Found Phys
DOI 10.1007/s10701-007-9153-5

Afshar's Experiment Does Not Show a Violation of Complementarity

Ole Steuernagel

Received: 1 February 2007 / Accepted: 15 June 2007

Abstract A recent experiment performed by S. Afshar [first reported by M. Chown, New Sci. 183:30, 2004] is analyzed. It was claimed that this experiment could be interpreted as a demonstration of a violation of the principle of complementarity in quantum mechanics. Instead, it is shown here that it can be understood in terms of classical wave optics and the standard interpretation of quantum mechanics. Its performance is quantified and it is concluded that the experiment is suboptimal in the sense that it does not fully exhaust the limits imposed by quantum mechanics.

1 Introduction

Bohr's principle of complementarity of quantum mechanics characterizes the nature of a quantum system as being dualistic and mutually exclusive in its particle and wave aspects [1]. In the famous debates between Einstein and Bohr in the late 1920's [1] complementarity was contested but finally the argument has settled in its favor [1, 2]. Some forty years later Feynman stated in his 1963 lectures quite categorically that "*No one has ever found (or even thought of) a way around the uncertainty principle*" [3].

Some recent discussions centered on the question of whether the principle of complementarity is founded on the uncertainty principle [4–12]. Irrespective of whether one does [6–8, 10, 13, 14] or does not [4, 9, 12] subscribe to Feynman's point of view that complementarity and uncertainty are essentially the same thing [13, 14], there is agreement that the principle of complementarity is at the core of quantum mechanics.

Steps towards the quantification of the principle can be found in Refs. [15–18], they have been extended by a formulation using the visibility, V, of interference

O. Steuernagel (✉)
School of Physics, Astronomy and Mathematics, University of Hertfordshire, Hatfield, AL10 9AB, UK
e-mail: O.Steuernagel@herts.ac.uk

patterns to quantify the wave nature of quantum particles and the difference between their path detector state overlaps as a measure, $\mathcal{D}$, of their particle nature. Together, these give rise to an inequality

$$V^2 + \mathcal{D}^2 \leqq 1, \tag{1}$$

derived by Englert in Ref. [12]. Here, this inequality will be taken as the basis for a quantitative discussion of the principle of complementarity.

A recent laser experiment performed by S. Afshar has been presented as a possible counterexample to the principle of complementarity in quantum mechanics, in particular inequality (1) was supposedly violated [19–21]. Since some confusion has arisen in the context of the interpretation of Afshar's experiment [22–29][1] a wave-optical analysis (using plane waves) is given to explain the mechanism of the experiment and derive the scattering amplitudes for all partial waves involved. Subsequently, these amplitudes are used to quantify equation (1) and demonstrate that, far from violating the principle of complementarity, the Afshar experiment is less than optimal in the sense that it does not fully exhaust the limits of quantum mechanics prescribed by inequality (1).

Afshar's experimental setup, procedure, and interpretation are introduced in the next section. The wave optical analysis in Sect. 3 is subdivided into an informal description of the experiment in Sect. 3.1 and a quantitative analysis of the scattering amplitudes in Sect. 3.2 and their effects on the observed intensities. After that a quantification of the principle of complementarity is derived for experiments with a setup like that of Afshar's in Sect. 4. Then, in Sect. 5, the specific setup of Afshar's experiment, as presented in Ref. [21], is analyzed to prove our conclusion: Afshar's experiment does not 'violate quantum mechanics'.

2 Afshar's Experiment and His Interpretation

Stages of the experiment: in stage (i) of the experiment, see Fig. 1, the grid G is not present. In this stage only one slit is left open, therefore, only the corresponding detector gets illuminated. Opening of the other slit (and illuminating the other detector simultaneously) does not change the image at the first detector.

In stage (ii) the setup is modified by inserting the grid G directly in front of the lens L, as shown in Fig. 1. The grid is carefully positioned in such a way that the grid wires sit at the minima of the interference pattern (which can be checked by inserting a screen at the position of G and opening the second slit). It is then observed that if one of the two slits is blocked the image of the remaining open slit is strongly modified due to the presence of the grid. The wires of the grid reflect and diffract the light, this leads to a reduction of the intensity and introduces the formation of stripes in the image of the slit [19–21]. Note that the positioning of the grid at the minima of the interference pattern implies that it constitutes a carefully matched grating that reflects little of the passing light but will always *diffract* some of it.

[1] Several more newspaper articles, weblogs, and online articles are devoted to discussions of Afshar's experiment. See e.g. [35, 36].

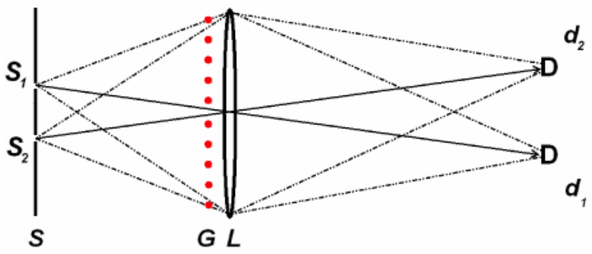


Fig. 1 (Color online) Setup of the 'Afshar experiment' [19–21]: a double-slit S is coherently and with equal intensity illuminated from the left. With the help of lens L the slits' images are recorded by detectors d_1 and d_2. At first, measurements are made *without* grid G (*dotted line*); subsequently, the grid is inserted and the measurements repeated. This sketch is a slightly simplified version of the original which contains an aperture around lens L and redirection mirrors in front of the detectors. Both simplifications are unimportant for the essentials of the experiment

In stage (iii) the other slit (previously blocked) is reopened. The crucial result at this stage is that the focal images of the two slits at detectors d_1 and d_2 look remarkably similar to those obtained in the absence of the grid in stage (i).

