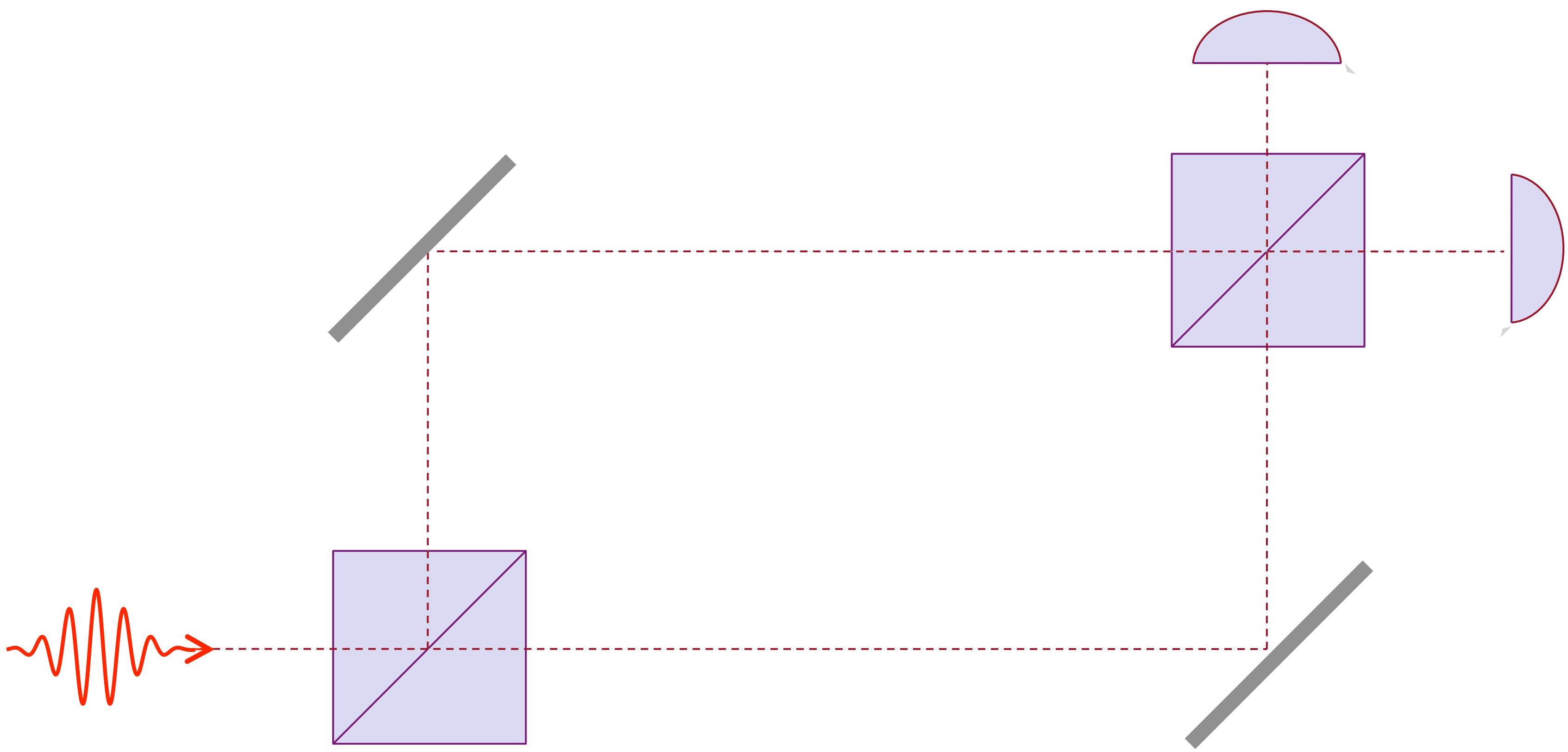
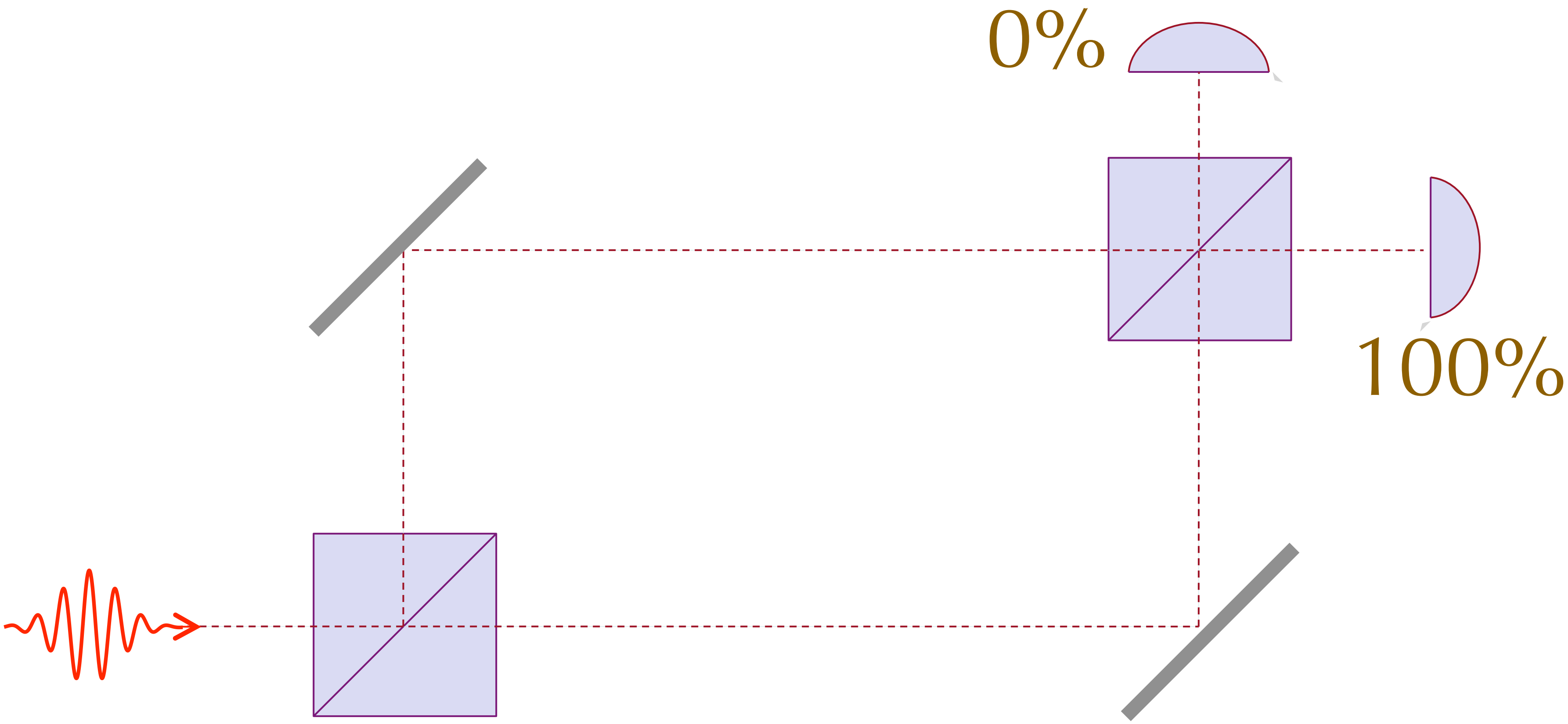
Why does nature like complex probability amplitudes?

