Quantum Weak Measurement

A Review of AAV and SSD Contributions

– Ritik Dubey (22MS208)

Under the supervision of **Prof. Nirmalya Ghosh**

Department of Physical Sciences, IISER Kolkata

Abstract

This report reviews the theory of quantum weak measurement based on the works of Aharonov, Albert, and Vaidman (AAV) and later analysis by Duck, Stevenson, and Sudarshan (SSD). Starting from the von Neumann measurement model, it explains how weak coupling between a system and a measuring device yields the weak value, which can exceed the observable's eigenvalue range under nearly orthogonal pre- and post-selections. SSD's treatment refines AAV's approximation and confirms its validity in the weak regime. The equivalence of both results is demonstrated graphically and analytically. Experimental realizations through the Stern–Gerlach setup and an optical analog are discussed, showing how weak value amplification enables detection of extremely small physical effects.

Keywords: Quantum weak measurement; Weak value; Pre-selection and post-selection; Weak-value amplification.

Contents

1 Theory

1.1 Von Neumann measurement Idea

AAV's discussion begins with the standard von Neumann measurement model. The key idea of **von Neumann** was, one can see that it is possible in principle to correlate the state of a microscopic quantum system with the value of a macroscopic classical variable, and we may take it for granted that we can perceive the value of the classical variable.
the hamiltonian discribing the coupling of the quantum system to the pointer has the form,

$$\hat{H} = \hat{H}_0 + \frac{1}{2m}\hat{P}^2 + \lambda \hat{A}\hat{q}$$

where,

$\hat{H}_0$**:** Unperturbed Hamiltonian of the quantum system.

$\frac{\hat{P}^2}{2m}$**:** Hamiltonian of the free pointer particle.

$\hat{A}$**:** Observable operator of the system (e.g., spin operator).

$\hat{q}$**:** Position operator of the pointer.

During the brief measurement interval, the interaction Hamiltonian can be assumed to dominate all other terms in the total Hamiltonian **(for more detail see supplementary info section)**. Under this assumption, the coupling between the quantum system and the pointer is described by an interaction Hamiltonian of the form:

$$\boxed{\hat{H} = -\, g(t)\, \hat{A}\hat{q}}$$

where $g(t)$ is a normalized coupling function defined over the measurement duration, $\hat{A}$ is Observable operator of the system (e.g., a spin operator) and $\hat{q}$ is Position operator of the pointer.

1.2 Strong v/s Weak Measurement

Let the measuring device start in state $|\phi\rangle$ with a Gaussian momentum-space wave function $\phi(p)$ centered at $p = 0$ and width Δp. Its position-space wave function $\phi(q)$, the Fourier transform of $\phi(p)$, is also Gaussian.

$$\phi(q) \equiv \langle q|\phi\rangle = \exp\left[-\frac{q^2}{4\Delta^2}\right],$$

$$\phi(p) \equiv \langle p|\phi\rangle = \exp\left[-\Delta^2 p^2\right], \tag{1.1}$$

where

$$\Delta q \equiv \Delta, \qquad \Delta p = \frac{1}{2\Delta},$$

Note: The system–pointer interaction is represented by a **unitary evolution**, since the combined system is isolated during the measurement. According to the Schrodinger, its time evolution is $U = \exp\left(-\frac{i}{\hbar}\hat{H}_{\text{int}}t\right)$, ensuring **conservation of total probability** (norm preservation).

say, Preselected state,

$$|\psi_{in}\rangle = \sum_n \alpha_n |a_n\rangle \qquad \text{(in superposition of eigensates of } \hat{A})$$

Postselected state,
After a weak interaction, a strong measurement is performed and only the outcomes corresponding to a chosen final state $|\psi_f\rangle$ are kept.
say,

$$|\psi_f\rangle = \sum_j \beta_j |a_j\rangle \qquad \text{(in superposition of some eigenstates of } \hat{A})$$

After Postselection, the complete system evolve as ,

$$\begin{aligned}\langle\psi_f|\Psi\rangle &= \langle\psi_f|\,\hat{U}(|\psi_{in}\rangle\otimes|\phi\rangle)\\ &= \langle\psi_f|\,e^{i\hat{q}\hat{A}}\,|\psi_{in}\rangle\otimes|\phi\rangle\\ &= \sum_n\sum_j \alpha_n\beta_j^*\;\delta_{j,n}\;e^{i\hat{q}a_n}\,|\phi\rangle\\ &= \sum_n \alpha_n\beta_n^*\;e^{ia_n\hat{q}}\;\;|\phi\rangle\end{aligned}$$

now, projecting on pointer's momentum, we get

$$\begin{aligned}\langle p|\Psi\rangle &= \sum_n \alpha_n\beta_n^*\,\langle p|\,e^{ia_n\hat{q}}\,|\phi\rangle\\ &= \sum_n \alpha_n\beta_n^*\,\phi(p-a_n)\\ &= \sum_n \alpha_n\beta_n^*\exp\left[-\frac{(p-a_n)^2}{4(1/2\Delta)^2}\right] \qquad \text{(No approximation used till now)}\\ \implies &\boxed{\langle p|\Psi\rangle = \sum_n \alpha_n\beta_n^* e^{-\Delta^2(p-a_n)^2}} \qquad \text{(SSD's Result)}\end{aligned}$$

Note: Superposition of Gaussians, each shifted by its eigenvalue a_n.

Strong vs Weak measurement

If the pointer's momentum spread δ is small compared to the spacing between the eigenvalues a_n, the measuring device ends up in well-separated peaks, each centered at an eigenvalue a_n (**Strong Measurement**). In the limit $\delta \to 0$, this corresponds to an ideal measurement where the outcome is always one of the eigenvalues, the probability of obtaining a_n is $|\alpha_n|^2$, and the system collapses to the corresponding eigenstate $|A = a_n\rangle$.
However, in the opposite limit, in which δ is much larger than the spread of the a_n's. AAV refer to this case as a **"weak measurement."** The final state of the "measuring" device is then a superposition of strongly overlapping, broad Gaussians. (SSD's Result)

Note: any single *weak measurement* gives almost no information, since $\Delta p \gg \langle A\rangle$. However, by repeating the experiment many times, one can map out the whole distribution and thus determine the centroid $\langle A\rangle$ to any desired accuracy.

