Single and joint spin measurements with a Stern-Gerlach device

1993 J. Phys. A: Math. Gen. 26 2001

(http://iopscience.iop.org/0305-4470/26/8/021)

J. Phys. A: Math. Gen. 26 (1993) 2001–2010. Printed in the UK

Single and joint spin measurements with a Stern–Gerlach device

Hans Martens and Willem M de Muynck

Department of Theoretical Physics, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands

Received 9 July 1992

Abstract. The measurement of spin-$\frac{1}{2}$ observables using a Stern–Gerlach type device is studied. A magnetic field with a dominant dipole component leads to a measurement of spin in one direction. The measurement is shown to be feasible (in principle) for electrons as well as for neutral particles. A quadrupole field leads to a joint measurement of two incompatible spin observables. Again, both the electron and the neutral case are presented. The accuracy of the joint measurement is compared to the limit imposed by the uncertainty principle.

1. Introduction

The Stern–Gerlach device is one of the classic illustrations of quantum mechanical measurement theory [1, 2]. Because of the relatively simple nature of the observables involved, especially for the spin-$\frac{1}{2}$ case, it indeed forms an ideal testing ground for quantum measurement formalisms. Since in recent years a lot of work has been done on such formalisms, especially on those going beyond the standard von Neumann theory, the problem of spin-$\frac{1}{2}$ measurements in general, and in particular the Stern–Gerlach (SG), still arouse interest [3–6].

The standard SG set-up is sketched in figure 1. Particles move in the 1-direction through an inhomogeneous magnetic field, perpendicular to their direction of motion. The magnetic field, with an overwhelming dipole component, is modelled by a vector potential $A(q, t) = (q_2(a - bq_3), 0, 0)$ for $t \in (0, \tau)$, and zero elsewhere. Thus the Hamiltonian, for spin-$\frac{1}{2}$ particles with mass m, magnetic moment μ and charge Q, is given by

$$\hat{H} = \frac{1}{2m}([\hat{p}_1 - Q\hat{q}_2(a - b\hat{q}_3)]^2 + \hat{p}_2^2 + \hat{p}_3^2) + \tfrac{1}{2}\hbar\mu(b\hat{q}_2\hat{\sigma}_2 + (a - b\hat{q}_3)\hat{\sigma}_3) \qquad (1)$$

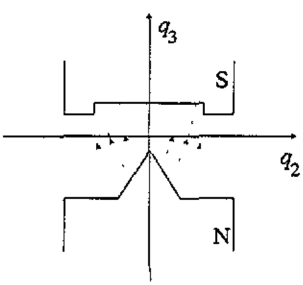

Figure 1. Schematic cross section of a Stern–Gerlach set-up. The inhomogeneous magnetic field is indicated by the dashed lines. The electrons move initially in the 1-direction, perpendicular to the plane of the drawing.

(carets denote operators; $\hat{\sigma}_1$, $\hat{\sigma}_2$ and $\hat{\sigma}_3$ denote the Pauli matrices, with eigenvalues ± 1). First we shall discuss neutral particles: $Q=0$. We choose our units such that $2m=b=\hbar=1$. Hence the Hamiltonian simplifies into

$$\hat{H}=\hat{p}_1^2+\hat{p}_2^2+\hat{p}_3^2+\tfrac{1}{2}\mu\hat{q}_2\hat{\sigma}_2+\tfrac{1}{2}\mu(a-\hat{q}_3)\hat{\sigma}_3. \tag{2}$$

The last two terms in equation (2) describe spin precession around an axis making an angle $\varphi=\tan^{-1}(q_2/(a-q_3))$ with the 3-direction. The dipole component of the magnetic field is usually taken very large, so that the point $(0,0,a)$ is far outside the beam area. Thus $\varphi \ll 1$, and $\hat{\sigma}_3$ is approximately conserved: we may neglect the $\hat{q}_2\hat{\sigma}_2$-term. Then, in the Heisenberg picture

$$\hat{p}_3(t)=\hat{p}_3(0)+\tfrac{1}{2}\mu\hat{\sigma}_3 t \tag{3}$$

whereas $\hat{p}_2$ and $\hat{p}_1$ are conserved.

In quantum measurement theory [1], originating with von Neumann, the object and the apparatus are initially independent. The object $\leftrightarrow$ apparatus interaction then brings about a correlation between the observable to be measured and the apparatus state. Finally some apparatus variable is read out (pointer observable), whose value gives information about the object observable that was to be measured. In the SG, the spatial variables function as the apparatus variables, whereas the spin observables represent the object variables. Therefore, given the way the SG is seen as a measuring instrument, it is natural to assume the particle state to be initially a product of the spatial part $|\xi\rangle$ and the spin part $|\psi\rangle$. The interaction, as (3) indicates, separates the particle beam into two sub-beams, the spin-up particles deflected upwards, the spin-down particles downwards. The measurement time is normally taken so large that the two sub-beams are fully separated [1]. The upper beam then contains only spin-up particles, the lower only spin-down particles. Thus reading out, e.g., $\hat{p}_3$ tells us the value of $\hat{\sigma}_3$. Blocking one of the beams leaves a beam of particles sharp with respect to $\hat{\sigma}_3$. As $\hat{\sigma}_3$ is conserved, this means that the measurement realizes a first kind measurement of $\hat{\sigma}_3$.