William K. Wootters
Williams College

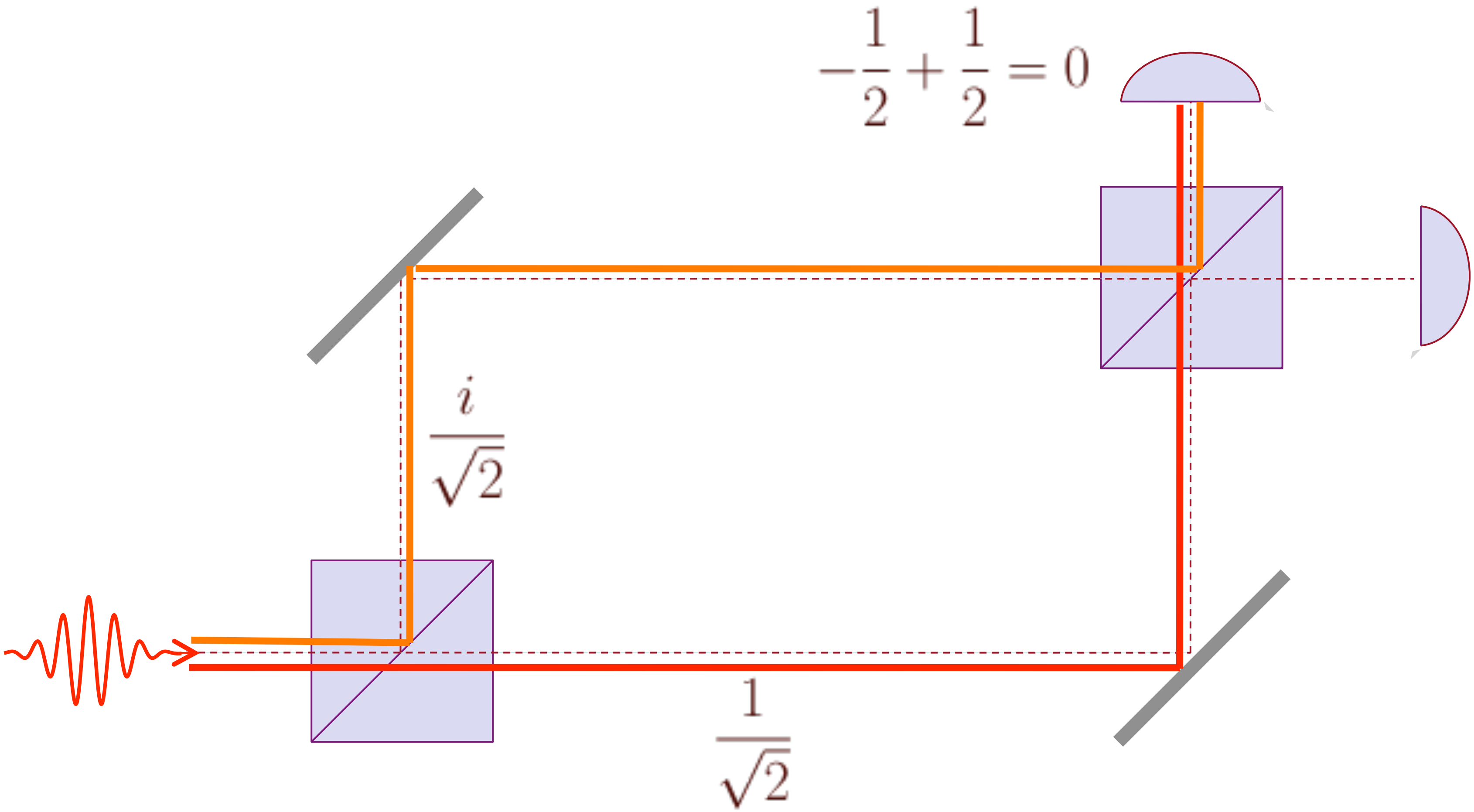
A simple quantum experiment



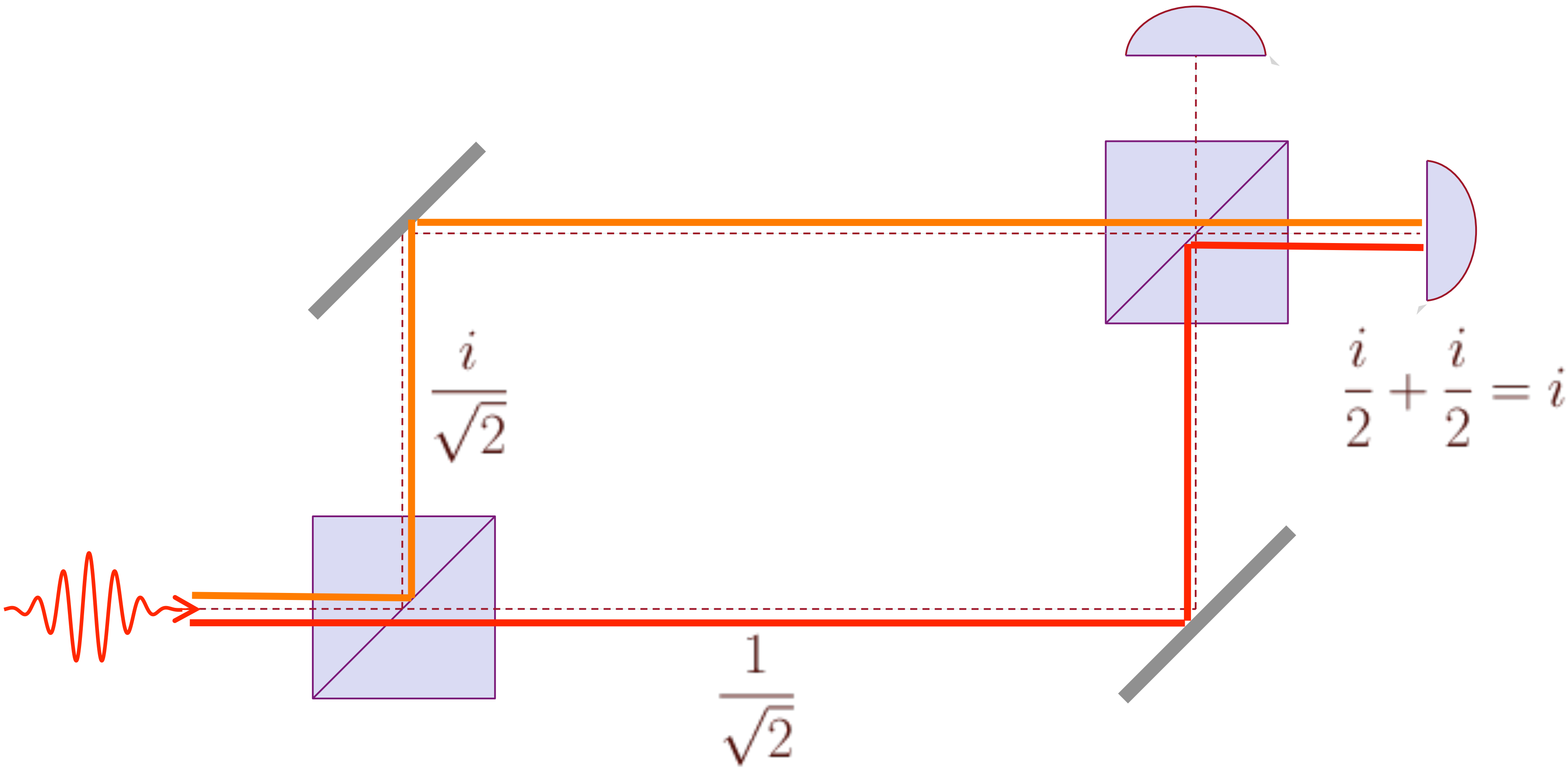
The experimental result



A quantum explanation of this result