AAV's work

$$\begin{aligned}\langle\psi_f|\Psi\rangle &= \langle\psi_f|\left(\hat{U}(|\psi_{in}\rangle\otimes|\phi\rangle)\right)\\ &= \underbrace{\langle\psi_f|\,e^{i\hat{A}\hat{q}}\,|\psi_{in}\rangle}_{\text{scalar}}\otimes|\phi\rangle\\ &= \langle\psi_f|\left(\sum_{n=0}^{\infty}\frac{(\hat{A}\hat{q})^n}{n!}\right)|\psi_{in}\rangle\,|\phi\rangle\\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\,\langle\psi_f|\,(\hat{A}^n\hat{q}^n)\,|\psi_{in}\rangle\,|\phi\rangle \qquad \text{(since } [\hat{A},\hat{q}]=0\text{, so, } (\hat{A}\hat{q})^n=\hat{A}^n\hat{q}^n\text{)}\end{aligned}$$

$$\begin{aligned}\langle\psi_f|\Psi\rangle &= \langle\psi_f|\psi_{in}\rangle\sum_{n=0}^{\infty}\left\{\frac{1}{n!}\frac{\langle\psi_f|\,\hat{A}^n\,|\psi_{in}\rangle}{\langle\psi_f|\psi_{in}\rangle}\hat{q}^n\right\}|\phi\rangle\\ &= \langle\psi_f|\psi_{in}\rangle\sum_{n=0}^{\infty}\left\{\frac{1}{n!}\langle\hat{A}^n\rangle_w\hat{q}^n\right\}|\phi\rangle \qquad \text{(defined a weak value)}\\ &= \langle\psi_f|\psi_{in}\rangle\left\{1+\langle\hat{A}\rangle_w\hat{q}+\frac{\langle\hat{A}^2\rangle_w\hat{q}^2}{2!}+\dots\right\}|\phi\rangle\end{aligned}$$

if expansion is truncated for $n \geq 2$, then ,

$$\begin{aligned}\langle p|\Psi\rangle &= \langle\psi_f|\psi_{in}\rangle\ e^{i\langle\hat{A}\rangle_w \hat{q}}\ \langle p|\phi\rangle \qquad \text{(becomes a Translator operator in } p\text{)}\\ &= \langle\psi_f|\psi_{in}\rangle \quad \underbrace{\phi(p - A_w)}_{\text{shift by weak value}}\end{aligned}$$

$$\implies \boxed{\langle p|\Psi\rangle = \sum_n \alpha_n \beta_n^*\ e^{-\Delta^2 (p-A_w)^2}} \qquad \text{AAV's result}$$

Note : this one is a single gaussian peaked at weak value.

$$\langle\hat{A}\rangle_w = \frac{\langle\psi_f|\,\hat{A}\,|\psi_{in}\rangle}{\langle\psi_f|\psi_{in}\rangle}$$

Weak values are **not bounded** by the eigenvalue range because they are not probability-weighted averages but ratios of complex amplitudes. When the pre- and post-selected states are nearly orthogonal, the denominator can be arbitrarily small, making the weak value arbitrarily large (or even complex).

Accurate Truncation condition (by SSD) :

since,

$$\frac{i\hat{q}^n}{n!}\langle\psi_f|\,\hat{A}^n\,|\psi_{in}\rangle \ll i\hat{q}\,\langle\psi_f|\,\hat{A}\,|\psi_{in}\rangle < \langle\psi_f|\psi_{in}\rangle \qquad ,\forall\ \mathrm{n} \geq 2$$

so, one and only conddition is,

$$\frac{i\hat{q}^2}{n!}\langle\psi_f|\,\hat{A}^n\,|\psi_{in}\rangle \ll i\hat{q}\,\langle\psi_f|\,\hat{A}\,|\psi_{in}\rangle \qquad ,\forall\ \mathrm{n} \geq 2$$

coparing order,

$$\begin{aligned}\Delta^n\langle\hat{A}^n\rangle_w &\ll \Delta\langle\hat{A}\rangle_w\\ \Delta^{n-1} &\ll \frac{\langle\hat{A}\rangle_w}{\langle\hat{A}^n\rangle_w}\\ \Delta &\ll \left(\frac{\langle\psi_f|\,\hat{A}\,|\psi_{in}\rangle}{\langle\psi_f|\,\hat{A}^n\,|\psi_{in}\rangle}\right)^{\frac{1}{n-1}}\\ \implies &\boxed{\Delta \ll \min_{n\geq 2}\left|\frac{\langle\psi_f|\,\hat{A}\,|\psi_{\mathrm{in}}\rangle}{\langle\psi_f|\,\hat{A}^n\,|\psi_{\mathrm{in}}\rangle}\right|^{\frac{1}{n-1}}}\end{aligned}$$

this is more accurate condition for the validity of AAV's Result !

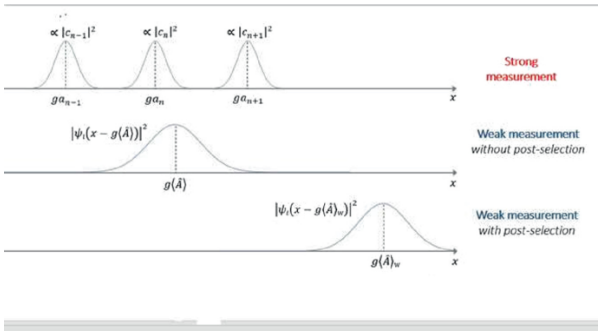

Figure 1.1: Pictorial Summery

1.3 The Paradox

• from AAV's paper,

$$\boxed{\langle p|\Psi\rangle = \sum_n \alpha_n \beta_n^* \; e^{-\Delta^2(p-A_w)^2}} \qquad \text{(single gaussian peaked at } A_w\text{)}$$

• from Duck's paper,

$$\boxed{\langle p|\Psi\rangle = \sum_n \alpha_n \beta_n^* e^{-\Delta^2(p-a_n)^2}} \qquad \text{(Superposition of gaussians each peaked at } a_n\text{)}$$

Paradox :
How can $|\Psi\rangle$ be a single gaussian peaked at A_w, while simultaneously being a superposition of gaussians each peaked at value a_n, where A_w may well be much greater than any of the a_n's? How can a large shift be produced from a superposition of small shifts?
-The resolution of the paradox lies in the fact that the superposition of Gaussian involves complex coefficient here.(we will see how both results are same in Stern-Garlach section)

2 Weak Stern-Gerlach Experiment

2.1 Introduction

A beam of spin-$\frac{1}{2}$ particles travels along the y-direction with spins prepared in the xz-plane at an angle α to the x-axis. The particles' spatial wavefunction is Gaussian in z with width Δ, giving a momentum spread $\delta = \frac{1}{2\Delta}$. The beam first passes through a weak Stern–Gerlach magnet, where the field gradient is small so that the momentum shift $\delta p_z \ll \Delta p_z$. As a result, the spatial wavefunction becomes a superposition of two slightly shifted Gaussians correlated with $\sigma_z = \pm 1$, but still strongly overlapping. The beam then enters a strong Stern–Gerlach magnet aligned along x, which separates the σ_x eigenstates, and only the $\sigma_x = +1$ component is post-selected. This selected beam travels freely to a distant screen. The screen is placed sufficiently far so that the displacement in the z-direction due to the average momentum p_z acquired during the weak interaction exceeds the initial position uncertainty Δz. On the screen, a wide spot is obtained whose displacement along z is measured, and this displacement yields the weak value of σ_z.