If a and/or τ are not large enough to satisfy the above conditions, the measurement is not perfect in the von Neumann sense. It is clear, however that from a pragmatic point of view, the SG still measures $\hat{\sigma}_3$ in some weaker sense. Accordingly, we will first discuss a formalism that can cope with such imperfect measurements, and then apply it to the SG case, using both neutral and charged particles. Finally, an SG-type device with quadrupole magnetic field is studied. This device realizes a joint measurement of incompatible spin observables. This latter set-up can, however, not reach perfection even asymptotically because of the limits imposed on joint measurements by the uncertainty principle.

2. Non-ideal measurements

In realistic situations, measurements are not described by self-adjoint operators, or projection-valued measures (PVMs), but rather by positive operator-valued measures (POVMs) [2, 7, 8]. The POVM notion forms an extension of von Neumann's axioms. For a discrete outcome set, a POVM $\{\hat{M}_k\}$ is a set of operators satisfying

$$\hat{M}_k \geqslant \hat{0} \qquad \sum_k \hat{M}_k=\hat{1}. \tag{4}$$

The operator $\hat{M}_k$ need not satisfy $\hat{M}_k = \hat{M}_k^2$, and is therefore not necessarily a projector. Moreover, in general $[\hat{M}_k, \hat{M}_l]_- \neq \hat{0}$ if $k \neq l$. The probability of outcome k is given by $\mathrm{Tr}(\hat{\rho}\hat{M}_k)$.

If we measure a POVM $\{\hat{M}_k\}$, this may be seen as a *non-ideal* measurement [9] of another POVM $\{\hat{N}_l\}$ if there exists a matrix (λ_{kl}) such that

$$\hat{M}_k = \sum_l \lambda_{kl}\hat{N}_l \qquad \lambda_{kl} \geqslant 0 \qquad \sum_k \lambda_{kl} = 1. \tag{5}$$

The matrix (λ_{kl}) is a stochastic matrix. We see that the $\{\hat{M}_k\}$-distribution $\mathrm{Tr}(\hat{\rho}\hat{M}_k)$ is a 'smeared' version of the $\{\hat{N}_l\}$-distribution. Even if the state is such that the latter is sharp, a measurement of $\{\hat{M}_k\}$ will in general not give one result with certainty. The non-ideality notion was introduced by Davies [7] and by Prugovečki [10]. It has been applied to several measurement schemes (see, e.g., [11] and references in [9, 12]). In the next section we shall see that the SG can also be treated by means of the non-ideality concept.

A major area of applications of (5) is formed by joint measurements of incompatible observables. Clearly, e.g. $\hat{\sigma}_2$ and $\hat{\sigma}_3$ are incompatible, so that they cannot be measured jointly. But non-ideal measurements of $\hat{\sigma}_2$ and $\hat{\sigma}_3$ may well be compatible. Consider a bivariate POVM $\{\hat{R}_{mn}\}$. If it satisfies

$$\begin{aligned}
&\sum_n \hat{R}_{mn} = \sum_k \lambda_{mk}\hat{M}_k \qquad \lambda_{mk} \geqslant 0 \qquad \sum_m \lambda_{mk} = 1 \\
&\sum_m \hat{R}_{mn} = \sum_l \mu_{nl}\hat{N}_l \qquad \mu_{nl} \geqslant 0 \qquad \sum_n \mu_{nl} = 1
\end{aligned} \tag{6}$$

we call a measurement of $\{\hat{R}_{mn}\}$ a *joint non-ideal measurement* of $\{\hat{M}_k\}$ and $\{\hat{N}_l\}$. Joint non-ideal measurements of several pairs of incompatible observables have been studied [12, 13] (cf also [14]). In particular, Prugovečki [10] and Busch [5] have discussed the spin-$\frac{1}{2}$ $\hat{\sigma}_2$, $\hat{\sigma}_3$ case. The joint measurement POVMs resulting from this theory have been connected to experiments for the analogous light polarization [15] and neutron interferometry [11] situations. Here we shall show that spin-$\frac{1}{2}$ joint measurements in the sense of (6) can also be realized by an SG set-up with a quadrupole magnetic field.

3. Dipole Stern–Gerlach

Consider once again the SG, described by (2), assuming that a is so large that the $\hat{q}_2\hat{\sigma}_2$ term may be neglected. As said above, after a correlation between $\hat{\sigma}_3$ and the apparatus, i.e. the spatial variables, is effected, the measurement is completed by reading out some apparatus observable [1, 2]. This read-out variable should if possible be chosen so as to optimize the measurement. In the following we shall, prompted by (3), take the momentum $\hat{p}_3$ to be that read-out variable. Qualitatively similar results, however, follow when position is used for this purpose, and sometimes this turns out to be preferable (namely, the electron quadrupole case, to be treated later on). As discussed above, we take the particle's state to be initially a product of the spatial part $|\xi\rangle$ and the spin part $|\psi\rangle$. After time τ, $\hat{p}_3$ has a probability distribution $\rho_{|\psi\rangle}(p_3, \tau)$, given by