A quantum explanation of this result



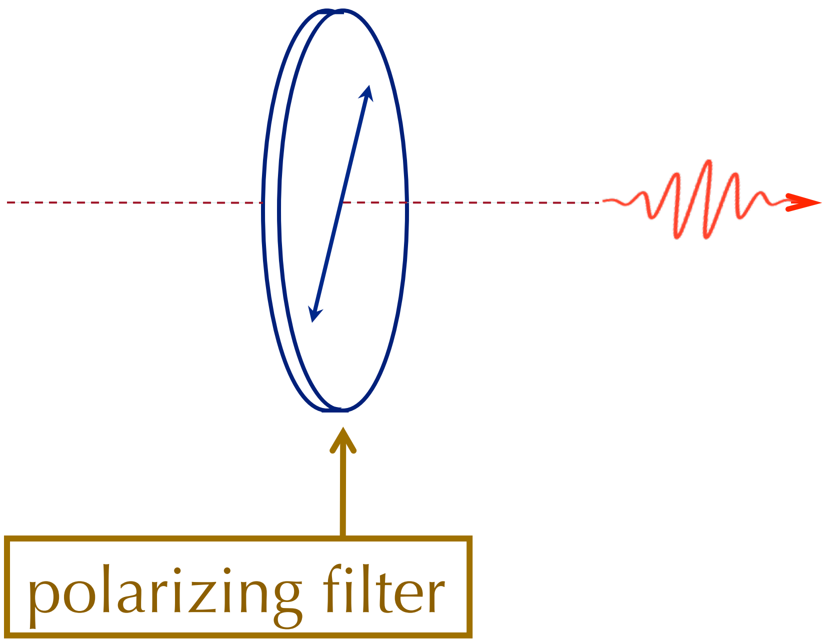
Questions

(i) Why do we have to work with "square roots" of probability? Is there a deeper explanation?