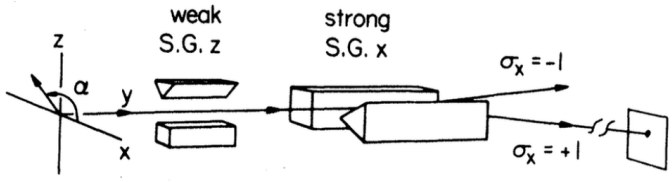

Figure 1.2: Stern-Gerlach setup layout for the AAV experiment.

For this case, the interaction Hamiltonian is given by,

$$\boxed{\hat{H} \;=\; -g(t)\mu_B\,\hat{\sigma}_z\,\frac{\partial B_z}{\partial z}\,\hat{z}}$$

which describes the coupling between the spin and the magnetic field gradient along the z-direction.
(for derivation, see Supplementary Info section)

2.2 AAV's result validation

The overall wavefunction of the system can be expressed as,

$$|\Psi\rangle = \underbrace{\left(\frac{1}{2\pi\Delta^2}\right)^{3/4}\exp\left[-\left(\frac{x^2+y^2+z^2}{4\Delta^2}\right)\right]}_{\text{Gaussian in 3D position}}\;\underbrace{\exp(-iP_0 y)}_{\text{momentum kick } P_0 \text{ along } y}\;|\psi_{\text{in}}\rangle$$

representing a Gaussian beam moving along the y-axis with momentum P_0.
where,

$$|\psi_{\text{in}}\rangle = \cos\left(\frac{\pi/2-\alpha}{2}\right)|\uparrow_z\rangle + \sin\left(\frac{\pi/2-\alpha}{2}\right)e^{i0}\,|\downarrow_z\rangle$$

$$|\psi_f\rangle = |\uparrow_x\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle + |\downarrow_z\rangle)$$

Spin Eigenstates in an Arbitrary Direction

The eigenstates of spin along an arbitrary direction $\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ are given by:

$$|\uparrow_n\rangle = \cos\left(\frac{\theta}{2}\right)|\uparrow_z\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|\downarrow_z\rangle\,,$$

$$|\downarrow_n\rangle = -e^{-i\phi}\sin\left(\frac{\theta}{2}\right)|\uparrow_z\rangle + \cos\left(\frac{\theta}{2}\right)|\downarrow_z\rangle\,.$$

These are eigenstates of the spin operator along $\mathbf{n}$:

$$\hat{S}_n\,|\uparrow_n\rangle = +\frac{\hbar}{2}\,|\uparrow_n\rangle\,, \qquad \hat{S}_n\,|\downarrow_n\rangle = -\frac{\hbar}{2}\,|\downarrow_n\rangle\,.$$

Using these definitions, the weak value of $\hat{\sigma}_z$ becomes,

$$(\hat{\sigma}_z)_w = \frac{\langle\psi_f|\,\hat{\sigma}_z\,|\psi_{\text{in}}\rangle}{\langle\psi_f|\psi_{\text{in}}\rangle} \quad\Longrightarrow\quad \boxed{(\hat{\sigma}_z)_w = \tan\Big(\frac{\alpha}{2}\Big)}.$$

from our theory we can directly say,

$$\begin{aligned}\langle p|\Psi\rangle &= \langle\psi_f|\psi_{\text{in}}\rangle\,\exp\Big[-\Delta^2\big(p_z - A_w\big)^2\Big] \\ &= \cos\frac{\alpha}{2}\,\exp\Big[-\Delta^2\big(p_z - \lambda\tan\frac{\alpha}{2}\big)^2\Big] \qquad \big(\lambda = \mu\frac{\partial B_z}{\partial z}\big)\end{aligned}$$

Note: at $\alpha = \pi \implies \tan(\alpha/2) \to \infty$
therefore,

$$\boxed{\delta p_z \;=\; \mu\,\frac{\partial B_z}{\partial z}\,\min\Big\{|\tan\tfrac{\alpha}{2}|,\;1\Big\} \;\ll\; \Delta p_z \sim \frac{1}{2\Delta}}$$

the particle travels a distance l along the y-axis while experiencing a small deflection δz in the z-direction. From the following geometry, $\tan(\theta) = \delta z/l = \delta p_z/p_0$, giving $\delta p_z = (\delta z/l)p_0$. **This relation shows that the shift in the pointer's momentum corresponds to the weak value, which can exceed the eigenvalue range when the pre- and post-selected states are nearly orthogonal.**

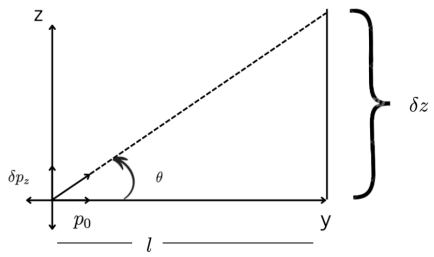

Figure 1.3: δp_z measurement schematics

2.3 Resolving the paradox

SSD's paper takes the analysis a step further, providing a deeper understanding of the AAV result. Since the interesting effects occur at $\alpha = \pi$, for a sufficiently small Δ we can approach $\alpha \approx \pi$. To analyze this more clearly, let us set $\alpha = \pi - 2\epsilon$ with $\epsilon \ll 1$.
then AAV's result reduces to,

$$\begin{aligned}\langle p|\Psi\rangle &= \cos\frac{\alpha}{2}\;\exp\Big[-\Delta^2\Big(p_z - \lambda\tan\frac{\alpha}{2}\Big)^2\Big] \\ &= \epsilon\,\exp\Big[-\Delta^2\Big(p_z - \frac{\lambda}{\epsilon}\Big)^2\Big]\end{aligned}$$

valid iff, shift $\ll$ spread, i.e,

$$\begin{aligned}\frac{\lambda}{\epsilon} &\ll \Delta p_z = \frac{1}{2\Delta} \\ \Longrightarrow\; &\boxed{\lambda\Delta \ll \epsilon \ll 1}\end{aligned}$$

now, Duck's result reduces to,

$$\begin{aligned}\langle p|\Psi\rangle &= \sum_n \alpha_n\beta_n^*\exp\Big[-\Delta^2(p_z - \lambda a_n)^2\Big] \\ &= \alpha_1\beta_1^* e^{-\Delta^2(p_z-\lambda a_1)^2} + \alpha_2\beta_2^* e^{-\Delta^2(p_z-\lambda a_2)^2}\end{aligned}$$

$$
\begin{aligned}
|\psi_{\text{in}}\rangle &= \frac{1}{\sqrt{2}}\Big[\big(\cos\frac{\alpha}{2}+\sin\frac{\alpha}{2}\big)\,|\uparrow_z\rangle + \big(\cos\frac{\alpha}{2}-\sin\frac{\alpha}{2}\big)\,|\downarrow_z\rangle\Big] \\
&= \frac{1}{\sqrt{2}}\Big[(\epsilon+1)\,|+\rangle + (\epsilon-1)\,|-\rangle\Big] \\
|\psi_f\rangle &= \frac{1}{\sqrt{2}}\big(|\uparrow_z\rangle + |\downarrow_z\rangle\big)
\end{aligned}
$$

