The minimum norm method for the determination of the charge density from elastic electron scattering data

H.G. Miller[1,2], Yiharn Tzeng[1], G.D. Yen[1], N. Canosa[3], R. Rossignoli[3], A. Plastino[3]

[1] Institute of Physics, Academia Sinica, Nankang, Taipei, 11529, Taiwan
[2] Department of Physics, University of Pretoria, Pretoria 0002, South Africa
[3] Departamento de Física, Universidad Nacional de La Plata, C.C. 67, (1900) La Plata, Argentina

Abstract. Unphysical behavior in the QR algorithm based least squares determination of the expansion coefficients of the charge density obtained from limited information about the charge form factor occurs when the spread of the singular values in the matrix relating these quantities becomes too large. Setting the smallest singular values equal to zero in the singular value decomposition used in the minimum norm method yields a much more reasonable determination of the charge density. Increasing the size of the basis without increasing the range of the prior information about the charge form factor leads to ambiguities in the determination of the charge density. Numerical results in an analytic model are presented.

PACS: 21.10.Ft; 25.30.Bf

Elastic electron scattering provides a means of determining the charge density of a nucleus, $\rho(r)$, from the experimentally determined charge form factor, $F_c(q)$. Since the electron-nucleus interaction is relatively weak, multiple scattering effects can be neglected and the scattering process can be described in first order perturbation theory. The connection between the charge density and the cross section is well understood and in plane wave Born approximation $F_c(q)$ is just the Fourier transform of $\rho(r)$ which for the case of even-even nuclei, which we shall consider, is simply given by

$$F_c(q) = 4\pi \int_0^\infty dr\, r^2 j_0(qr)\, \rho(r) \tag{1}$$

where j_0 is the spherical Bessel function of zero order and q is the absolute value of the three momentum transfer. Given that the experimental measurements are performed over a limited range at a finite number of values of the momentum transfer, q, a unique determination of $\rho(r)$ is not possible since the resulting inverse problem is ill posed. The generally accepted procedure for determining

$\rho(r)$ is to expand it in a Fourier Bessel (FB) basis [1–3] and then determine the expansion coefficients from a least squares fit to the experimentally determined values of $F_c(q)$. One of the problems in the aforementioned procedure is that increasing the number of terms in the expansion generally leads to unphysical oscillations in the charge density in spite of the fact that the charge form factor is well reproduced at the experimentally determined values of q. These unphysical oscillations arise, as we shall show, if one or more of the singular values of the non-square matrix to be inverted becomes abnormally small. In this case the least squares problem does not have a unique solution [4]. Since the inversion in the least squares problem is generally accomplished by means of the QR algorithm [4], this fact is overlooked. If the minimum norm method, which makes use of a singular value decomposition, is used, a unique solution of the least squares problem always exists [4] and the unphysical oscillations can be eliminated.

In order to demonstrate this we make use of the following analytical model. For a charge density given by a symmetrized Fermi distribution [5]

$$\rho(r) = \alpha \frac{\cosh(R/d)}{\cosh(R/d) + \cosh(r/d)} \tag{2}$$

an analytical expression for the corresponding charge form factor can easily be obtained [6, 7]:

$$F_c(q) = -\frac{4\pi^2 \alpha d}{q} \frac{\cosh(R/d)}{\sinh(R/d)}$$

$$\times \left[\frac{R\cos(qR)}{\sinh(\pi qd)} - \frac{\pi d \sin(qR)\cosh(\pi qd)}{\sinh^2(\pi qd)} \right]. \tag{3}$$