$$\begin{aligned}
\rho_{|\psi\rangle}(p_3, \tau) &= \int_{-\infty}^{\infty} \mathrm{d}p_2\, \rho_{|\psi\rangle}(p_2, p_3; \tau) \\
&= |\langle\uparrow|\psi\rangle|^2 |\langle p_3 - \tfrac{1}{2}\mu\tau|\xi\rangle|^2 + |\langle\downarrow|\psi\rangle|^2 |\langle p_3 + \tfrac{1}{2}\mu\tau|\xi\rangle|^2
\end{aligned} \tag{7}$$

$$\rho_{|\psi\rangle}(p_2, p_3; \tau) = \mathrm{Tr}[|p_2, p_3\rangle\langle p_2, p_3| \exp(-\mathrm{i}\tau\hat{H})|\xi\rangle\langle\xi| \otimes |\psi\rangle\langle\psi| \exp(\mathrm{i}\tau\hat{H})] \tag{8}$$

($|\uparrow\rangle, |\downarrow\rangle$ denote $\hat{\sigma}_3$ eigenstates; $|p_3 \pm \frac{1}{2}\mu\tau\rangle$ is the eigenstate of $\hat{p}_3(0)$ with eigenvalue $p_3 \pm \frac{1}{2}\mu\tau$). The measurement's POVM on the spin-$\frac{1}{2}$ space $\mathbb{C}^2$ is uniquely determined by $\rho_{|\psi\rangle}(p_3, \tau) := \langle\psi|\hat{M}(p_3)|\psi\rangle$, so that

$$\hat{M}(p_3) = \lambda_+(p_3)|\uparrow\rangle\langle\uparrow| + \lambda_-(p_3)|\downarrow\rangle\langle\downarrow| \qquad \lambda_\pm(p_3) := |\langle p_3 \mp \tfrac{1}{2}\mu\tau|\xi\rangle|^2 \tag{9}$$

and $\lambda_-(p_3) \neq 0$ for $p_3 > 0$ in general. Since

$$\int_{-\infty}^{\infty} \lambda_\pm(p_3)\,\mathrm{d}p_3 = 1 \qquad \lambda_\pm(p_3) \geqslant 0 \tag{10}$$

the measurement is for any value of τ a non-ideal $\hat{\sigma}_3$ measurement in the sense of (5). At this point it is perhaps more convenient to divide the p_2, p_3-plane into two areas, $p_3 > 0$ and $p_3 < 0$, corresponding in the ideal case to spin-up and spin-down, respectively. The POVM then reduces to a two-element discrete ('yes–no') POVM $\{\hat{M}_k\}$ ($k = \pm$),

$$\hat{M}_\pm = \lambda_{\pm+}|\uparrow\rangle\langle\uparrow| + \lambda_{\pm-}|\downarrow\rangle\langle\downarrow| \tag{11a}$$

$$\begin{bmatrix} \lambda_{++} & \lambda_{+-} \\ \lambda_{-+} & \lambda_{--} \end{bmatrix} = \begin{bmatrix} \int_0^\infty \lambda_+(p_3)\,\mathrm{d}p_3 & \int_0^\infty \lambda_-(p_3)\,\mathrm{d}p_3 \\ \int_{-\infty}^0 \lambda_+(p_3)\,\mathrm{d}p_3 & \int_{-\infty}^0 \lambda_-(p_3)\,\mathrm{d}p_3 \end{bmatrix}. \tag{11b}$$

In a 2×2 non-ideality matrix, such as (11b) to which the functions $\lambda_\pm(p_3)$ have been reduced, the off-diagonal elements λ_{+-} and λ_{-+} correspond to the probabilities of wrong results. Thus we define

$$\kappa := \tfrac{1}{2}(\lambda_{+-} + \lambda_{-+}) \tag{12}$$

as a measure for the lack of quality of this SG measurement. In the worst possible case, realized at $\tau = 0$, the outcome does not depend on the input spin state at all (i.e. $\lambda_{ij} = \tilde{\lambda}_i$) so that $\kappa = \frac{1}{2}$. Ideally the off-diagonal elements vanish: $\lambda_{ij} = \delta_{ij}$ and $\kappa = 0$. Only when τ is much larger than the $\hat{p}_3$-width of the beam will this be the case. Only then the measurement is a first kind measurement, and the POVM reduces to a PVM. But the fact that the POVM or, equivalently, the measurement outcome probability distribution can be represented in the form of non-ideality (5), is independent of this condition on τ. Even when the beams are not fully separated, the final $\hat{p}_3$-distribution is unambiguously related to the initial $\hat{\sigma}_3$-distribution.