$$
\begin{aligned}
\langle p|\Psi\rangle &= \left(\frac{1}{\sqrt{2}}(\epsilon+1)\right)\frac{1}{\sqrt{2}}e^{-\Delta^2(p_z-\lambda)^2} + \left(\frac{1}{\sqrt{2}}(\epsilon-1)\right)\frac{1}{\sqrt{2}}e^{-\Delta^2(p_z+\lambda)^2} \\
&= \left(\frac{1}{\sqrt{2}}\right)^2\Big[(\epsilon+1)e^{-\Delta^2(p_z-\lambda)^2} - (1-\epsilon)e^{-\Delta^2(p_z+\lambda)^2}\Big]
\end{aligned}
$$

therefore,

$$
\langle p|\Psi\rangle = \left(\frac{1}{\sqrt{2}}\right)^2\Big[(\epsilon+1)e^{-\Delta^2(p_z-\lambda)^2} - (1-\epsilon)e^{-\Delta^2(p_z+\lambda)^2}\Big] \quad \text{(No condition applied)}
$$

$$
\langle p|\Psi\rangle = \epsilon\,\exp\Big[-\Delta^2\Big(p_z-\frac{\lambda}{\epsilon}\Big)^2\Big] \quad \text{(AAV's truncation condition applied)}
$$

They demonstrated through graphical analysis that the SSD result smoothly approaches the AAV expression in the weak–interaction limit. In other words, AAV's formula is valid strictly within the *weak measurement regime*, while the SSD treatment is more general and automatically reduces to the AAV expression when the coupling becomes weak.

2.4 Graphical and Analytical proof

Below, we present both the graphical and analytical reasoning that resolves the apparent paradox. For the graphical analysis, we simulated only the SSD result for different parameter choices to illustrate how the peak position evolves across regimes.

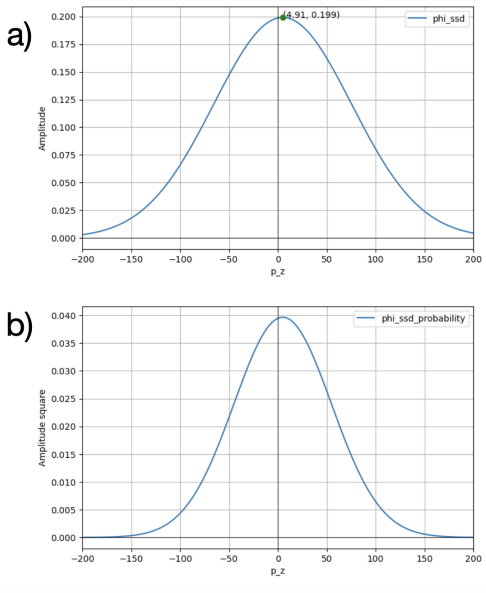

Figure 1.4: Simulated with parameters $\epsilon = 0.2$ and $\lambda\Delta = 0.01$. The resulting Gaussian is shifted by $4.91 \approx 5$.

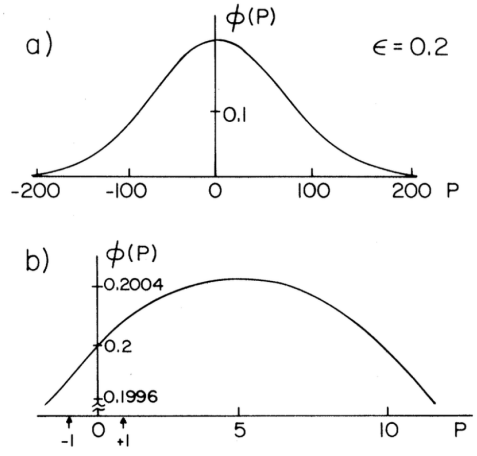

FIG. 2. (a) Graph of the function $\varphi(p;\epsilon,\Delta,\lambda)$ [see Eq. (33)] as a function of $P \equiv p/\lambda$ with $\lambda\Delta = 0.01$ and $\epsilon = 0.2$. Note that the function resembles a single Gaussian whose peak, shown in a closeup in (b), is shifted to $P = 1/\epsilon = 5$.

Figure 1.5: from original SSD's paper

Description :
For $\epsilon = 0.2$, the SSD pointer function lies in the weak–interaction regime and produces a single Gaussian peak at $\cot\epsilon \approx 4.91$, in excellent agreement with the expected weak value shift.

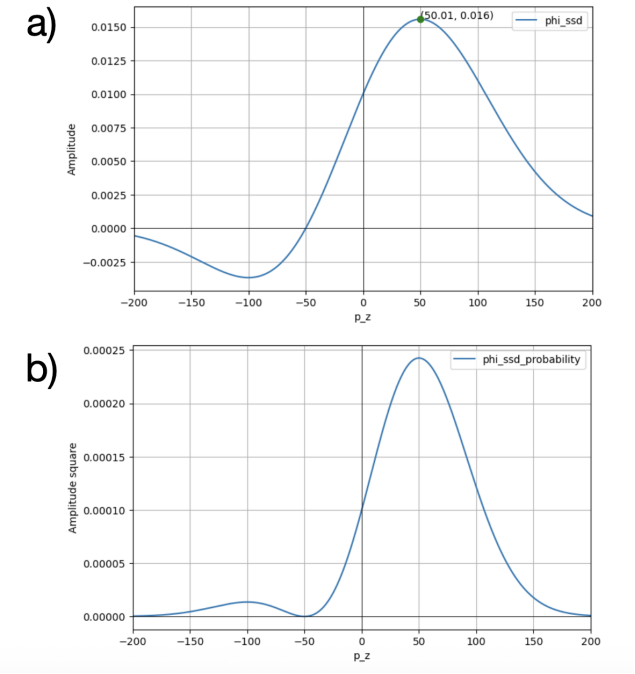

Figure 1.6: Simulated with parameters $\epsilon = 0.01$ and $\lambda\Delta = 0.01$. The resulting Gaussian is shifted by 50.01.

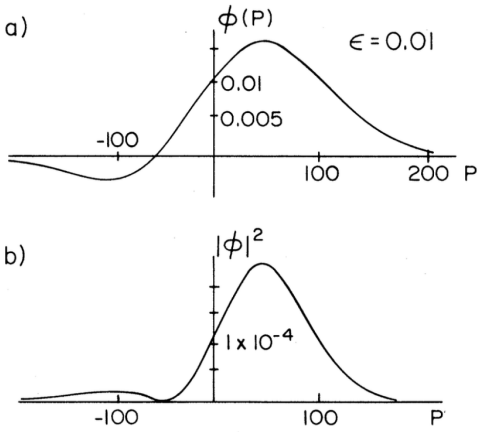

FIG. 3. (a) As Fig. 2, but for $\epsilon = \lambda\Delta = 0.01$. Note that although the component Gaussians peak at $+1$ and -1, the composite function is shifted by about half the width $1/\Delta$. (b) The resulting probability distribution $|\phi(P)|^2$.