Only two of the parameters α, R and d are independent since the charge density must fulfill the normalization condition

$$4\pi \int dr\, r^2\, \rho(r) = Z. \tag{4}$$

Since the charge density is a single-valued function defined in a finite domain it can be expanded in a basis of orthogonal functions. In the FB expansion [1–3] use is made of the following orthogonality relation between spherical Bessel functions in a finite domain

$$\int_0^{R_c} dr\, r^2 j_l(q_n r) j_l(q_m r) = \frac{R_c^3}{2}\, [j_{l+1}(q_n R_c)]^2\, \delta_{n,m} \tag{5}$$

where the q_n are defined such that

$$j_l(q_n R_c) = 0. \tag{6}$$

For an even-even nucleus this yields the following FB expansion of the charge density

$$\rho(r) = \begin{cases} \sum a_n j_0(q_n r) & r \le R_c \\ 0 & r > R_c \end{cases} \tag{7}$$

where, from Eq. (6),

$$q_n = \frac{n\pi}{R_c}. \tag{8}$$

If the form factor in Eq. (1) is known at q_n, the coefficients a_n can easily be obtained and are given by:

$$a_n = \frac{F_c(q_n)}{2\pi R_c^3 (j_1(q_n R_c))^2} \tag{9}$$

In general, however, the cross section is measured at momentum transfers different from q_n. Using the expansion of the charge density given in Eq. (7) the charge form factor is then given by

$$F_c(q) = \frac{4\pi R_c^2}{q} \sum_n a_n \frac{(-1)^n}{(qR_c)^2 - (n\pi)^2} \sin(qR_c). \tag{10}$$

Now, given that n measurements are made at momentum transfers $\mathbf{q}' = (q'_1, q'_2, \dots, q'_n)$, we wish to determine the m expansion coefficients $\mathbf{a} = (a_1, a_2, \dots, a_m)$. In this case Eq. (10) leads to the following system of equations

$$\mathbf{F}_c = A\mathbf{a} \tag{11}$$

where $\mathbf{F}_c \in R^n$ is a column vector with elements $F_c(\tilde{q})$ with $\tilde{q} \in \mathbf{q}'$ and the corresponding matrix $A \in R^{m+n}$. If $n \ge m$ and rank$(A) = m$, a unique solution $\mathbf{a}$ for which the L^2 norm $\|\mathbf{F}_c - A\mathbf{a}\|_2$ is minimized exists [4]. This is just the least squares problem and it is usually solved by inverting the square matrix $A^T A$ since

$$A^T \mathbf{F}_c = A^T A\, \mathbf{a} \tag{12}$$

where A^T is the transpose of A. Numerically the most efficient way to accomplish this is by means of the QR algorithm which transforms $A^T A$ into a right upper triangular matrix by means of a set of orthogonal transformations. The inverse of this transformed matrix is easily obtained by simple back substitutions.

However, if the rank$(A) < m$, there are many solutions $\mathbf{a} \in R^m$ for which $\|\mathbf{F}_c - A\mathbf{a}\|_2$ is minimized. A unique solution to this least squares problem may be obtained if the additional requirement that $\|\mathbf{a}\|_2$ be minimized is also included [4]. This is nothing more than the minimum norm solution of the least squares problem and is obtained from a singular value decomposition of A. In this case A is decomposed in the following manner

$$A = U^T \Sigma V \tag{13}$$

where

$$\Sigma = \begin{bmatrix} \hat{\Sigma} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sigma_1 & & & & 0 \\ & \sigma_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & \sigma_r \\ \hline & & 0 & & \end{bmatrix} \Big| \begin{matrix} 0 \\ 0 \end{matrix} \in R^{m+n} \tag{14}$$

where $U \in R^{m+n}$ and $V \in R^{m+n}$ are orthogonal and the square of the singular values σ_i^2 are the eigenvalues of $A^T A$ with $\sigma_1 \ge \sigma_2 \dots \sigma_r > 0$. Since this decomposition is numerically expensive, the QR algorithm is usually used to solve the least squares problem.

It should also be noted that the matrix $\tilde{A} = (A^T A)^{-1} A^T$ needed to determine a in Eq. (11) is nothing more than the Moore-Penrose inverse of A [8, 9] which often is denoted as the pseudoinverse. In general, for any n, m, the pseudoinverse of A is given in terms of the aforementioned singular value decomposition as

$$\tilde{A} = V^T \Sigma^{-1} U \tag{15}$$

where the non-zero diagonal elements of the matrix Σ^{-1} are σ_i^{-1}. For $m = n$ it corresponds to the inverse of A. Furthermore, a unique solution of Eq. (11) can be obtained with the minimum norm method for $n > m$ when the system of equations is overdetermined [4] as well as for undetermined inverse problems with $n < m$. Lastly, it is interesting to note that under certain conditions Maximum Entropy methods for solving undetermined inverse problems also lead to the minimum norm method [10–12].