Afshar's interpretation, as cited from Ref. [19]:

> Laser light falls on two pinholes in an opaque screen. On the far side of the screen is a lens that takes the light coming through each of the pinholes (another opaque screen stops all other light hitting the lens) and refocuses the spreading beams onto a mirror that reflects each onto a separate photon detector. In this way, Afshar gets a record of the rate at which photons are coming through each pinhole. According to complementarity, that means there should be no evidence of an interference pattern. But there is, Afshar says.
>
> He doesn't look at the pattern directly, but has designed the experiment to test for its presence. He places a series of wires exactly where the dark fringes of the interference pattern ought to be. Then he closes one of the pinholes. This, of course, prevents any interference pattern from forming, and the light simply spreads out as it emerges from the single pinhole. A portion of the light will hit the metal wires, which scatter it in all directions, meaning less light will reach the photon detector corresponding to that pinhole.
>
> But Afshar claims that when he opens up the closed pinhole, the light intensity at each detector returns to its value before the wires were set in place. Why? Because the wires sit in the dark fringes of the interference pattern, no light hits them, and so none of the photons are scattered. That shows the interference pattern is there, says Afshar, which exposes the wave-like face of light. And yet he can also measure the intensity of light from each slit with a photon detector, so he can tell how many photons pass through each slit—the particle-like face is there too.
>
> "This flies in the face of complementarity, which says that knowledge of the interference pattern always destroys the which-way information and vice versa," says Afshar. "Something everyone believed and nobody questioned for 80 years appears to be wrong."

In other words, stage (iii) of the experiment is interpreted as a demonstration of simultaneous 'perfect' particle ($\mathcal{D} = 1$) and wave behavior ($V = 1$) thus violating the complementarity inequality (1) [19, 20]. Indeed, Ref. [21], erroneously (see Sect. 5 below), reports a value of $V^2 + \mathcal{D}^2 \geq 1.35$.

3 Wave Optical Analysis

Afshar's interpretation relies on the fact that the slits' images obtained in the presence of the grating, in stage (iii), are very similar to those obtained in stage (i) without a grating. In the article [19] they are describes in the figure caption of the slit images in stage (iii) of the experiment as returning to their 'original value', that of stage (i). But, close inspection of the images themselves reveals that in stage (iii) they display residues of the same disturbances that are seen in stage (ii), see picture in [19], Fig. 8 of [20], and Fig. 1 of [21]. These disturbances were considered negligible and without fundamental significance. Here, it is shown that these residual disturbances must not be neglected, they are key to understanding and correctly interpreting Afshar's experiment.

3.1 Qualitative Analysis

Putting the grid into the area where the dark fringes are to be expected in stage (iii) of Afshar's experiment is an elegant way of proving that some interference contrast is present, while quite efficiently avoiding the reflection or absorption of passing photons by the grid. But avoiding the reflection or absorption of photons is not enough to guarantee that their behavior in relation to the complementarity principle is unaffected. The grid's other main effect is the elastic scattering of passing (and reflected) photons due to *diffraction*. Even if the grid G was perfectly absorbing, and hence did not reflect any photons, would it still *diffract passing photons* on their way from slit S_2 to detector d_2. Some of those photons get scattered towards detector d_1; this is because matching the diffraction grating to the minima of the interference pattern implies that the direction of its first diffraction order points towards detector d_1, see Fig. 2(a) below. For symmetry reasons, analogous perturbations affect photons on their way from slit S_1 to d_1, redirecting some towards detector d_2, see Fig. 2(b).

This explains some features of Afshar's experiment. In stage (ii) of the experiment, reflection and absorption, and diffraction by the grid distort the slit images at the detectors. In stage (iii) the other slit is opened and first-order diffraction of photons from that newly opened slit apparently restores the slit images. It also shows qualitatively how the complementarity principle is at work: the path detection in the presence of the grid becomes less reliable since photons are diffracted towards the 'wrong' detector thus compromising path detection.

3.2 Quantitative Analysis of Diffraction

In the forthcoming analysis we introduce three simplifications that do not affect the underlying mechanism of the experiment.

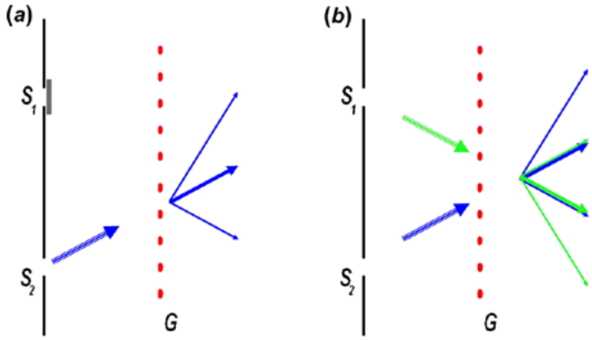


Fig. 2 (Color online) Geometry of wave vectors in the vicinity of the grid plane G: (**a**) only slit S_2 is open, (**b**) both slits are open. Some forward scattered light is deflected sideways by the grid (**a**). This loss of forward scattered light is partly compensated for by opening the other slit (**b**) which 'adds on' more light with the correct phase and transverse momentum. Here, only zeroth and first order transmitted light wave vectors (not to scale) are plotted

Firstly, we assume the light is monochromatic with angular frequency $\omega = ck$ and wave number k, and we only analyze features in the plane defined by the optical axis and the two slits S_1 and S_2, that is, we treat the problem in two spatial dimensions.

Secondly, we assume that the we can use the paraxial wave approximation, namely, we describe the light in terms of (scalar) plane waves and all angles are small. This approximation is suitable because the Afshar experiment allows for the clear separation of various diffraction orders n, see Figs. 1(c) and (d) of Ref. [21].

Thirdly, we model the grid by a planar reflecting film (e.g., a suitably etched mirror) consisting of flat strips of metal rather than an array of reflecting round wires. This last assumption would lead to a slightly better performance than Afshar's setup, which is using round wires, because reflection (although not diffraction) of photons towards the 'wrong' detector would become completely suppressed—instead, photons that hit the grating would be back-reflected. This modification would therefore not only enhance performance, it also simplifies our analysis since the two *mutually exclusive subensembles* of photons, transmitted and back-reflected, are cleanly separated.