Figure 1.7: from original SSD's paper

Description :

For $\epsilon = 0.01$, the system enters the deep weak–interaction regime in which the two SSD component Gaussians almost completely overlap and interfere, leaving a single, effectively Gaussian pointer whose peak no longer follows $1/\epsilon$ but instead saturates at $p_{\text{peak}} \approx 50$ (for $\Delta = 0.01$), consistent with the SSD prediction near the saturation limit.

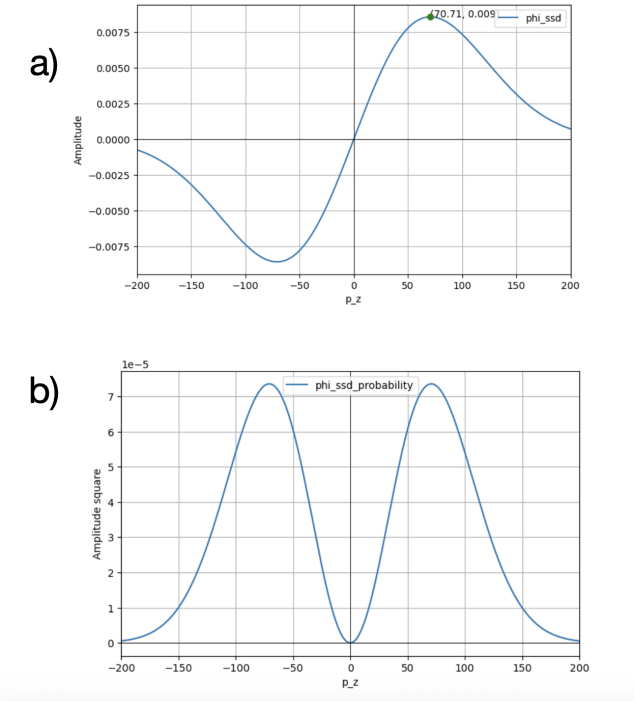

Figure 1.8: Simulated with parameters $\epsilon = 0.0$ and $\lambda\Delta = 0.01$. The resulting Gaussian is shifted by 70.

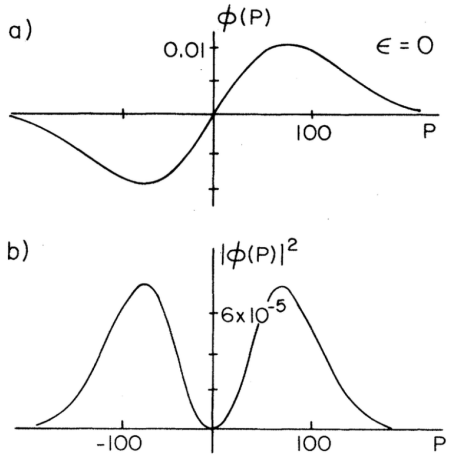

FIG. 4. (a) As in Fig. 2, but for $\epsilon = 0$. (b) The resulting probability distribution $|\varphi(P)|^2$.

Figure 1.9: from original SSD's paper

Description :

For $\epsilon = 0$ the SSD solution becomes antisymmetric, and the probability distribution develops two peaks at $p = \pm 1/(\sqrt{2}\Delta)$; for $\Delta = 0.01$ this gives $p_{\text{peak}} \approx 70.7$, exactly matching the simulated value.

For the full simulation videos, refer to the following Google Drive folder: https://drive.google.com/drive/folders/XXXXXXXXXXXX. Here, only the spread of the wavepacket increases, and this eventually produces the weak regime where, after destructive interference, we are left with a very small-amplitude single Gaussian peaked at the weak value (for our simulation, $1/\epsilon = 10$).

Analytical Proof

We start from Duck's expression:

$$\langle p|\Psi\rangle = \frac{1}{2}\Big[(\epsilon+1)e^{-\Delta^2(p_z-\lambda)^2} - (1-\epsilon)e^{-\Delta^2(p_z+\lambda)^2}\Big].$$

Factor out a common Gaussian envelope:

$$\begin{aligned}\langle p|\Psi\rangle &= \frac{e^{-\Delta^2 p_z^2}}{2}\Big[(\epsilon+1)e^{2\Delta^2 p_z\lambda - \Delta^2\lambda^2} - (1-\epsilon)e^{-2\Delta^2 p_z\lambda - \Delta^2\lambda^2}\Big] \\ &= \frac{e^{-\Delta^2(p_z^2+\lambda^2)}}{2}\Big[(\epsilon+1)e^{2\Delta^2 p_z\lambda} - (1-\epsilon)e^{-2\Delta^2 p_z\lambda}\Big].\end{aligned}$$

In the *weak regime* $\Delta\lambda \ll 1$ we may Taylor-expand $e^{\pm 2\Delta^2 p_z\lambda} \approx 1 \pm 2\Delta^2 p_z\lambda$
Substituting gives:

$$\begin{aligned}\langle p|\Psi\rangle &\approx \frac{e^{-\Delta^2(p_z^2+\lambda^2)}}{2}\Big[(\epsilon+1)(1+2\Delta^2 p_z\lambda) - (1-\epsilon)(1-2\Delta^2 p_z\lambda)\Big] \\ &= \frac{e^{-\Delta^2(p_z^2+\lambda^2)}}{2}\big[2\epsilon + 4\Delta^2 p_z\lambda\big] \\ &= e^{-\Delta^2(p_z^2+\lambda^2)}\big[\epsilon + 2\Delta^2 p_z\lambda\big].\end{aligned}$$

Factor out ϵ and use $1 + x \approx e^x$ for small x:

$$\begin{aligned}\langle p|\Psi\rangle &\approx \epsilon\, e^{-\Delta^2(p_z^2+\lambda^2)}\left(1 + \frac{2\Delta^2 p_z\lambda}{\epsilon}\right) \\ &\approx \epsilon\, e^{-\Delta^2(p_z^2+\lambda^2)}\exp\Big(\frac{2\Delta^2 p_z\lambda}{\epsilon}\Big) \\ &= \epsilon\exp\Big[-\Delta^2\Big(p_z^2 - \frac{2p_z\lambda}{\epsilon} + \lambda^2\Big)\Big].\end{aligned}$$

Complete the square inside the exponent:

$$p_z^2 - \frac{2p_z\lambda}{\epsilon} + \lambda^2 = \Big(p_z - \frac{\lambda}{\epsilon}\Big)^2 + \Big(\lambda^2 - \frac{\lambda^2}{\epsilon^2}\Big).$$

Hence:

$$\langle p|\Psi\rangle \approx \epsilon\,\exp\Big[-\Delta^2\Big(p_z - \frac{\lambda}{\epsilon}\Big)^2\Big]\,\exp\Big[-\Delta^2\Big(\lambda^2 - \frac{\lambda^2}{\epsilon^2}\Big)\Big].$$

The second exponential is p_z-independent and can be absorbed into the overall normalization. Therefore, up to normalization:

$$\boxed{\langle p|\Psi\rangle \propto \epsilon\exp\Big[-\Delta^2\Big(p_z - \frac{\lambda}{\epsilon}\Big)^2\Big]},$$

which is precisely the AAV form. □

3 Optical Analog

Mathematical Analogy: Electron's spin states ↔ Light's polarization states.