If rank$(A) = m$, there is no benefit in obtaining the minimum norm solution of the least squares problem as it is the same as that obtained with the QR algorithm. However, generally one does not know the rank of A. The simplest way to determine the rank of A is to determine the singular values of A. The number of non-zero singular values determines the rank of the matrix. In practice, however, if one or more of the singular values become small, the rank of the matrix A is usually less than the dimension of the square matrix $A^T A$ and as we shall show unphysical behavior will occur in the least squares fit obtained by means of the QR algorithm. In this case the problem can be avoided by using the minimum norm method and setting the small singular values to zero in the singular value decomposition of A. Unfortunately there are no hard and fast rules to determine how small a singular value should be before it is set equal to zero. Furthermore, large gaps in the values of the singular values do not always occur.

We have performed a numerical calculation for ^{12}C in the aforementioned analytical model with $R_c = 8$ fm, $d = 0.626$ fm and $R = 1.1A^{1/3}$. As input we have taken the values of $F_c(q)$ determined at a set of nine momentum
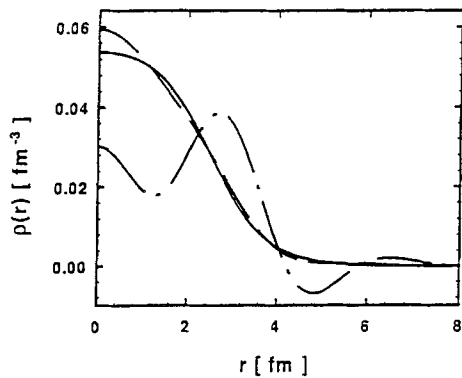
Fig. 1. Charge density as a function of r. The *solid curve* is the exact density, while the *dot-dashed curve* designates the least squares fit with the QR algorithm, and the *dashed curve* represents the result of the minimum norm method for 7 basis states
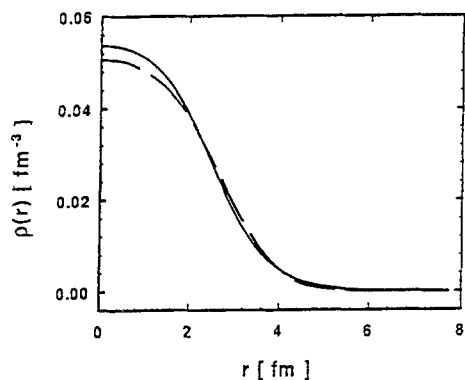

Fig. 2. Charge density as a function of r. The *solid curve* is the exact density, while the *dashed curve* represents the result of the minimum norm method for 30 basis states

Fig. 3. The FB expansion coefficients as a function of n. The circles connected by a *solid curve* are the exact values of the expansion coefficients while the triangles connected by the *short dashed curve* designates the values obtained from the least squares method with the QR algorithm and the squares connected by the *long dashed curve* represents the result obtained from the minimum norm method for 7 basis states

transfers $(0.001 \text{ fm}^{-1}, 0.5 \text{ fm}^{-1}, 1.0 \text{ fm}^{-1}, 2.0 \text{ fm}^{-1}, 3.0 \text{ fm}^{-1}, 4.0 \text{ fm}^{-1}, 5.0 \text{ fm}^{-1}, 6.0 \text{ fm}^{-1}, 7.0 \text{ fm}^{-1})$. If we require that the spread in the singular values of A be less than 10^5 the QR algorithm may be used to determine the set of a_m for $m \leq 6$ and the fits to the charge density look quite reasonable. For $m = 7$ (see Fig. 1) the smallest singular value $(.00307165)$ is less than 10^{-5} the value of the largest singular value (701.129) and we have set it equal to zero in the singular value decomposition of A. In this case the use of the QR algorithm results in large unphysical oscillations in the charge density while the minimum norm method provides a reasonable fit to the charge density. In both cases we reproduce the values of $F_c(q)$ at the nine given values of the momentum transfer.