If we do not neglect the $\hat{q}_2\hat{\sigma}_2$-term in (2), equation (7) is no longer generally valid. But note that the Hamiltonian still has the symmetry property

$$[\hat{I}_2 \otimes \hat{\sigma}_3, \hat{H}]_- = \hat{0} \tag{13}$$

$\hat{I}_2$ denoting reflection of the position coordinates in the 1, 3-plane. Therefore, if the initial spatial state is 1, 3-reflection symmetric ($\hat{I}_2|\xi\rangle = |\xi\rangle$), it can be shown that

$$\rho_{|\psi\rangle}(p_2, p_3; \tau) = \rho_{\hat{\sigma}_3|\psi\rangle}(-p_2, p_3; \tau). \tag{14}$$

Combining (14) with (8) and (7), and with $\{\hat{M}(p_3)\,\mathrm{d}p_3\}$'s definition, we get

$$[\hat{M}(p_3), \hat{\sigma}_3]_- = \hat{0} \tag{15}$$

so that the measurement POVM can still be written in the form (5), with $\{\hat{N}_l\}$ a $\hat{\sigma}_3$-measurement. For initially reflection symmetric wavepackets, the measurement is a non-ideal spin measurement regardless of the values of a or τ.

The Schrödinger equation was integrated numerically†, using the Hamiltonian of (2). From the final state $\exp(-\mathrm{i}\tau\hat{H})|\xi\rangle\otimes|\psi\rangle$ thus obtained, the POVM and the corresponding measurement inaccuracy were calculated as outlined above. We took the initial spatial state $|\xi\rangle$ to be Gaussian in the 2- and 3-directions, and monochromatic in the 1-direction. Explicitly,

$$\langle p_1, p_2, p_3|\xi\rangle = c\delta(p_1 - p_{\mathrm{in}})\exp(-\tfrac{1}{4}p_2^2/v_2 - \tfrac{1}{4}p_3^2/v_3) \qquad c = (2\pi)^{-1/4}(v_2 v_3)^{-1/4}. \tag{16}$$

Thus 1-momentum is sharp, whereas in the other directions the variances are given by $\langle\Delta^2\hat{p}_2\rangle = v_2$ and $\langle\Delta^2\hat{p}_3\rangle = v_3$. Starting from the state (16), the Hamiltonian (2) gives for $a\to\infty$, when we may neglect the $\hat{q}_2\hat{\sigma}_2$-term,

$$\kappa_\infty(\tau) := \kappa(\tau)|_{a\to\infty} = \tfrac{1}{2} - \tfrac{1}{2}\operatorname{erf}(\tfrac{1}{2}\sqrt{2}\,\tfrac{1}{2}\mu\tau v_3^{-1/2}). \tag{17}$$

The value of κ, calculated for the measurement with initial state (16), is plotted in figure 2 for several values of a, together with the limit curve (17). Quality increases (κ decreases) with τ. Indeed, as expected, the limit value of κ achieved for large τ, decreases with a. Moreover, as a increases, the curves approximate (17) better and better. (The $a=0$ case will be treated in more detail in the next section.)

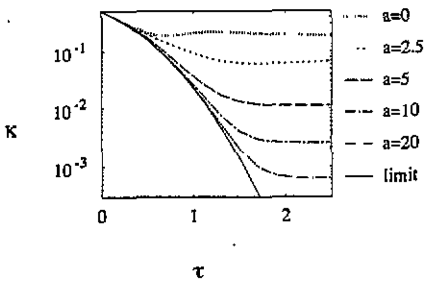

Figure 2. Numerical results for SG: spin measurement quality κ versus interaction time τ and field strength a. The strong field limit ($a\to\infty$) is also plotted (solid line). (Neutrals: $\mu=2$, initially $|\xi\rangle$ as in (16) with $v_2 = v_3 = \frac{1}{4}$.)

For electrons $Q=-e$, $m=m_e$ and $\mu = g\mu_{\mathrm{B}}$ (μ_{B} is the Bohr magneton). Now we choose units such that $\hbar = b = e = 2m_e = \frac{1}{2}g\mu_{\mathrm{B}} = 1$. Hence the Hamiltonian becomes

$$\hat{H} = [\hat{p}_1 + \hat{q}_2(a-\hat{q}_3)]^2 + \hat{p}_2^2 + \hat{p}_3^2 + \hat{q}_2\hat{\sigma}_2 + (a-\hat{q}_3)\hat{\sigma}_3. \tag{18}$$

Again, if a is taken sufficiently large, $\hat{q}_2\hat{\sigma}_2$ may be neglected and $\hat{\sigma}_3$ is conserved, so that the measurement is a non-ideal measurement of $\hat{\sigma}_3$ in the sense of (5). Thus the non-ideality formalism may be applied to electrons as well as to neutrals.

But the behaviour of electrons is also different from that of neutrals in certain respects. This can be seen from figure 3. Here a typical result of a numerical integration of the Schrödinger equation corresponding to the Hamiltonian (18) is plotted. Initial conditions are the same as for neutrals (apart from the values of the variances). But the wavepackets are not symmetrically deflected. This (numerical) result can be understood by means of the following heuristic argument. Consider first an electron in a homogeneous magnetic field of strength B in the 3-direction. Classically, it will move

† The integration of the Schrödinger equation was done using two methods. In the first method the position representation was used, with a discrete mesh. Convergence of the procedure was first order in time, and second order in position. In the second method we worked in the representation of the Landau states [cf (19)]. Both methods agreed to within numerical precision. The calculations were done on an IBM RS/6000 computer.