A similar clean separation of these mutually exclusive subensembles of photons could be achieved by making the grating strips perfectly absorbing; of course in this case only the transmitted photons are any more available for simultaneous path measurements.

With these simplifying assumptions we can determine the interference pattern at the grid to be proportional to $\cos(k_\perp x)^2 = \cos(\frac{ks}{g}x)^2$, where k is the wave number of the light and its transverse component $k_\perp = ks/g$ arises from the geometry of the setup ('forward' direction along the z-axis, 'x' parameterizes transverse coordinate). The distance between the slits is $2s$ and g is the distance from the double-slit to the grid; small angles (paraxial beams) are assumed throughout. Consequently, the grid has spacing $\Lambda = 2\pi/(2k_\perp) = \pi g/(ks)$ and can be expanded into a discrete Fourier-series with periodicity Λ. Our model for the grid $G(x)$ is a periodic comb of reflecting stripes; in other words, the reflectivity alternates between values of unity in regions

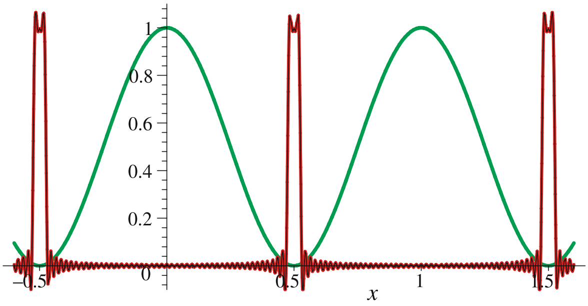


Fig. 3 (Color online) The grid function $G(x)$ of (2) for a grid with covering ration $a = 0.05$ described by a discrete Fourier series expanded up to 50th order together with the interference pattern intensity $\cos(k_\perp x)^2$ of the two slits S_1 and S_2 in the grid plane. The x-axis is scaled in terms of the grid spacing Λ

centered at odd multiples of $\Lambda/2$ over a distance Λa, $(a < 1)$ and zero over the remaining distance $\Lambda(1-a)$: a is the covering ratio of the grating. $G(x)$ is plotted, together with the interference pattern, in Fig. 3.

The grid's functional description in terms of a discrete Fourier series is given by

$$G(x) = c_0 + \sum_{n=1}^{\infty} c_n \cos\left(\frac{2\pi}{\Lambda} x n\right), \tag{2}$$

with

$$c_0 = a, \quad \text{and} \quad c_n = 2(-1)^n \frac{\sin(a\pi n)}{\pi n}. \tag{3}$$

We will now work out how this grating affects incoming plane waves. For simplicity we assume that the wave vector of the incident light ψ_{in} is perpendicular to the grating. We will later include the slight deviation due to the fact that the slits S_1 and S_2 lie half a diffraction order off the optical axis, compare Fig. 2. A plane wave traveling in the z-direction with wave vector $\vec{k} = k \cdot \hat{\mathbf{z}}$ is described by $\psi_{in} = e^{i(kz-\omega t)}$. When the grid diffracts this plane wave into the n-th positive or negative order its wave vector becomes $\vec{k}_{\pm n} = \pm k_{x,n} \cdot \hat{\mathbf{x}} + k_{z,n} \cdot \hat{\mathbf{z}}$ where $k_{x,n} = n \cdot 2k_\perp$ and $k_{z,n} = \sqrt{k^2 - k_{x,n}^2}$.

After interaction with the grating G we thus find the plane-wave modes $\psi_{\pm n,t} = e^{i(k_{z,n}z \pm k_{x,n}x - \omega t)}$ for light transmitted into orders $\pm n$, and $\psi_{\pm n,r} = e^{i(-k_{z,n}z \pm k_{x,n}x - \omega t)}$ for reflection. With the tacit understanding that the grating function (2) can be viewed as an operator $\hat{G}$ that imparts transverse momentum kicks of size $\hbar k_{x,n} = n \cdot \hbar 2k_\perp \doteq n p_\perp$, we thus introduce the corresponding momentum transfer operator $\hat{p}_\perp$ which allows us to write down the effect of the grating on an incoming wave as

$$\hat{G}(x) = c_0 \hat{1} + \sum_{n=1}^{\infty} c_n \frac{e^{i\frac{nx\hat{p}_\perp}{\hbar}} + e^{-i\frac{nx\hat{p}_\perp}{\hbar}}}{2} \tag{4}$$

$$= a\hat{1} + \sum_{n=1}^{\infty} \frac{(-1)^n \sin(a\pi n)}{\pi n} 2\cos\left(\frac{nx\hat{p}_\perp}{\hbar}\right). \tag{5}$$

Including the π phase jump associated with reflection [30] we therefore find for the reflection amplitudes

$$r_0 = -c_0 = -a \tag{6}$$

and

$$r_n = r_{-n} = \frac{-c_n}{2} = (-1)^{n+1}\frac{\sin(a\pi n)}{\pi n}. \tag{7}$$

Following the same logic [30], we find the transmission amplitudes obey

$$t_0 = 1 + r_0 \quad \text{and} \quad t_n = t_{-n} = r_n. \tag{8}$$

After a multiplication with $e^{i\omega t}$ to remove the time-dependence we thus arrive at the result that the reflected and transmitted partial waves have the form

$$\begin{aligned}[\psi_t] + [\psi_r] = &\left[(1-c_0)e^{ikz} - \sum_{n=1}^{\infty}\frac{c_n}{2}\cdot(\psi_{+n,t} + \psi_{-n,t})\right] \\ &+ \left[-c_0 e^{-ikz} - \sum_{n=1}^{\infty}\frac{c_n}{2}\cdot(\psi_{+n,r} + \psi_{-n,r})\right].\end{aligned} \tag{9}$$

With $\sum_{n=1}^{\infty} c_n^2 = 2(a-a^2)$ we can check the normalization and find $\sum_{n=-\infty}^{\infty}(r_n^2 + t_n^2) = r_0^2 + t_0^2 + \sum_{n=1}^{\infty} c_n^2 = a^2 + (1-a)^2 + 2(a-a^2) = 1$, as required.