Furthermore, with the same requirement on the spread of the singular values and for the same prior information about the charge form factor, a well-behaved charge density with $m = 30$ may also be obtained with the minimum norm method (see Fig. 2). Here again we reproduce the values of $F_c(q)$ at the nine given values of the momentum transfer. It is interesting to note that although the tail of the charge density agrees well with that obtained with $m = 7$ it does not agree well at smaller values
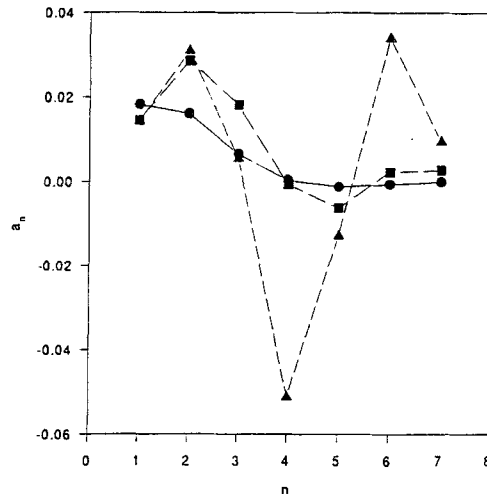
of r. This ambiguity arises from the fact that not enough information about the charge form factor has been provided to uniquely determine the charge density. The fact that no information has been given about the charge form factor for $q \geq 7 \text{ fm}^{-1}$ nicely demonstrates the fact that a unique determination of the charge density is not possible at smaller values of r.

In order to ascertain in a more quantitative manner the underlying reasons for the large unphysical oscillations in charge density obtained we compare the values of the expansion coefficients obtained from the least squares and minimum norm method with the exact values given by Eq. (9). The results for $m = 7$ are given in Fig. 3. Note that the coefficients for small values of n are not well determined because of the sparse amount of information initially provided about the form factor. The exact coefficients become very small very quickly and oscillate slowly about zero. The least squares method produces much larger oscillations which persist as m is increased while the minimum norm method provides a much smoother representation of the behavior of the coefficients. The deviations in the calculated expansion coefficients, which may be substantial, lead to deviations in ρ as can be seen in Fig. 2. Fortunately, and this is the nice thing about the choice of the FB basis, the coefficients with larger values of n are small and large deviations in these coefficients are not that important except at small values of r. If this were not the case the cut-off procedure used in the Fourier Bessel method would certainly not work as well as it does. Note also that setting the smallest singular value of A equal to zero is not equivalent to setting any of the values of the expansion coefficients equal to zero but rather a simple way of guaranteeing that no excessive structure is introduced into these coefficients. For this reason for larger values of n they are in better agreement
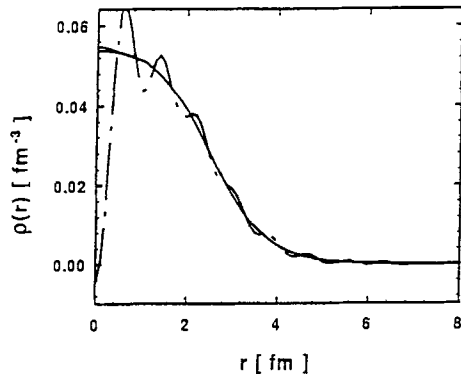
Fig 4. Charge density as a function of r. The *solid curve* is the exact density, while the *dot-dashed curve* designates the least squares fit with the QR algorithm, and the *dashed curve* represents the result of the minimum norm method for 20 basis states. The minimum norm results are distinguishable from the exact calculations only at small values of r

with the exact values of the expansion coefficients. This type of behavior persists if one increases m without increasing the amount of initial information.