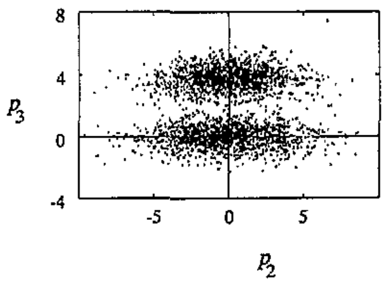

Figure 3. Typical output distribution for the electron SG set-up. The density of dots indicates the p_2, p_3-probability density. (Electrons: $a = 16$; $\tau = 1.9$; initially $|\psi\rangle = \frac{1}{2}\sqrt{2}\,(|\uparrow\rangle + |\downarrow\rangle)$, $|\xi\rangle$ as in (16) with $v_2 = 8$, $v_3 = \frac{1}{2}$ and $p_{\text{in}} = 0.03$.)

in spiral trajectories. Quantum mechanically, these correspond to the so-called Landau orbits [16]. The energy in the 1, 2-degrees of freedom is quantized into discrete levels

$$E_n = (2n+1)B \qquad n = 0, 1, 2, \ldots. \tag{19}$$

If the field inhomogeneities are sufficiently small with respect to the beam width (in the scaled units we use, $\langle \Delta^2 \hat{q}_2 \rangle \ll 1$), we may neglect any transition between different Landau levels (adiabatic approximation). Then only the drift of the levels due to the inhomogeneities in the 3-direction is to be taken into account: we approximate the magnetic field strength by $B = a - q_3$. Straightforward computation of the energy gradient then gives the force

$$\boldsymbol{F}_n = -\boldsymbol{\nabla} E_n \approx [(2n+1) + \sigma_3]\boldsymbol{e}_3 \tag{20}$$

where we have included the spin dependent force. Accordingly, $n = n_0 + 1$ spin-down electrons experience the same upward force as $n = n_0$ spin-up electrons. Therefore the initial spatial state should be constructed so as not to contain both $n = n_0$ and $n = n_0 + 1$ electrons. We may, for example, match the beam to the field strength in order to approximate the $n = 0$ Landau state. In the computations we therefore took the 2-variance of the Gaussian (16) to be $v_2 = \frac{1}{2}a$. Moreover, (20) shows that $n = 0$ spin-down electrons will experience a zero net force: they are not vertically deflected (in accord with figure 3). Therefore for the electron case the two integration areas in (11) are replaced by $p_3 > \tau$ and $p_3 < \tau$.

If we include the $\hat{q}_2\hat{\sigma}_2$ term, however, the symmetry property (13) does not hold any longer, unless $p_{\text{in}} = 0$. Consequently, for electrons the measurement cannot be written in the non-ideality form (5) for arbitrary a. But if a is large enough either to allow the neglect of $\hat{q}_2\hat{\sigma}_2$ or of the term $2p_{\text{in}}\hat{q}_2 a$ ($a^2 \gg p_{\text{in}}$), the POVM still approximately satisfies (5), taking $\{\hat{N}_l\}$ to be a $\hat{\sigma}_3$-measurement, and we may compute the measure (12). Results are shown in figure 4. For the conditions used there, the numerically computed final electron states are in accord with the symmetry (13) with an error of

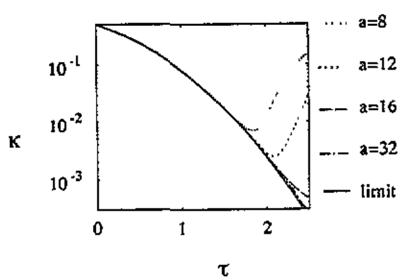

Figure 4. Numerical results for electron SG: spin measurement quality κ versus interaction time τ and field strength a. (Electrons: initially $|\xi\rangle$ as in (16) with $v_2 = \frac{1}{2}a$, $v_3 = \frac{1}{2}$ and $p_{\text{in}} = 0.03$.)

less than 1%. Note that the separation of the two integration areas at $p_3 = \tau$ only works until the spin-up packet comes in the $q_3 \approx a$ area. Then the wavepacket will stop moving. But the separation point will not stop, so that measurement quality will eventually decrease rather than approach a constant limit. This effect, visible in figure 4, is thus an artifact caused by the crude choice of integration areas.

The need for the approximation discussed in the beginning of the previous paragraph can, however, be overcome. We may, for example, add an electric field to compensate for the $2p_{\text{in}}\hat{q}_2$-term. Then the Hamiltonian would satisfy (13) exactly, and the measurement would be a non-ideal $\hat{\sigma}_3$-measurement, exactly satisfying (5) for $\hat{I}_2$-invariant wavepackets. Thus the restriction $a^2 \gg p_{\text{in}}$ would be removed.

We see therefore that spin measurements are possible for electrons, at least in principle. This contradicts a widely reproduced opinion of Bohr's [17, 18] (cf also [19]). Bohr argued, using qualitative arguments, that the variation of the Lorentz force across the beam width must be smaller than the spin-dependent force for the SG to work. This leads to the condition $\Delta q_2 \leqslant \lambda/2\pi$ (where Δq_2 is the beam's width in the 2-direction and λ is the de Broglie wavelength of the electrons). According to Bohr this condition implies that diffraction effects blur the beam, so that the SG would in effect be impossible for electrons ([17], p 335). But Bohr ignores that the Lorentz force causes the Landau quantization, so that the electrons are no longer 'free', strictly speaking. Thus the variation of the Lorentz force over the beam width, which is small, only causes some transitions between Landau levels, rather than a complete blurring of the beam, as is evidenced by the numerical results (figure 3).