After the discussion of the effect of the grating on a single plane wave $\psi_{in} = e^{i(kz-\omega t)}$ at normal incidence, we can extend the discussion by modeling the waves emanating from the two slits by a superposition of two plane waves as sketched in Fig. 2 above, namely

$$\psi_{in} = \frac{1}{\sqrt{2}}(\exp[i(k_{z,\frac{1}{2}}z + k_{x,\frac{1}{2}}x - \omega t)] + e^{i\phi}\exp[i(k_{z,\frac{1}{2}}z + k_{x,-\frac{1}{2}}x - \omega t)]). \tag{10}$$

Here, ϕ is the relative phase angle between the modes emanating from the two slits; $\phi = 0$ describes the Afshar experiment with the wires positioned at the interference minima (other cases correspond to nonzero values of ϕ, for maximal illumination of the wires $\phi = \pm\pi$).

The transverse components $k_{x,\pm\frac{1}{2}} = \pm\hbar k_\perp = \pm\Delta p_\perp/2$ of the incoming waves sketched in Fig. 2(b) offset the waves by half orders up or down. This is why the diffracted (and reflected) waves fall into half odd integer orders, compare e.g. Fig. 5(a). The stage (ii) reflection and transmission probabilities are given by $R^{(\text{ii})}_{n\pm 1/2} = |r_{n\pm 1/2}|^2$ and $T^{(\text{ii})}_{n\pm 1/2} = |t_{n\pm 1/2}|^2$; here, the shifts '$\pm 1/2$' label slits S_2 and S_1 respectively. Their effect shows up as a relative transverse displacement of the curves in Figs. 4 and 5 by one unit.

In stage (iii) the probabilities become considerably modified by the presence of the second wave in (10). For instance, the partial waves of order $n + 1/2$ originating from slit S_2 overlap with the adjacent order $n + 1/2 - 1$ partial waves originally emanating from slit S_1. The associated probabilities are therefore given by the coherent sums

$$R^{(\text{iii})}_{n+1/2}(\phi) = |r_{n-1/2} + e^{i\phi}\cdot r_{n+1/2}|^2 \tag{11}$$

and

$$T^{(\text{iii})}_{n+1/2}(\phi) = |t_{n-1/2} + e^{i\phi}\cdot t_{n+1/2}|^2. \tag{12}$$

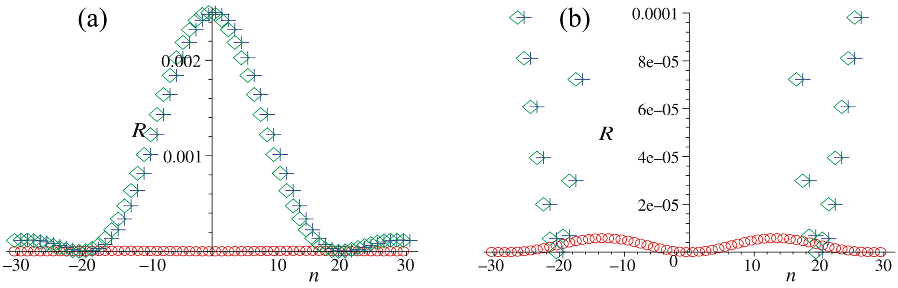


Fig. 4 (Color online) Reflection probabilities for diffraction up to 30th order n: for stage (ii) of the experiment green diamonds and blue crosses describe the reflection probabilities $R^{(ii)}_{n-1/2}$ and $R^{(ii)}_{n+1/2}$ for beams originating from slit S_1 and S_2 respectively. When both slits are opened simultaneously in stage (iii) (red circles) the reflection is very much suppressed, plot (**a**), but the reflection probabilities $R^{(iii)}_{n+1/2}(\phi=0)$ are not zero, plot (**b**). Covering ratio of the grid is $a=0.05$

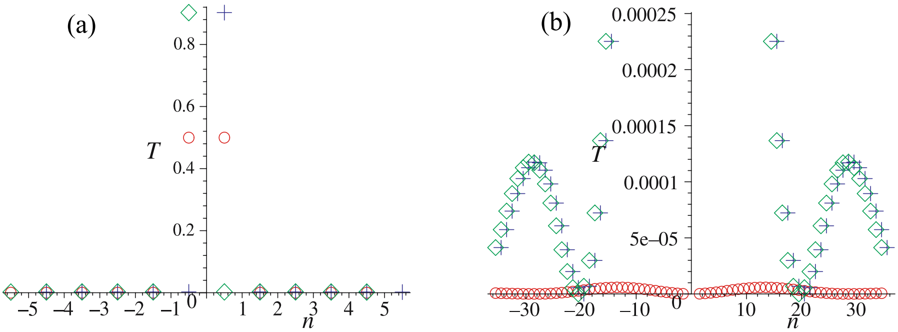


Fig. 5 (Color online) Transmission probabilities as in Fig. 4 above. In stage (ii) of the experiment grid reflection and scattering reduces the probability of a photon to arrive at the 'correct' path detectors to roughly 90%, plot (**a**). When both slits are opened simultaneously in stage (iii) (red circles) the path detectors receive nearly all light. Their combined share $T^{(iii)}_{\pm 1/2}$ of very nearly 50% each accounts for more than 99% of the total intensity. Concurrently, diffraction into higher orders is greatly suppressed. One has to zoom in to see that the transmission probabilities $T^{(iii)}_{n+1/2}(\phi=0)$ for orders other than $\pm 1/2$ are not zero, plot (**b**)