In Fig. 4 we show the results for which there is a large amount of prior information as is usually the case when the charge form factor has been determined experimentally. The values of $F_c(q)$ are given at 140 equally spaced values for $0.001 \leq q \leq 7.001$ fm^{-1}. If no assumptions are made about $F_{ch}(q)$ for $q > 7$ fm^{-1} reasonable and identical results are obtained for $\rho(r)$ for $m \leq 19$ from both the least squares and the minimum norm method. Note that m may take on much larger values than the suggested cut-off for m ($\leq q_{max}R_c/\pi$) of Eq. (8) [1]. For $m \geq 20$ the spread in the singular becomes too large and oscillations occur in the least squares fit. In spite of the fact that no prior information is available for $q > 7$ fm^{-1} the minimum norm method provides a means of extrapolating to higher momentum which is consistent with the prior information and provides a reasonable determination of $\rho(r)$. The results are, of course, model dependent as the short ranged behavior of $\rho(r)$ depends on the choice of m as has been previously demonstrated.

If, as is usually the case, the experimental data, F_c, is assumed to have Gaussian errors then the uncertainties in the determination $\rho(r)$ may be determined in the manner suggested in [2]. In this case the variance in $\rho(r)$ due to the errors in the charge form factor, ΔF_c are given by

$$\Delta^2 \rho(r) = \sum_i (\Delta^2 F_c)_i \left(\frac{\partial \rho(r)}{\partial (F_c)_i} \right)^2$$

$$= \sum_i (\Delta^2 F_c)_i \left(\sum_j \frac{\partial \rho(r)}{\partial a_j} \frac{\partial a_j}{\partial (F_c)_i} \right)^2 \tag{16}$$

and derivatives can be determined from Eqs. (7) and (12) (in terms of pseudoinverse $\tilde{A}$ used to determine the expansion coefficients). The square root of $\Delta^2 \rho(r)$ corresponds to the width of a Gaussian distribution which describes the uncertainty in the determination of $\rho(r)$. Within the framework of Maximum Entropy methods a highly non-linear method for treating inverse problems involving data with Gaussian errors has been given by Gull and Daniell [13].

In the present note we have shown that unphysical oscillations of the FB expanded charge density determined from information of the charge form factor occur when the spread in the singular values of the matrix relating these quantities becomes too large. These oscillations may occur in any basis and are not the result of using the FB basis [7]. When they occur, the minimum norm method with the small singular values set equal to zero should be used to obtain the expansion coefficients. Increasing the size of the basis without increasing the range of the prior information about the charge form factor leads to ambiguities in the determination of the charge density.

References

1. J.L. Friar and J. W. Negele, Nucl. Phys. A **212**, 93 (1973)
2. B. Dreher et al., Nucl. Phys. A **235**, 219 (1974)
3. J. Heisenberg and H.P. Blok, Ann. Rev. Nucl. Part. Sc. **33**, 569 (1983)
4. D.S. Watkins, Fundamentals of Matrix Computations (Wiley, New York, 1991)
5. M.E. Grypeos, G.A. Lalazissis, S.E. Massen, and C.P. Panos, J. Phys. G **17**, 1093 (1991)
6. R.E. Kozak, Am. J. Phys. **59**, 74 (1991)
7. R. Anni, G. Co', and P. Pellegrino, preprint (1994)
8. A. Bjerhammaar, Theory of of Errors and Generaalized Matrix Inverses (Elsevier, Amsterdam, 1991)
9. C.R. Rao and S.K. Mitra, Generalized Inverse of Matrices and its Applications (Wiley, New York, 1971)
10. J. Baker-Jarvis, J. Math. Phys. **30**, 302 (1989)
11. J. Baker-Jarvis, M. Racine, and J. Alameddine, J. Math. Phys. **30**, 1459 (1989)
12. N. Canosa, H.G. Miller, A. Plastino and R. Rossignoli, Physica A**220**, 611 (1995)
13. S.F. Gull and G.J. Daniell, Nature **272**, 686 (1978)