ($z \leq 20$ cm) scatter parameters $\sigma_k (k = 0, 1, 2)$ and the energy fluence absorption curve $I(z)$ are given by

$$
\sigma_0 = 0.035 \cdot z; \quad \sigma_1 = 0.210 \cdot z; \quad \sigma_2 = 0.525 \cdot z
$$

$$
I(z) = I_0 \cdot \left[ A_0 \cdot \exp(-\mu_0 \cdot z) + A_1 \cdot \exp(-\mu_1 \cdot z) \right].
$$

Based on Monte Carlo calculations (Ulmer et al 2005), the following parameters have been applied: $A_0 = 0.748$, $\mu_0 = 0.01502$ cm$^{-1}$, $A_1 = 0.252$, $\mu_1 = 0.02204$ cm$^{-1}$, $c_0 = 0.605$, $c_1 = 0.246$, $c_2 = 0.149$. Figure 8 presents the ideal absorption curve of the energy fluence (20 $\times$ 20 cm$^2$ field) and the attenuation curves of the small fields. The transverse profiles ($z = 5, 10$ and $15$ cm) with/without scatter (deconvolution) are shown in figures 9–11. With reference to deconvolutions, it has to be pointed out that (a) the chosen integration domain has to be very large (cutoff, if the relative energy fluence obeys $\leq 0.002$) and (b) the calculation order of $n$ and $m$ has to be very high ($N = M = 20$) to obtain a field-independent altitude at each position $z$ (figure 12).

### 3.4. Applications of deconvolutions to image problems (removal of scatter)

The following example has been taken from a CT image of a test phantom (figure 13), comprising an inner cylinder with bone-equivalent material (diameter: 4 cm) and an outer one with water-equivalent material (total diameter: 16 cm). The incident photon beam (125 kV) is a beamlet with Gaussian $\sigma_0 = 0.5$ mm. In order to apply deconvolution procedures, it is first necessary to determine the scatter produced within the phantom. The energy spectrum of the incident photon has been determined by Monte Carlo calculations (a detailed discussion is beyond the scope of this work). Based on the scatter behavior, as a function of the depth $z$, we are able to approximate the depth-dependent scatter with two Gaussian kernels with

![Graph showing the decrease of energy fluence in water and attenuation curves of small field sizes.](Image)

**Figure 8.** Decrease of the energy fluence in water (solid line: $20 \times 20$ cm$^2$ field, absorption); dashes and dots are attenuation curves of small field sizes ($1 \times 1$–$3 \times 3$ cm$^2$) with inclusion of scatter.
**Figure 9.** Transverse profiles of the $1 \times 1$ cm$^2$ field at $z = 5, 10,$ and $15$ cm and related deconvolution (dashes: boxes with rounded corners).

**Figure 10.** Transverse profiles of the $2 \times 2$ cm$^2$ field at $z = 5, 10,$ and $15$ cm and related deconvolution (dashes: boxes with rounded corners).

**Figure 11.** Transverse profiles of the $3 \times 3$ cm$^2$ field at $z = 5, 10,$ and $15$ cm and related deconvolution (dashes: boxes with rounded corners).
$c_0 = 0.81$ and $c_1 = 0.19$. Each photon pencil beam reaches the detector plane with $\sigma_0 = 0.42$ cm and $\sigma_1 = 3.05$ cm.

Instead of the repeated calculation of all photon beamlets—for the reason of symmetry in the above case we can reduce the calculation to one pencil—a deconvolution with the help of the scaling transformation according to equation (25) appears to be much more successful. The proportionality between the density and the scatter functions $\sigma_0(x, y, z)$ and $\sigma_1(x, y, z)$ holds (the coefficients $c_0$ and $c_1$ remain unchanged). However, the end values of the scatter functions at the detector plane are not valid. The scaling transformation has to be corrected by the detector influence and the initial scatter of the photon beamlet at the impinging position

$$
\begin{align*}
\sigma_0(x, y, z) &= \sigma_{i0} \cdot \rho(x, y, z) \\
\sigma_1(x, y, z) &= \sigma_{i1} \cdot \rho(x, y, z) \\
\sigma_{i0} &= 0.324 \text{ cm}; \sigma_{i1} = 1.638 \text{ cm}
\end{align*}
$$

One should point out that the scaling factors $\sigma_{i0}$ and $\sigma_{i1}$ correspond to mean values between beam entrance and exit at the phantom, corrected by detector influences and initial scatter.
occurring in the head. An integration based on the voxel density with the inverse kernels can be performed. The complete deconvolution problem was carried out in some areas of interest (incident surface, center of the cylinder, the borders water $\rightarrow$ bone and bone $\rightarrow$ water). Figure 14 shows a profile of the extracted Hounsfield units (box, inclusion of scatter and deconvolution).
It is rather interesting to see that, in the transition domains, the image is blurred. The complete 2D calculation is presented in figures 15 (calculated source) and 16 (image); although the procedure is more complex (deconvolution of 2D problems), the result is worth noticing.