Since in stage (iii) amplitudes of adjacent orders have similar magnitudes but alternating signs, see (7), the resulting interference is destructive at the position of the wires: opening the second slit in stage (iii) of the experiment seemingly counteracts the perturbations due to the grating so clearly visible in stage (ii). Figures 4 and 5 illustrate the transition from stage (ii) to stage (iii). Afshar, erroneously, interpreted this as a recovery of the slit images to their form encountered in stage (i) [19]. In Ref. [20] he states with respect to stage (iii) of the experiment that "*no [. . .] diffraction takes place, since no attenuation in the transmitted light is observed*". Instead, our analysis and Figs. 4 and 5 show that in stage (iii) light from the 'wrong' slit mostly compensates for the disturbances introduced by the grating in stage (ii). This is also clearly seen in the functional form of the transition amplitudes (12): making use of the fact that amplitudes of adjacent orders are real and have opposite signs, see (7), we can also write $T^{(\mathrm{iii})}_{n+1/2}(\phi = 0) = (|t_{n-1/2}| - |t_{n+1/2}|)^2$. The amplitudes cancel out partially but not completely, because $|t_{n-1/2}| \neq |t_{n+1/2}|$: small residues of the perturbations must remain in stage (iii). Such residues are indeed seen in the Figure in Ref. [19] and Fig. 8(c) of Ref. [20]—just as explained by our analysis and to be expected from our Figs. 4(b) and 5(b) above.

We have to conclude that in the transition from stage (ii) to stage (iii) of the experiment the increase of the count rates in stage (iii) to levels similar to those observed in stage (i) is due to a partial compensation of the effects of the grid but not their complete circumvention. This partial compensation arises at the expense of redirecting light to the 'wrong' detector: not only $t_{n+1/2}$ contributes to $T^{(\mathrm{iii})}_{n+1/2}(\phi = 0)$, but also $t_{n-1/2}$—which originates from the other slit. This fact was overlooked by Afshar et al., instead, they claim to employ an entirely novel kind of "*non-perturbative measurement technique [...] conceptually different from quantum non-demolition or non-destructive techniques which do not destroy, but perturb the photon wavefunction directly*", see introduction of Ref. [21].

This rather strong claim appears to be unsubstantiated and in the light of our analysis, which explains all features of the experiment using elementary methods, quite unnecessary.

Before closing this section let us revisit our assumptions. The monochromaticity assumption is good for the narrow-band laser light used in the experiment, the extension to temporal wave packets is straightforward, so is an extension to wider slits. The reduction to two dimensions applies to the plane depicted in Fig. 1 and to a setup with long slits rather than round holes, it does not affect the essence of the experiment. An analysis of the exact setup described in Afshar's experiment is only more tedious but not different in principle, see Sect. 5.2 below. Modeling the grid by a planar reflecting film rather than round reflecting wires allows us to write down expressions for transmitted and reflected waves in simple form. This is obviously helpful for tracking the flow of light in this setup and in a real experiment this setup would perform better than the one used by Afshar because none of the transmitted light gets reflected towards the 'wrong' detector, it only reaches the wrong detector via diffraction.

4 Complementarity in Afshar's Experiment

4.1 Inferring is Not Measuring

The presence of the interference pattern in Afshar's experiment is meant to be 'inferred' [19–21], but according to the orthodox interpretation of quantum mechanics it has to be measured in order to be described by the quantum mechanical formalism. This requirement is well captured by Wheeler's famous dictum to the effect "that only a registered event is a real event" [31]. When studying the complementary aspects of a quantum particle's behavior both aspects have to be measured *simultaneously* for every particle [1, 32]. We will see below, that Afshar's setup actually allows to measure all quantities relevant for the quantification of inequality (1).

4.2 Two Ensembles: Transmitted and Reflected Light

Note that transmitted and back-reflected photons form two mutually exclusive subensembles to which the complementarity principle must therefore be applied separately—precisely because they are subjected to different simultaneous path and wave measurements. This important point was violated in the interpretation given in Ref. [21], see discussion in Sect. 5.1.

In the following, just as in Afshar's implementation of the experiment, we mostly deal with the transmitted light only, but everything we say can analogously be rephrased for the ensemble of back-reflected photons.

4.3 Qualitative Discussion of Complementarity

The interference pattern at the position of the grid, in stage (iii), is responsible for the reduced scattering of photons by the grid in this stage of the experiment. In order to measure the contrast of the interference pattern without compromising the simultaneous path detection we have to change the relative phase ϕ between light emanating from the two slits (by, say, changing the relative path length between the two holes with the help of a phase shifter, or, by moving the grating) and then collect all light *transmitted* by the grating. This is no problem since the path detectors collect most of the transmitted light. Therefore only the light which diffracts into the remaining higher orders, thus missing the two path detectors, has to be collected somehow behind (so as not to interfere with the simultaneous path determination) the path detectors.

For transmitted light the finite slit widths of the grating *reduce the observed fringe contrast* because the slits of the grating are not infinitesimally small and therefore act as bucket detectors. For an ideal characterization of the interference pattern fine spatial resolution is required, the interference pattern would have to be scanned with a detector equipped with a very narrow aperture. Instead, in the experiment the other extreme is implemented, the detectors are the two path detectors (other detectors could be set up behind these two to capture all transmitted photons diffracted into higher orders): jointly, these detectors operate as one bucket detector for the transmitted photons, covered by very wide slits that allow a large fraction of the light, namely $1 - a$, to pass. With such a large detector aperture good path detection of the photons is possible but almost no interference contrast can be observed because wide grating slits sample large parts of the interference pattern denying good resolution, for a quantification see curve for V_t^2 in Fig. 6 below.

Clearly this is the tradeoff between path detection and determination of the interference pattern contrast which lies at the heart of the principle of complementarity. The other extreme would be to employ very narrow grating slits (a near unity) which allow us to observe the modulation due to the interference pattern with great accuracy, but at the expense of strong diffraction of passing photons, thus erasing their path information.