An optical analog of the previous experiment is constructed using polarized light. A laser beam, expanded with lenses, provides the broad coherent source. A polarizer at angle α and an analyzer at angle β define the initial and final polarizations, while a weakly birefringent crystal serves as the weak measurement device. It introduces a small lateral displacement between the ordinary (x-polarized) and extraordinary (z-polarized) components, much smaller than the beam width and determined by the crystal properties.

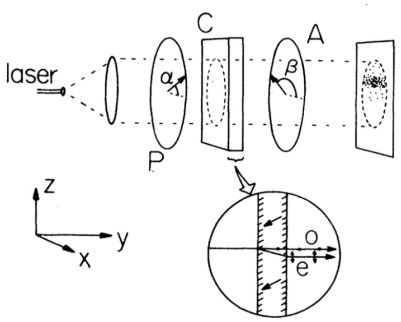

Figure 1.10: An optical analog schematic. A broad, coherent beam passes through a polarizer (P) and an analyzer (A). Between them, a birefringent crystal (C) is placed.

Note: The shift is lateral rather than angular (as in the Stern–Gerlach case), the analysis focuses on the spatial z-distribution instead of momentum.

The input beam is assumed to have a wide Gaussian profile.
After passing through polarizer, the preselected state becomes,

$$|\psi_{in}\rangle = \cos(\alpha)\,|\hat{x}\rangle + \sin(\alpha)\,|\hat{z}\rangle$$

following the theory, we can say,

$$|\Psi_{in}\rangle = |\psi_{in}\rangle \otimes |\phi_{in}\rangle$$

now, projecting on pointer's position,

$$\begin{aligned}\langle q|\Psi_{in}\rangle &= |\psi_{in}\rangle \otimes \langle q|\phi_{in}\rangle \\ &= |\psi_{in}\rangle\, e^{-\frac{z^2}{4\Delta^2}} \\ &= \left(\cos(\alpha)\,|\hat{x}\rangle + \sin(\alpha)\,|\hat{z}\rangle\right)\, e^{-\delta^2 z^2} \qquad \text{,gaussian peaked at } z=0\end{aligned}$$

After passing through the crystal, the two orthogonal polarization components experience lateral displacements a_1 and a_2 along the $\hat{z}$-direction. Hence, we can express the field as:

$$\langle q|\Psi\rangle = \underbrace{\cos(\alpha)\; e^{-\delta^2(z-a_1)^2}}_{\text{gaussian peaked at } z=a_1}\; |\hat{x}\rangle + \underbrace{\sin(\alpha)\; e^{-\delta^2(z-a_2)^2}}_{\text{gaussian peaked at } z=a_2}\; |\hat{z}\rangle$$

after passing through Analyser,

$$\begin{aligned}\langle \psi_f|\Psi\rangle &= \Big(\cos(\beta)\,|\hat{x}\rangle + \sin(\beta)\,|\hat{z}\rangle\Big).\Big(\cos(\alpha)\; e^{-\delta^2(z-a_1)^2}\,|\hat{x}\rangle + \sin(\alpha)\; e^{-\delta^2(z-a_2)^2}\,|\hat{z}\rangle\Big) \\ &= \underbrace{\cos(\beta)\cos(\alpha)\; e^{-\delta^2(z-a_1)^2} + \sin(\beta)\sin(\alpha)\; e^{-\delta^2(z-a_2)^2}}_{\text{superposition of two gaussian peaked at } a_1 \text{ and } a_2}\end{aligned}$$

Claim: In the near-orthogonal limit, this expression reduces to Duck's result.

$$\langle q|\Psi\rangle = \cos(\alpha+\beta)\left\{\frac{1}{2}\Big[(1+\epsilon)\ e^{-\delta^2(z-a_1)^2} - (1-\epsilon)\ e^{-\delta^2(z-a_2)^2}\Big]\right\} \qquad ,\epsilon \ll 1$$

Proof

$$\begin{aligned}
\langle\psi_f|\Psi\rangle &= \cos(\beta)\cos(\alpha)e^{-\delta^2(z-a_1)^2} + \sin(\beta)\sin(\alpha)e^{-\delta^2(z-a_2)^2} \\
&= \frac{\cos(\alpha+\beta)}{2}\Big[\frac{2\cos(\alpha)\cos(\beta)}{\cos(\alpha+\beta)}e^{-\delta^2(z-a_1)^2} + \frac{2\sin(\alpha)\sin(\beta)}{\cos(\alpha+\beta)}e^{-\delta^2(z-a_2)^2}\Big] \\
&= \frac{\cos(\alpha+\beta)}{2}\left[\Big(1+\frac{\cos(\alpha-\beta)}{\cos(\alpha+\beta)}\Big)e^{-\delta^2(z-a_1)^2} - \Big(1-\frac{\cos(\alpha-\beta)}{\cos(\alpha+\beta)}\Big)e^{-\delta^2(z-a_2)^2}\right] \\
&= \langle q|\Psi\rangle = \cos(\alpha+\beta)\frac{1}{2}\Big[(1+\epsilon)e^{-\delta^2(z-a_1)^2} - (1-\epsilon)e^{-\delta^2(z-a_2)^2}\Big], \\
&\quad \text{for } \alpha=\pi/4,\ \beta=3\pi/4-\epsilon,\ \epsilon\ll 1
\end{aligned}$$

From our theoretical analysis, under the weak regime condition, this expression simplifies to a single Gaussian centered at the weak value. Remarkably, this weak value can lie far outside the range of the eigenvalues (for instance, in the spin case, a value of 100 is possible, whereas the maximum allowed eigenvalue is 1/2). Consequently, these tiny lateral shifts can be significantly amplified—though still smaller than the overall position spread—when operating in the weak regime with nearly orthogonal pre- and post-selected states. This amplification makes it possible to detect extremely small birefringence in the crystal.

4 Conclusion

In this report, we reviewed the theory of quantum weak measurement through the complementary perspectives of Aharonov, Albert, and Vaidman (AAV) and the refinement provided by Duck, Stevenson, and Sudarshan (SSD). Starting from the von Neumann measurement model, we showed how a weak system–pointer coupling, when combined with nearly orthogonal pre- and post-selections, produces the weak value—a quantity that can lie far outside the eigenvalue spectrum of the measured observable.