4. Conclusions

Following a previous publication (Ulmer and Kaissl 2003), we have derived a rigorous algorithm for the computation of the inverse kernel (deconvolution) of a Gaussian convolution. The presented algorithm avoids the application of arbitrary filter functions in connection with the Fourier transform (see e.g. Ming Fang et al 1994). In this paper, we have extended the deconvolution problem to linear combinations of Gaussian kernels and to a ‘Mexican hat’ kernel. The so-called pitfall cases in section 3.1 give a clear indication that we must have accurate knowledge about the image functions and the parameters involved, in order to be able to perform reliable deconvolutions. Section 2.1 basically provides an understanding of the deconvolution of a ‘Mexican hat’ with respect to the underlying different physical processes of the energy transfer in the Bragg peak region registered either by a diode detector or a flat ionization chamber. The presented example of a closed voxel deconvolution of an image phantom appears to be challenging, since the whole deconvolution calculation consists of a scaling procedure. Further research in this direction seems to be justified to avoid pitfalls and artifacts.

Acknowledgments

Many thanks go to Barbara Schaffner, ETH Zürich. She has made available measurements of proton depth dose curves at the Harvard Cyclotron with a diode detector and ionization chamber. Thanks also go to the KNNC (Korean National Cancer Center) for providing a transverse profile and to the reviewer of this journal for helpful comments.

Appendix

A.1. Some further examples

The purpose of figures 2 and 3 is mainly based on a test and verification of the deconvolution procedure, if the preceding source functions are subjected to convolutions with Gaussian kernels. Some possible applications are given by backcalculations of fluence profiles obtained by MLCs (Das et al 2008, Pappas et al 2006). For interested readers we present similar examples in this appendix (figures A1–A3). The parameters of these figures are stated in table A1. Thus, a comparison between figures A1 (triangle) and A2 (three adjoining boxes) is of interest, since a possible pitfall is apparent. The solid curve in figure A2 also shows a triangle-like shape, and one may anticipate a similar source function as in figure A1. Thus, only by the suitable choice of the \( \text{rms} \sigma \) value the deconvolution procedure leads to rather different source functions. The small roundness of the corners of figure A3 results from the finite values \( N \) and \( M \) at the deconvolution.

In all preceding cases of the convolution/deconvolution with three Gaussian kernels we have restricted ourselves to the standard case, where all coefficients of the three Gaussian kernels have a positive sign. With regard to deconvolutions this presumption can be abandoned, and we now consider the case with \( c_1 < 0 \). In physical applications to scatter problems, the only restriction holds that the resulting kernel \( K_s \) has always to be positively definite, i.e. in every case \( \sigma_2 > \sigma_1 \) must hold. Figure A4 presents such an example (the parameters are stated in
Figure A1. Convolution of a triangle (solids) with one Gaussian (dashes) and deconvolution (dots).

Figure A2. Convolution of three boxes (box length: 0.1 cm, space length between the boxes: 0.05 cm, height of the source functions: 1 (middle part) and 0.5 (at both sides).


<table>
<thead>
<tr>
<th>Figure</th>
<th>(c_0)</th>
<th>(c_1)</th>
<th>(c_2)</th>
<th>(\sigma_0) (cm)</th>
<th>(\sigma_1) (cm)</th>
<th>(\sigma_2) (cm)</th>
<th>(N)</th>
<th>(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>0.25</td>
<td>–</td>
<td>–</td>
<td>4</td>
<td>–</td>
</tr>
<tr>
<td>A2 solid / A3</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>0.10</td>
<td>–</td>
<td>–</td>
<td>8</td>
<td>–</td>
</tr>
<tr>
<td>A2 dots / A3</td>
<td>0.8</td>
<td>0.20</td>
<td>–</td>
<td>0.025</td>
<td>0.075</td>
<td>–</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>A4</td>
<td>0.9</td>
<td>–0.38</td>
<td>0.48</td>
<td>0.40</td>
<td>0.82</td>
<td>–</td>
<td>1.50</td>
<td>–</td>
</tr>
<tr>
<td>A5–A8</td>
<td>0.9</td>
<td>–0.38</td>
<td>0.48</td>
<td>0.40</td>
<td>0.82</td>
<td>1.50</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Table A1, which also provides information on some calculations with boxes as source functions). A possible application may be multiple scatter theory of Molière, where continuously bending scatter profiles may not always hold. It should be noted that the cases (A5–A8) based on \(c_1 < 0\) may be of advantage with respect to the convergence in deconvolution calculations, since a negative \(c_1\) value may also imply increasing of \(c_0\), and the powers \(c_0^m / c_0^{m+1}\) occurring in equation (42) become very small more rapidly. The convolution/deconvolution is not the procedure of a Mexican hat, because \(\sigma_0 < \sigma_1 < \sigma_2\).
The main difference between this consideration with the negative $c_1$ value and the standard case can be verified at the outbound branches of figures A6 and A7, which show different bending due to the negative $c_1$ value compared to the standard situation with positive values of all coefficients.