Regrettably, this important point was incorrectly analyzed by Afshar et al. in Ref. [21] where they mixed the analysis for the mutually exclusive ensembles of reflected and transmitted photons, for details see Sect. 5.1 below.

4.4 Quantification of Visibility

The light in the grating plane forms a sinusoidal field distribution, with the intensity distribution $I(x)=\cos(\frac{\pi}{\Lambda}x+\phi)^2$, where ϕ is the relative phase between the two slits S_1 and S_2. To find out how much light gets transmitted we have to integrate over the grating's slit opening(s). We find that the transmitted intensity is given by $I_{t,\max}=\int_{-\Lambda/2\cdot(1-a)}^{+\Lambda/2\cdot(1-a)}dx\cos(\frac{\pi}{\Lambda}x)^2=(\pi-a\pi+\sin(a\pi))/(2\pi)$ in the maximum case (grating positioned at interference pattern minima, $\phi=0$) and $I_{t,\min}=\int_{-\Lambda/2\cdot(1-a)}^{+\Lambda/2\cdot(1-a)}dx\sin(\frac{\pi}{\Lambda}x)^2=(\pi-a\pi-\sin(a\pi))/(2\pi)$ in the minimum case (grating positioned at interference pattern maxima, $\phi=\pm\pi$). The ensuing measurable visibility of transmitted light V_t thus is

$$V_t(a)\doteq\frac{I_{t,\max}-I_{t,\min}}{I_{t,\max}+I_{t,\min}}=\frac{\sin(a\pi)}{\pi(1-a)}. \tag{13}$$

An analogous calculation for reflected light is easily performed and yields the result that the effective slit width in this case is given by $1-a$, namely the visibility of reflected light $V_r(a)=V_t(1-a)$. This is to be expected, for back-reflected photons the roles are reversed: small slits correspond to wide reflecting stripes and vice versa.

4.5 Quantification of Distinguishability

For a balanced interferometric setup such as Afshar's the distinguishability $\mathcal{D}$ of paths is determined by the power of the path detectors d_1 and d_2 to discriminate the two paths. Formally, it is given by half the distance between detector states in the trace class norm [12]. More intuitively, it is the sum of the differences in overlap $|\langle S_j|d_k\rangle|^2$ between the light mode emanating from slit S_j and the two detector modes d_1 and d_2. In order to illustrate this important point, let us concentrate on one slit, say, S_1. If the grid is not present (stage (i) of the experiment) no light is scattered towards the other detector, and the detectors can perfectly discriminate the slits. Mathematically this can be formulated by the fact that the mode-overlap $|\langle S_1|d_1\rangle|^2=1$ (all photons from slit S_1 reach detector d_1) and that no photons go astray and hit the wrong detector $|\langle S_1|d_2\rangle|^2=0$. The discrimination power of the detector d_1 is therefore given by the difference 'correct counts minus false counts', i.e. $|\langle S_1|d_1\rangle|^2-|\langle S_1|d_2\rangle|^2$. Since we deal with two detectors we arrive at Englert's expression [12] for the distinguishability of transmitted light $\mathcal{D}_t$, the weighted sum

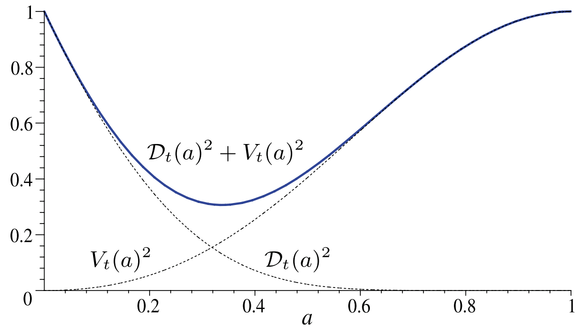
$$\mathcal{D}_t(a)=\frac{1}{2}\sum_{j=1}^{2}\left||\langle S_j|d_1\rangle|^2-|\langle S_j|d_2\rangle|^2\right| \tag{14}$$

$$=\frac{1}{2}\left(\left||t_0|^2-|t_1|^2\right|+\left||t_1|^2-|t_0|^2\right|\right) \tag{15}$$

$$=(1-a)^2-\left(\frac{\sin(a\pi)}{\pi}\right)^2. \tag{16}$$

In this derivation we used that only undeflected light and light that was diffracted into first order can reach either detector (compare Fig. 2(b)). Because of this unwanted

Fig. 6 (Color online) Complementarity is not violated in Afshar-type experiments: $\mathcal{D}_t^2(a) + V_t^2(a) \leq 1$ for all grid covering ratios a



scattering of photons from slit S_1 towards detector d_2 (and from S_2 towards d_1) the t_1-terms appear, showing that the path discrimination in Afshar's experiment is imperfect.

Just like in the case of visibility, V, distinguishability of transmitted and reflected light obeys $\mathcal{D}_t(a) = \mathcal{D}_r(1-a)$.

4.6 Quantification of Complementarity

Combining (13) and (16) yields

$$\mathcal{D}_t^2(a) + V_t^2(a) \leq 1, \tag{17}$$

where the limit of unity is only reached for the extreme cases of $a = 0$ (complete absence of a grid) or $a \approx 1$ (very small slits remain open allowing us to resolve the interference pattern very well: $V_t \to 1$), see Fig 6.

We have shown for all cases of plane reflecting grids that Afshar's experiment does not violate inequality (1).

4.7 More Ensembles: Transmitted, Lost and Reflected Light

One could, in principle, divide the ensembles of investigated particles further and, say, discard not only the back-reflected (or absorbed) photons but also those that are lost to higher diffraction orders, i.e. pass the grating but then miss the path detectors. The results qualitatively still obey an Englert-type inequality. For small grid covering ratios of $a = 0.05 \ldots 0.1$, employed in Afshar's implementation, little light gets deflected towards higher diffraction orders [21] and the above theory works well enough.