We then reviewed SSD's contribution, which removes the apparent paradox in AAV's original result by clearly showing when the first-order approximation is valid. The SSD approach is more general, and in the limit of very weak coupling it naturally gives back the AAV formula. Using both graphs and analytic calculations, we showed that the two treatments agree completely in the weak-measurement regime.

The Stern–Gerlach weak-measurement experiment was analyzed as a concrete physical realization of the theory. By studying the pointer's spatial response, we showed how spin-dependent beam separation in an inhomogeneous magnetic field naturally implements the weak coupling mechanism. The optical analog further illustrated that weak-value amplification is a general wave-interference effect not restricted to quantum particles.

Overall, our study highlights that weak measurements do not contradict standard quantum mechanics; instead, they provide a powerful interferometric tool for extracting subtle physical information with minimal system disturbance. Their applications—from precision metrology to foundational tests of quantum theory—continue to grow, showing that weak values remain an important concept in both theoretical and experimental quantum physics.

References

1. Y. Aharonov, D. Z. Albert, and L. Vaidman,
"How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100,"
Phys. Rev. Lett. **60**, 1351 (1988).
DOI: https://doi.org/10.1103/PhysRevLett.60.1351

2. I. M. Duck and P. M. Stevenson,
"The sense in which a 'weak measurement' of a spin-1/2 particle's spin component yields a value 100,"
Phys. Rev. D **40**, 2112 (1989).
DOI: https://doi.org/10.1103/PhysRevD.40.2112

3. J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd,
"Colloquium: Understanding Quantum Weak Values: Basics and Applications,"
Rev. Mod. Phys. **86**, 307 (2014).
DOI: https://doi.org/10.1103/RevModPhys.86.307

4. B. Zwiebach,
Quantum Physics II: Lecture Notes 2 – Spin One-Half, Bras, Kets, and Operators,
MIT OpenCourseWare, Fall 2013.
Available at: https://ocw.mit.edu/courses/8-05-quantum-physics-ii-fall-2013/resources/mit8_05f13_chap_02/

— The End —

Supplementary info.

1. When and why can we neglect $\hat{H}_0$ and $\frac{\hat{p}^2}{2m}$?

When and why can we neglect $\hat{H}_0$?

During the measurement time t_{meas}, the pointer's position evolves as

$$\hat{X}(t) = \hat{X}(0) + \frac{\hat{P}(0)}{m}t.$$

Because of the pointer's quantum nature, its position changes inherently.

If the pointer is very massive ($m \to$ large, classical limit), then

$$\hat{X}(t) \approx \hat{X}(0),$$

so there is no inherent change in position, and any shift corresponds only to the eigenvalue of the measured observable $\hat{M}$.

Even for a quantum pointer, this can be accounted for: during the measurement interval t_{meas}, the free evolution of the pointer is known, so subtracting this deterministic motion from the measured position isolates the contribution arising solely from the system observable $\hat{M}$.

However, there is a more fundamental issue that must be addressed. To see this, consider the time evolution of the pointer fluctuations:

$$\begin{aligned}
\Delta X(t) &= \hat{X}(t) - \langle \hat{X}(t) \rangle \\
&= \hat{X}(0) + \frac{t}{m}\hat{P}(0) - \langle \hat{X}(0) + \frac{t}{m}\hat{P}(0) \rangle \\
&= (\hat{X}(0) - \langle \hat{X}(0) \rangle) + \frac{t}{m}(\hat{P}(0) - \langle \hat{P}(0) \rangle) \\
&= \Delta X(0) + \frac{t}{m}\Delta P(0).
\end{aligned}$$

The variance is then

$$\begin{aligned}
\sigma_X^2(t) &= \langle \Delta X^2(t) \rangle \\
&= \left\langle \left(\Delta X(0) + \frac{t}{m}\Delta P(0) \right)^2 \right\rangle \\
&= \langle \Delta X^2(0) \rangle + \frac{t^2}{m^2}\langle \Delta P^2(0) \rangle + \frac{t}{m}\langle \Delta X(0)\Delta P(0) + \Delta P(0)\Delta X(0) \rangle.
\end{aligned}$$

Introducing the symmetric covariance,

$$\text{Cov}_s(\Delta X(0), \Delta P(0)) = \frac{1}{2}\langle \Delta X(0)\Delta P(0) + \Delta P(0)\Delta X(0) \rangle,$$

the variance can be rewritten as

$$\sigma_x^2(t) = \sigma_x^2(0) + \frac{t^2}{m^2}\sigma_p^2(0) + \frac{2t}{m}\text{Cov}_s(\Delta X(0), \Delta P(0)).$$

Note: For a symmetric, real Gaussian initial pointer state, the X–P correlation vanishes, so

$$\text{Cov}_s(\Delta X(0), \Delta P(0)) = 0.$$

Hence, the variance simplifies to

$$\sigma_x^2(t) = \sigma_x^2(0) + \frac{t^2}{m^2}\sigma_p^2(0), \qquad \sigma_x(t) = \sqrt{\sigma_x^2(0) + \frac{t^2}{m^2}\sigma_p^2(0)}.$$

If at $t = 0$, the standard deviation of the Gaussian wavepacket is $\sigma_x(0)$, then it increases with time. During the measurement time t_{meas}, the wavepacket spreads, and the resulting δX_{spread} can become comparable to or larger than the shift caused by the coupling to the measured observable $\hat{M}$.

Conclusion

To resolve the eigenvalue of $\hat{M}$, one must either take $m \to \infty$ (classical pointer) or $t_{\text{meas}} \to 0$ (short measurement time) to make the inherent spreading negligible.

When and why can we neglect $\frac{\hat{p}^2}{2m}$?

Case I: $[\hat{M}, \hat{H}_0] = 0$ **(commuting)**

For the observable $\hat{M}$,

$$\frac{d\hat{M}}{dt} = -\frac{i}{\hbar}[\hat{M}, \hat{H}_0] + \frac{\partial \hat{M}}{\partial t}.$$

Assuming $\hat{M}$ is not explicitly time-dependent and $[\hat{M}, \hat{H}_0] = 0$, we have

$$\frac{d\hat{M}}{dt} = 0 \quad \Rightarrow \quad \hat{M} \text{ is time-independent.}$$

⇒ During the measurement, the eigenvectors of $\hat{M}$ do not change with time, and thus are perfectly correlated with the pointer's position $\hat{X}$ (von Neumann measurement idea).

Neglecting $\hat{H}_0$: The total time-evolution operator is

$$U(t) = e^{-\frac{i}{\hbar}\hat{H}t} = e^{-\frac{i}{\hbar}(\hat{H}_0 + \lambda \hat{M}\hat{P})t}.$$

Since $[\hat{M}, \hat{H}_0] = 0$ and $[\hat{P}, \hat{H}_0] = 0$, we can factorize:

$$U(t) = e^{-\frac{i}{\hbar}\hat{H}_0 t} \cdot e^{-\frac{i}{\hbar}\lambda \hat{M}\hat{P}t}.$$

The first term represents the free evolution of the system, while the second term generates the entanglement between the pointer and the eigenvalues of $\hat{M}$. Since the free evolution does not affect the measurement correlation, we can neglect $\hat{H}_0$.