A.2. Multiple scatter of Molière and an approximation by two Gaussians

The lateral scatter of protons can be described in a similar fashion to Compton scatter of high energy photons. Usually, the spatial behavior of the multiple-scatter theory involves a Gaussian and Hermite polynomials $H_{2n}$ for the long-range correction:

$$K (r, \sigma_I(z)) = N \cdot \exp(-r^2/\sigma_I(z)^2) \cdot \left[ a_0 + a_2 \cdot H_2(r/\sigma_I(z)) + a_4 \cdot H_4(r/\sigma_I(z)) \right] + a_6 \cdot H_6(r/\sigma_I(z)) + 0 \text{(higher order)}$$

$$N = 1/\left(\sqrt{\pi} \cdot \sigma_I(z)\right); \quad r^2 = x^2 + y^2. \quad (A.1)$$

The parameters of the Hermite polynomial expansion of multiple-scatter theory are $a_0 = 0.932; a_2 = 0.041; a_4 = 0.019; a_6 = 0.008$. The Gaussian half-width is the same as assumed
Figure A5. Convolution and deconvolution of three boxes (box length: 1 cm, space between them: 1 cm).

Figure A6. See figure A5 (space between the boxes reduced to 0.5 cm).

for the inner part. Based on an optimization problem, the task is now to determine a linear combination of two Gaussians according to equation (A.2) with different half-widths. This way, we are able to go beyond the Highland approximation. It is now the task to represent the lateral scatter described by equation (A.1) by a linear combination of two Gaussians, i.e.

\[
\begin{align*}
\int \int [K(r, \sigma_1(z)) - c_0 \cdot K_0(r, \sigma_0(z)) - (1 - c_0) \cdot K_1(r, \sigma_1(z))]^2 \, dy \, dx &= \text{Minimum} \\
& \quad c_0 + c_1 = 1.
\end{align*}
\]

Thus, \( \sigma_0 \) accounts for the inner part and \( \sigma_1 \) for the long-range tail of the proton scatter. With the help of the optimized approximation of the Molière scatter by two Gaussian kernels, the
The solution procedure of equation (A.2) is a standard task in numerical mathematics.

References


Bethe H A 1953 Molière’s theory of multiple scattering *Phys. Rev.* **89** 1256

Crop F, Reynaert N, Pitomvils G, Paelinck L, De Wagner C, Vakaet L and Thierens H 2009 The influence of small field sizes, penumbra, spot sizes and measurement depth on perturbation factors for microionization chambers *Phys. Med. Biol.* **54** 2951

Feynman R P 1962 *Quantum Electrodynamics Lecture Notes and Reprints* (New York: Benjamin)
Feynman R P and Hibbs A R 1965 *Quantum Mechanics and Path Integrals* (New York: Mac Graw Hill)
Gröbner W and Knapp H 1967 Contributions to the method of Lie series (Bibliographisches Institut, Mannheim)
Kawrakow I and Rogers D W O 2000 The EGSnrc code system: Monte Carlo simulation of electron and photon transport *NRCC Report PIRS-701* NRC, Canada
Matsuura T and Saitoh S 2006 Analytical and numerical inversion formulas in the Gaussian convolution by using the Paley-Wiener spaces *Appl. Anal.* **85** 901
Primas H 1989 *Induced Nonlinear Time Evolution of Open Quantum Systems* (New York: Plenum)
Saitoh S 2004 Approximate real inversion formula of a Gaussian convolution *Appl. Anal.* **83** 727
Schaffner B 2008 Proton dose calculation based on in-air fluence measurements *Phys. Med. Biol.* **53** 1545
Tuminonen T 2004 Perfusion deconvolution via EM algorithm *PhD Thesis* University of Helsinki, Research Project
Ulmer W 2007 Theoretical aspects of energy-range relations, stopping power and energy straggling of protons *Rad. Phys. Chem.* **76** 1089
Ulmer W and Kaisil W 2003 The inverse problem of a Gaussian convolution and its application to the finite size of the measurement chambers/detectors in photon and proton dosimetry *Phys. Med. Biol.* **48** 707