Another subtlety that was glossed over in the above treatment is due to the fact that depending on whether the grating is positioned at the minima or the maxima of the interfering light beams, $\phi = 0$ or $\pm\pi$, different numbers of photons get reflected or transmitted. In other words, the constitution or relative weight of the ensembles varies with ϕ even for fixed a. For small grid covering ratios a the effect is not too dramatic and we will therefore neglect this further complication. It shows once more that the Afshar setup is not an optimal design when it comes to investigations of complementarity.

5 Quantification of Afshar's Experiment

We will now quantify various quantities appearing in the particular experimental implementation reported by Afshar et al. [21]. It turns out that the data generated in this experiment conform quite well with the predictions of our theory, this leads me to conclude that an interpretation of the data as a demonstration of a violation of the principle of complementarity is unsupported.

5.1 Reflected and Transmitted Light

In Sect. 4.2 it was already pointed out that transmitted and back-reflected photons form two mutually exclusive subensembles to which the complementarity principle must be applied separately. Regrettably, this was not done by Afshar et al. In Sect. 4.1 of Ref. [21] they determine the fringe visibility of their setup by considering the irradiation of the grid wires—in the terminology employed here they estimated the visibility V_r of the interference pattern of the reflected photons.

But, photons that hit the wire will not reach the detector[2]!

The photons are therefore not available further downstream for their *simultaneous* path measurement. In other words, Afshar et al. combined the expression for the visibility of the reflected photons V_r with that for the distinguishability $\mathcal{D}_t$ of the transmitted photons. This combination is neither bound by unity nor does it tell us much about *simultaneous* measurements required for the investigation of the principle of complementary in quantum mechanics. Their estimate of '$V^2 + K^2 \geq 1.35$', where K corresponds to our $\mathcal{D}$, is therefore ill-conceived. In our language it amounts to combining $V_r(a)^2 + \mathcal{D}_t(a)^2$, a combination which can reach values of 2 and which does not have much meaning in the context of discussing complementarity. In contrast to a violation of inequality (1), as reported in Ref. [21], this combination is well behaved but quite meaningless.

5.2 Measured Values in Afshar's Experiment

In Ref. [21] it is reported that the grating consists of $N = 6$ wires of width 0.127 mm each which are spaced $(2\pi/2.462)$ mm ≈ 2.55 mm apart. With $0.127/2.55 \approx 0.04980$ the grating has a covering ratio $a \approx 0.05$ of roughly 5% and approximately covers a strip $N \times 2.55$ mm $= 15.3$ mm wide. If it is centered on the beam, such a strip blocks out roughly 84% of the beam which is reported to have an Airy-disk radius of $R = 10.7$ mm. We can conclude that roughly 16% of the photons reach the detectors without encountering the grating at all.

In order to cross-check the validity of our theory, let us determine the probability of a photon to be scattered towards the 'wrong' detector (compare Fig. 2(a)), namely $p_{\text{false}} = 0.84 \times |t_1(a)|^2 = 0.84 \times 2(\sin(a\pi)/\pi)^2$ which amounts to $p_{\text{false}} = 0.84 \times 2(\sin(0.05 \cdot \pi)/\pi)^2 \approx 0.0042$ and agrees with the smaller value of 0.41% for the scattering of photons emerging from pinhole A and hitting detector B reported in

[2] Here we disregard the imperfections due to the fact that the wires are round and reflecting: they could be made into flat strips as suggested in Sect. 3 or simpler still, they could be made absorbing.

the caption of Fig. 1 in Ref. [21]. This fortuitous agreement may well be somewhat coincidental, since our theory does not really apply to the case of non-uniform illumination as encountered in the cross section of the diffraction limited image of the circular apertures. The other reported value of 0.46% for a first order scatting rate implies that $a \approx 0.0526$.

These values for a do not match up with the reported reduction of the intensity at the path detectors of 14.14% and 14.62%. We worked out that the forward scattering amplitude is $t_0(a) = 1 - a$. The expected intensity at stage (ii) should therefore be equal to $0.16 \times 1 + 0.84 \times (1 - 0.05)^2 = 0.9181$ but instead is on the order of 0.855. We therefore use an effective value $a^\star$ which is meant to encompass the effects due to the fact that we analyzed the simplified plane-wave scenario. We find that $a^\star = 0.09$ yields the forward scattering probability $0.16 + 0.84 \cdot t_0(a^\star)^2 = 0.16 + 0.84 \cdot (1 - a^\star)^2 = 0.855604$ which conforms well with the reported intensity peaks.

As another check we now determine the same peak intensity in stage (iii), namely the forward scattering probability $0.16 + 0.84 \cdot (t_0(a^\star) + t_1(a^\star))^2 = 0.16 + 0.84 \cdot (1 - a^\star + \sin(a^\star\pi)/\pi)^2 = 0.997994$, this implies a relative reduction as compared to stage (i) by 0.2006%. In Ref. [21] it was reported that the photon count for the two path detectors in stage (iii) compared to stage (i) was reduced by 0.31% and 1.13% respectively, our conservative estimate of 0.2006% is not in conflict with these numbers. Despite the very considerable variation in these two measured values we can at least state that our conservative theoretical estimate predicts even smaller intensity reductions than observed in stage (iii) of the experiment but this does not allow us to conclude that these small numbers are due to a novel perturbation-free measurement approach, as was claimed in [21].

Let us finally quantify the complementarity aspects of this experiment.

For the 16%-or-so fraction of photons that do not come close to the grating no interference measurement has been performed $V_t = 0$ and the path distinguishability, $\mathcal{D}_t$, should equal unity since the path discrimination is effectively free of errors.

For the 84% of photons that do encounter the grating a grating covering ratio of $a = 0.05$ or $a^\star = 0.09$ implies $V_t(0.05)^2 + \mathcal{D}_t(0.05)^2 \approx 0.813$ or $V_t(0.09)^2 + \mathcal{D}_t(0.09)^2 \approx 0.682$. Clearly, neither the individual ensembles nor their weighted average, such as $0.16 \times 1 + 0.84 \times 0.813 = 0.84292$, violates inequality (1), as was claimed in [21].