4. Quadrupole field

If $a = 0$, the magnetic field has a quadrupole configuration. Obviously, for neutrals, the Hamiltonian (2) still satisfies the inversion symmetry (13). But additional symmetries arise. Denote rotation over an angle θ around the 1-axis for the space variables by $\hat{R}(\theta) = \exp(-\mathrm{i}\theta\hat{L}_1)$, and for the spin variables by $\hat{S}(\theta) = \exp(-\mathrm{i}\frac{1}{2}\theta\hat{\sigma}_1)$. Then we see that, for neutrals, the Hamiltonian remains unchanged if we rotate the spatial variables one way, and the spin variables the other:

$$[\hat{R}(\theta)\otimes\hat{S}(-\theta)]\hat{H}|_{a=0}[\hat{R}(\theta)\otimes\hat{S}(-\theta)]^{\dagger} = \hat{H}|_{a=0}. \tag{21}$$

Rewriting $(p_2, p_3) = (p\cos\theta, p\sin\theta)$, the probability density $\rho_{|\psi\rangle}(\theta, \tau)$ of finding the particle at angle θ in the p_2, p_3 plane at $t = \tau$ is

$$\rho_{|\psi\rangle}(\theta, \tau) := \langle\psi|\hat{M}(-\theta)|\psi\rangle = \int_0^{\infty} p\,\mathrm{d}p\,\rho_{|\psi\rangle}(p\cos\theta, p\sin\theta; \tau) \tag{22}$$

$\rho_{|\psi\rangle}(p_2, p_3; \tau)$ being given by (8). The minus sign in the definition of $\{\hat{M}(\theta)\,\mathrm{d}\theta\}$ means that we interpret the detection of a particle at angle $-\theta$ in the p_2, p_3-plane as measurement result θ. This compensates for the minus sign in the argument of $\hat{S}$ in (21). The rotational symmetry (21) of the Hamiltonian implies (see appendix) that, if the initial spatial part of the particle state is $\hat{R}(\theta)$-invariant, the POVM $\{\hat{M}(\theta)\,\mathrm{d}\theta\}$ can be parametrized as

$$\begin{aligned} \hat{M}(\theta) &= \int_0^{2\pi} \frac{\mathrm{d}\theta'}{2\pi}(\hat{1} + \hat{\sigma}_2\cos\theta' + \hat{\sigma}_3\sin\theta')f(\theta'-\theta) \\ f(\theta) &= \frac{1-\alpha}{2\pi} + \alpha\delta(\theta-\theta_0) \qquad 0 \leqslant \alpha \leqslant 1. \end{aligned} \tag{23}$$

Note that not only is $\{\hat{M}(\theta)\,\mathrm{d}\theta\}$ a POVM, but so is $\{(\hat{1}+\hat{\sigma}_2\cos\theta+\hat{\sigma}_3\sin\theta)\,\mathrm{d}\theta\}$. The latter was given by Helstrom [8] as a measurement of spin direction. Because

$$\int_0^{2\pi} f(\theta)\,\mathrm{d}\theta=1 \qquad f(\theta)\geqslant 0 \tag{24}$$

the quadrupole SG is a non-ideal measurement of Helstrom's spin direction observable. The smearing function f still depends on two parameters, α and θ_0. They characterize the measurement's quality and bias, respectively, and are determined by the specifics of the set-up. In particular the inversion symmetry (13) implies that $\theta_0=0$. The remaining parameter α depends on τ and the input state.

Alternatively, we can separate the p_2, p_3 plane into two half-planes, like we did in the previous section. Let the line $p_2\cos(\eta)-p_3\sin(\eta)=0$ divide the two areas. The probability is distributed over these two areas according to the POVM $\{\hat{N}_m(\eta)\}$ $(m=\pm)$, which is obtained from $\{\hat{M}(\theta)\,\mathrm{d}\theta\}$ by integration over a semi-circle (using $\theta_0=0$),

$$\begin{aligned}\hat{N}_+(\eta)&:=\int_{\eta-1/2\pi}^{\eta+\frac{1}{2}\pi}\hat{M}(\theta)\,\mathrm{d}\theta=\tfrac{1}{2}\hat{1}+\frac{\alpha}{\pi}(\hat{\sigma}_2\cos\eta+\hat{\sigma}_3\sin\eta)\\ \hat{N}_-(\eta)&:=\hat{1}-\hat{N}_+(\eta).\end{aligned} \tag{25}$$

Thus, the POVM $\{\hat{N}_m(\eta)\}$ is a non-ideal measurement of spin in direction η, with the analogue of the matrix (11) given by

$$\begin{bmatrix}\lambda_{++}(\eta) & \lambda_{+-}(\eta)\\ \lambda_{-+}(\eta) & \lambda_{--}(\eta)\end{bmatrix}=\begin{bmatrix}\frac{1}{2}+\alpha/\pi & \frac{1}{2}-\alpha/\pi\\ \frac{1}{2}-\alpha/\pi & \frac{1}{2}+\alpha/\pi\end{bmatrix}. \tag{26}$$

Accordingly, the measurement's quality is characterized by $\kappa(\eta)=\frac{1}{2}-\alpha/\pi$.