—

Case II: $[\hat{M}, \hat{H}_0] \neq 0$ **(non-commuting)**

If the measurement is very fast, i.e., the measurement time t_{meas} is short, then even though $\hat{M}$'s eigenvectors change with time, the change is negligible during t_{meas}:

$$t_{\text{meas}} \ll \text{timescale of evolution due to } \hat{H}_0.$$

Conclusioon

⇒ Fast measurement $\implies$ we can approximately treat $[\hat{M}, \hat{H}_0] \approx 0$, and neglect $\hat{H}_0$.

---

2. Coupling Hamiltonian for the Stern–Gerlach Measurement

From our theory of von Neumann–type measurements, the interaction Hamiltonian takes the generic form

$$\hat{H}_{\text{int}} = -\, g(t)\, \hat{q}\, \hat{A},$$

where $\hat{A}$ is the system observable and $\hat{q}$ is the pointer variable. To identify the corresponding coupling in the **Stern–Gerlach measurement**, we start from the physical interaction energy between a magnetic moment and an inhomogeneous magnetic field.

Magnetic moment of the electron

The magnetic dipole moment operator is

$$\vec{\mu} = g\, \mu_B\, \frac{\hat{\vec{S}}}{\hbar}$$

where $g \simeq 2$ for the electron and μ_B is the Bohr magneton. Using $\hat{\vec{S}} = \frac{\hbar}{2}\hat{\vec{\sigma}}$, we obtain

$$\vec{\mu} = \mu_B\, \hat{\vec{\sigma}} \qquad \Longrightarrow \qquad \mu_z = \mu_B\, \hat{\sigma}_z.$$

Expansion of the magnetic field

In the Stern–Gerlach device, the magnetic field varies along the z–axis. For a sufficiently weak and slowly varying gradient:

$$B_z(z) \simeq B_0 + \left(\frac{\partial B_z}{\partial z}\right) z,$$

which is simply the first-order Taylor expansion around $z = 0$.

Interaction Hamiltonian

The potential energy of a magnetic moment in a magnetic field is

$$\hat{H} = -\vec{\mu} \cdot \vec{B} = -\mu_z \, B_z(z).$$

Substituting $\mu_z = \mu_B \hat{\sigma}_z$ and the expanded field:

$$\hat{H} = -\mu_B \, \hat{\sigma}_z \left(B_0 + \frac{\partial B_z}{\partial z} \, \hat{z} \right).$$

The first term

$$-\mu_B \hat{\sigma}_z B_0$$

produces only a constant spin-dependent energy shift and does *not* cause spatial separation of the beams. Therefore, the relevant coupling term for the SG measurement is

$$\hat{H}_{\text{coupling}} = -\mu_B \, \hat{\sigma}_z \left(\frac{\partial B_z}{\partial z} \right) \hat{z}.$$

Including a finite measurement duration

Because the SG interaction acts only while the atom is inside the magnet, we introduce a time-window function $g(t)$ that is nonzero only during the measurement:

$$\boxed{\hat{H}_{\text{int}}(t) = -\, g(t) \, \mu_B \left(\frac{\partial B_z}{\partial z} \right) \hat{z} \, \hat{\sigma}_z}$$

This has the exact von Neumann structure

Code

```
"""
Code for a particular value of epsilon and delta
AAV vs SSD pointer functions
"""

# ================================
# Import packages
# ================================
import numpy as np
import matplotlib.pyplot as plt


# ================================
# Parameters
# ================================
lembda = 1                      # constant, mu*(/z (B_z))
epsilen = 0.01
alpha = np.pi - 2 * epsilen     # angle between pre- and post-selected spin states
delta =  0.01                  # spread in position of pointer

a = 1000                        # max range on x axis
c = 80000                       # number of grid points on x axis
b = 200                           # x-axis limits


# ================================
# Function definitions
# ================================

# AAV's pointer function
def phi_aav_func(alpha, delta, lembda, p):
    return np.cos(alpha / 2) * np.exp((-delta**2) * (p - lembda * np.tan(alpha / 2))**2)


# SSD's pointer function
def phi_ssd_func(alpha, delta, lembda, p):
    c1 = 0.5 * (np.cos(alpha / 2) + np.sin(alpha / 2))
    c2 = 0.5 * (np.cos(alpha / 2) - np.sin(alpha / 2))
    return c1 * np.exp((-delta**2) * (p - lembda)**2) + c2 * np.exp((-delta**2) * (p + lembda)**2)


# ================================
# Variables
# ================================
p = np.linspace(-a, +a, c)

# Arrays
phi_aav = np.zeros(len(p))
phi_ssd = np.zeros(len(p))

# Compute values
for i in range(len(p)):
    phi_aav[i] = phi_aav_func(alpha, delta, lembda, p[i])
    phi_ssd[i] = phi_ssd_func(alpha, delta, lembda, p[i])


# ================================
# Find peaks
# ================================
idx_aav = np.argmax(phi_aav)
idx_ssd = np.argmax(phi_ssd)

peak_aav = (p[idx_aav], phi_aav[idx_aav])
peak_ssd = (p[idx_ssd], phi_ssd[idx_ssd])


# ================================
# Plot: AAV vs SSD
# ================================
plt.figure(figsize=(8, 5))
#plt.plot(p, phi_aav, label='phi_aav')
plt.plot(p, phi_ssd, label='phi_ssd')
plt.xlabel('p_z')
```

```
plt.ylabel('Amplitude')

# Mark peaks
#plt.scatter(*peak_aav, color='red', zorder=5)
#plt.annotate(f"({peak_aav[0]:.2f}, {peak_aav[1]:.3f})", xy=peak_aav)
plt.scatter(*peak_ssd, color='green', zorder=5)
plt.annotate(f"({peak_ssd[0]:.2f}, {peak_ssd[1]:.3f})", xy=peak_ssd)

# Always show axes at origin
plt.axhline(0, color="black", linewidth=0.6)
plt.axvline(0, color="black", linewidth=0.6)

plt.grid()
plt.legend()
plt.xlim(-b, b)
plt.show()


# ================================
# Weak regime check
# ================================
if delta < epsilen:
    print('weak regime')
else:
    print('strong regime')


# ================================
# Probability plot
# ================================
phi_aav_square = phi_aav**2
phi_ssd_square = phi_ssd**2

plt.figure(figsize=(8, 5))
# plt.plot(p, phi_aav_square, label='phi_aav_probability')
plt.plot(p, phi_ssd_square, label='phi_ssd_probability')
plt.xlabel('p_z')
plt.ylabel('Amplitude square')
plt.xlim(-b, b)

# Always show axes at origin
plt.axhline(0, color="black", linewidth=0.6)
plt.axvline(0, color="black", linewidth=0.6)

plt.grid()
plt.legend()
plt.show()
```