The combination $V_r(0.09)^2 + \mathcal{D}_t(0.09)^2 \approx 1.646$ does of course exceed the value 'one' and indeed an estimate for its lower bound of '1.35' as given in Ref. [21], but it is not a useful quantity for our considerations, see Sect. 4.3.

To summarize, our theoretical values roughly conform with the experimentally observed values. All experimentally determined values remain comfortably above or below the respective theoretical bounds. No violations of established theory has been found. The variation in the experimentally generated data is too large to draw further, more detailed, conclusions.

5.3 Some Difficulties with Afshar's Experiment

Before we conclude let us mention some subtleties of the Afshar experiment which could be quantified—if one was so inclined.

The experiment is not very well designed to perform complementarity experiments because the amount of transmitted and reflected light depends on the relative position of the grating with respect to the interference pattern.

Another weakness of the setup are losses, partly due to the use of round wires in Afshar's implementation. But even in the simpler version discussed here, we still encounter losses into higher diffraction orders. Although they could be included into the analysis they only diminish the path-resolution further and the experimental setup would become more involved.

Finally, the greatest weakness in the analysis given by Afshar et al. [21] is the reliance on an *inference* that an interference pattern 'must' be present. Quantum mechanics is not an ontological theory, it must not be approached as such, only measured events are quantified by quantum mechanics [31, 32].

6 Conclusion

The quantification of the principle of complementarity in Sect. 4 shows that Afshar's experiment must not be interpreted as an example of a possible violation of the principle of complementarity of quantum mechanics. It is actually suboptimal in the sense that it does not fully exhaust the limits stipulated by quantum mechanics; unlike, say, the nearly optimal experiment performed by Dürr et al. [33].

The analysis given here also applies to some quantum-optical cases such as single photon, thermal or Glauber coherent light [34], it is therefore not to be expected that substantially different results could be found when performing Afshar's experiment with non-classical states of light.

Acknowledgement I wish to thank Dimitris Tsomokos for discussions.

References

1. Wheeler, J.A., Zurek, W.H.: Quantum Theory and Measurement. Princeton University Press, Princeton (1983)
2. Feynman, R.P., Leighton, R., Sands, M.: The Feynman Lectures on Physics. Addison–Wesley, Reading (1977)
3. Feynman, R.P., et al.: In: Feynman, R.P., Leighton, R., Sands, M. (eds.) The Feynman Lectures on Physics, vol. 1. Addison–Wesley, Reading (1977), Chap. 37-6
4. Scully, M.O., Englert, B.-G., Walther, H.: Nature **351**, 111 (1991)
5. Zajonc, A.G., Wang, L.J., Zou, X.Y., Mandel, L.: Nature **353**, 507 (1991)
6. Bhandari, R.: Phys. Rev. Lett. **69**, 3720 (1992)
7. Tan, S.M., Walls, D.F.: Phys. Rev. A **47**, 4663 (1993)
8. Storey, E.P., Tan, S.M., Collet, M.J., Walls, D.F.: Nature **367**, 626 (1994)
9. Englert, B.-G., Scully, M.O., Walther, H.: Nature **375**, 367 (1995)
10. Storey, E.P., Tan, S.M., Collet, M.J., Walls, D.F.: Nature **375**, 368 (1995)
11. Wiseman, H., Harrison, F.: Nature **377**, 584 (1995)
12. Englert, B.-G.: Phys. Rev. Lett. **77**, 2154 (1996)
13. Björk, G., Söderholm, J., Trifonov, A., Tsegaye, T., Karlsson, A.: Phys. Rev. A **60**, 1874 (1999)
14. Steuernagel, O.: quant-ph/9908011 (1999)

15. Wootters, W.K., Zurek, W.H.: Phys. Rev. D **19**, 473 (1979)
16. Greenberger, D.M., Yasin, A.: Phys. Lett. A **128**, 391 (1988)
17. Mandel, L.: Opt. Lett. **16**, 1882 (1991)
18. Jaeger, G., Shimony, A., Vaidman, L.: Phys. Rev. A **51**, 54 (1995)
19. Chown, M.: New Scientist **183**, 30 (2004)
20. Afshar, S.S.: quant-ph/0701027 (2007)
21. Afshar, S.S., Flores, E., McDonald, K.F., Knoesel, E.: Found. Phys. **37**, 295 (2007)
22. Kastner, R.E.: quant-ph/0502021 (2005)
23. Srinivasan, R.: quant-ph/0504115 (2005)
24. Drezet, A.: quant-ph/0508091 (2005)
25. O'Hara, P.: quant-ph/0608202 (2006)
26. Georgiev, D.: Prog. Phys. **2**, 97 (2007)
27. Qureshi, T.: quant-ph/0701109 (2007)
28. Chown, M.: New Sci. **193**, 13 (2007)
29. Flores, E., Knoesel, E.: quant-ph/0702210 (2007)
30. Feynman, R.P., et al.: In: Feynman, R.P., Leighton, R., Sands, M. (eds.) The Feynman Lectures on Physics, vol. 1. Addison–Wesley, Reading (1977), Chap. 33-6
31. Cramer, J.G.: Phys. Rev. D **22**, 362 (1980)
32. Feynman, R.P., et al.: In: Feynman, R.P., Leighton, R., Sands, M. (eds.) The Feynman Lectures on Physics, vol. 1. Addison–Wesley, Reading (1977), Chap. 38-6
33. Dürr, S., Nonn, T., Rempe, G.: Phys. Rev. Lett. **81**, 5705 (1998)
34. Steuernagel, O.: Phys. Rev. A **65**, 013809 (2002)
35. Chown, M.: 'The Independent' newspaper, 6 October 2004
36. Espinasse, P.: SPIE OE Mag. **5**, 8 (2005)