Now divide the p_2, p_3 plane into its four quadrants. The probability distribution over these four regions can be summarized in a bivariate POVM $\{\hat{R}_{mn}\}$, $m, n=\pm$. This POVM can be obtained by integrating $\hat{M}(\theta)$ over the respective quarter circles. The marginal POVMs of $\{\hat{R}_{mn}\}$ then correspond to a division of the p_2, p_3 into two half-planes by the $p_2=0$ and $p_3=0$ lines, respectively. It is easily seen that they are given by

$$\sum_n \hat{R}_{mn}=\hat{N}_m(0) \qquad \sum_m \hat{R}_{mn}=\hat{N}_n(\tfrac{1}{2}\pi) \qquad m,n=\pm. \tag{27}$$

Hence the bivariate POVM $\{\hat{R}_{mn}\}$ represents a joint non-ideal measurement of the incompatible observables $\hat{\sigma}_2$ and $\hat{\sigma}_3$ [5], in the sense of (6). Moreover, as η in (25) is arbitrary, the quadrupole SG may be considered to measure spin in *all* the directions in the 2, 3-plane jointly, in a non-ideal way.

The numerically calculated value for $\kappa(\eta)$ of (26) is plotted in figure 5 (dashed curve). We again took the initial spatial state to be Gaussian as in (16), but now $v_2=v_3$,

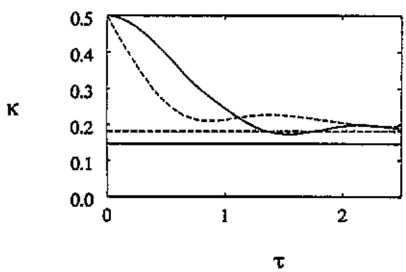

Figure 5. Numerical results for quadrupole SG: spin measurement quality κ versus interaction time τ, both for neutrals (dashed) and for electrons (solid). For electrons position is used as read-out variable. The fundamental quantum bound and the limit valid for spin direction measurements, (32) and (30) respectively, are indicated by horizontal lines. (Neutrals: $\mu=2$, initially $|\xi\rangle$ as in (16) with $v_2=v_3=\frac{1}{2}$; electrons: initially $|\xi\rangle$ as in (16) with $v_2=v_3=\frac{1}{2}$ and $p_{\rm in}=0.03$.)

so that the state is rotation invariant. The κ curve of figure 5 is, in contrast to those of figure 3, not monotonic. This is a consequence of the way in which the correlation between spin angle and momentum angle comes about, as sketched in the following qualitative argument. In the uncorrelated initial state, the spin part of the wavefunction can at each point (except the origin) $(p\cos\theta, p\sin\theta)$ be separated in a part with spin direction $-\theta$ on the one hand, and a part with direction $\pi-\theta$ on the other. Roughly speaking, the former propagates outward whereas the latter moves inward. As the particles with $\pi-\theta$ spins approach the origin from all sides, interference occurs. This causes the local minimum of the curve in figure 5. Eventually, a final state results that consists of a ring-like distribution of outward moving particles in the 2, 3-plane, in which the spin direction is $-\theta$.

If $a=0$, the Hamiltonian (18) for electrons reads

$$\hat{H}=[\hat{p}_1+\hat{q}_2\hat{q}_3]^2+\hat{p}_2^2+\hat{p}_3^2+\hat{q}_2\hat{\sigma}_2-\hat{q}_3\hat{\sigma}_3. \tag{28}$$

The full rotation symmetry (21) is out of the question. Nevertheless, the Hamiltonian remains invariant if we transform the spatial part according to $\hat{q}_2 \leftrightarrow \pm\hat{q}_3$, and the spin part according to $\hat{\sigma}_2 \leftrightarrow \mp\hat{\sigma}_3$. In other words: reflection symmetries analogous to (13) hold for both the $\theta=\frac{1}{4}\pi$ and $\theta=\frac{3}{4}\pi$ directions. Since, as shown in the dipole section, non-ideality (5) directly follows from the symmetry property (13) if the initial state also has that symmetry, we still have the possibility of realizing a joint measurement of two incompatible spin observables, namely $\hat{\sigma}_{\pi/4}$ and $\hat{\sigma}_{3\pi/4}$

$$\hat{\sigma}_{\pi/4}:=\tfrac{1}{2}\sqrt{2}\,(\hat{\sigma}_2+\hat{\sigma}_3) \qquad \hat{\sigma}_{3\pi/4}:=\tfrac{1}{2}\sqrt{2}\,(\hat{\sigma}_2-\hat{\sigma}_3). \tag{29}$$

In figure 5 the calculated value for κ in one of the two directions is plotted versus that of neutrals. For electrons it turns out to be better in this particular configuration to use $\hat{q}_2$ and $\hat{q}_3$ as read-out variables, rather than momentum. This is possible because the symmetry properties that we use [namely, the $\theta=\frac{1}{4}\pi$ and $\theta=\frac{3}{4}\pi$ reflection symmetries analogous to (13)] hold for position as well as for momentum. As in zeroth order $\hat{q}_3(t)=\hat{q}_3(0)+\hat{\sigma}_3 t^2$, the κ-curve starts quadratically, rather than linearly.

Note that, because $0\leqslant\alpha\leqslant 1$, the quality κ of the matrix (26) is limited by

$$\kappa(\eta)\geqslant\frac{1}{2}-\frac{1}{\pi}. \tag{30}$$

For neutrals the SG jointly measures spin in *all* directions in the 2, 3-plane, but for electrons, on the other hand, the measurement is a joint measurement of only *two* spin observables. It is not realized via a spin angle measurement, such as (23). Therefore only the uncertainty principle for joint measurements of two incompatible observables [5, 10–12] gives a limit for the latter case, i.e.

$$(\kappa_{\pi/4}-\tfrac{1}{2})^2+(\kappa_{3\pi/4}-\tfrac{1}{2})^2\leqslant\tfrac{1}{4}. \tag{31}$$

For a rotation symmetric initial state, such as the Gaussian (16) with $v_2=v_3$ which was employed in the quadrupole set-up calculations for both electrons and neutrals, $\kappa_{\pi/4}=\kappa_{3\pi/4}$. Then (31) implies

$$\kappa_{\pi/4}\geqslant\tfrac{1}{2}-\tfrac{1}{4}\sqrt{2}. \tag{32}$$

This limit is weaker than (30). Indeed, as figure 5 indicates, electrons may achieve better quality in this joint measurement than do neutrals.

Acknowledgments

The authors thank R Daling and J Cuypers for their computational assistance, and H Stoof for his careful reading of the manuscript. One of the authors (HM) further acknowledges the support of the Foundation for Philosophical Research (SWON), which is subsidized by the Netherlands Organization for Scientific Research (NWO).

Appendix

Symmetry (21), together with rotational invariance of the initial spatial state $|\xi\rangle$ and definition (22), implies

$$\hat{S}(\theta)\hat{M}(\theta')\hat{S}^{\dagger}(\theta)=\hat{M}(\theta'+\theta). \tag{A.1}$$

In particular,

$$\hat{M}(\theta)=\hat{S}(\theta)\hat{M}(0)\hat{S}^{\dagger}(\theta). \tag{A.2}$$

The operator $\hat{M}(0)$, like any positive operator (*effect*) on $\mathbb{C}^2$, can be parametrized as

$$\hat{M}(0)=\frac{\beta}{2\pi}\hat{1}+\frac{\alpha}{2\pi}(\hat{1}+\hat{\sigma}_1\sin\phi+\hat{\sigma}_2\cos\theta_0\cos\phi+\hat{\sigma}_3\sin\theta_0\cos\phi) \tag{A.3}$$

$(\alpha,\beta\geqslant 0)$. Integrating (A.2) over θ gives, using the POVM normalization (4),

$$\hat{1}=\int_0^{2\pi}\hat{M}(\theta)\,\mathrm{d}\theta=(\alpha+\beta)\hat{1}+\alpha\hat{\sigma}_1\sin\phi. \tag{A.4}$$

From (A.4) it follows that $\alpha+\beta=1$ and $\phi=0$, so that $\hat{M}(\theta)$ can be written in the form (23).

References

[1] Bohm D 1951 *Quantum Theory* (Englewood Cliffs, NJ: Prentice-Hall)
[2] Ludwig G 1976 *Einfuehrung in die Grundlagen der Theoretischen Physik* vol 3 (Braunschweig: Vieweg)
[3] Scully M, Englert B-G and Schwinger J 1989 *Phys. Rev.* A **40** 1775
[4] Schroeck F 1982 *Found. Phys.* **12** 479
[5] Busch P 1986 *Phys. Rev.* D **33** 2253
[6] Levine R and Tucci R 1990 *Found. Phys.* **19** 175
[7] Davies E 1976 *Quantum Theory of Open Systems* (London: Academic)
[8] Helstrom C 1976 *Quantum Detection and Estimation Theory* (London: Academic)
[9] Martens H and de Muynck W 1990 *Found. Phys.* **20** 255 (see also references therein)
[10] Prugovecki E 1977 *J. Phys. A: Math. Gen.* **10** 543
[11] de Muynck W and Martens H 1990 *Phys. Rev.* A **42** 5079
[12] Martens H and de Muynck W 1990 *Found. Phys.* **20** 357; 1991 *Phys. Lett.* **157A** 441
[13] Yuen H and Shapiro J 1980 *IEEE J. Inf. Th.* **IT-26** 78
[14] Arthurs E and Goodman M 1988 *Phys. Rev. Lett.* **60** 2447
[15] Mittelstaedt P, Prieur A and Schieder R 1987 *Found. Phys.* **17** 891
[16] Landau L and Lifshitz E 1977 *Quantum Mechanics* (*non-relativistic theory*) 3rd edn (Oxford: Pergamon)
[17] Bohr N 1985 *Collected Works* vol 6, ed J Kalckar (Amsterdam: North-Holland)
[18] Mott N and Massey H 1977 *The Theory of Atomic Collisions* 3rd edn (Oxford: Clarendon)
[19] Darrigol O 1984 *Historical Studies in the Physical Sciences* **15**, no 1 38