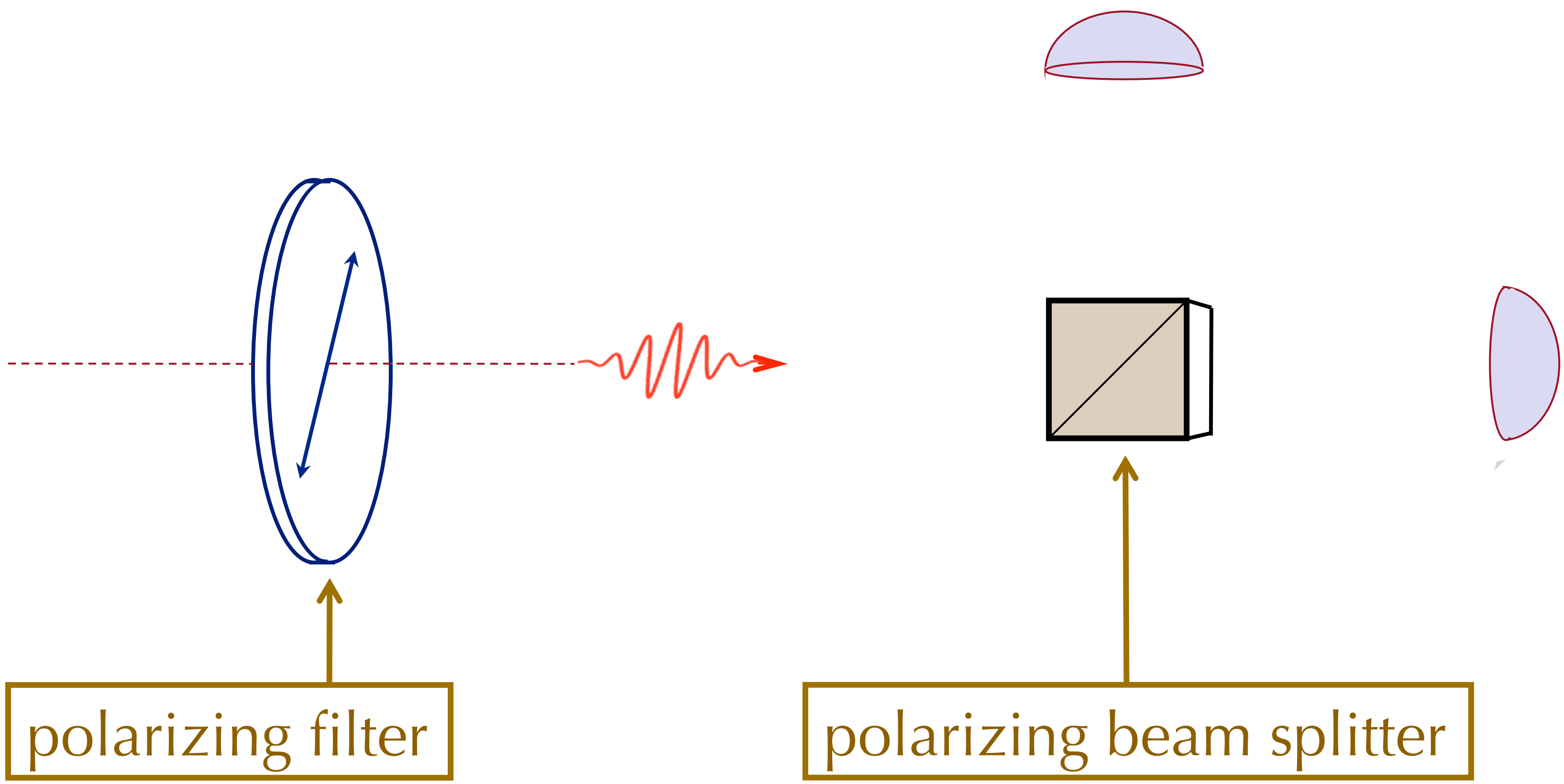
(ii) And why are these "square roots" complex?

I will try to answer the first question—why square roots? But my answer will make the second question more puzzling.

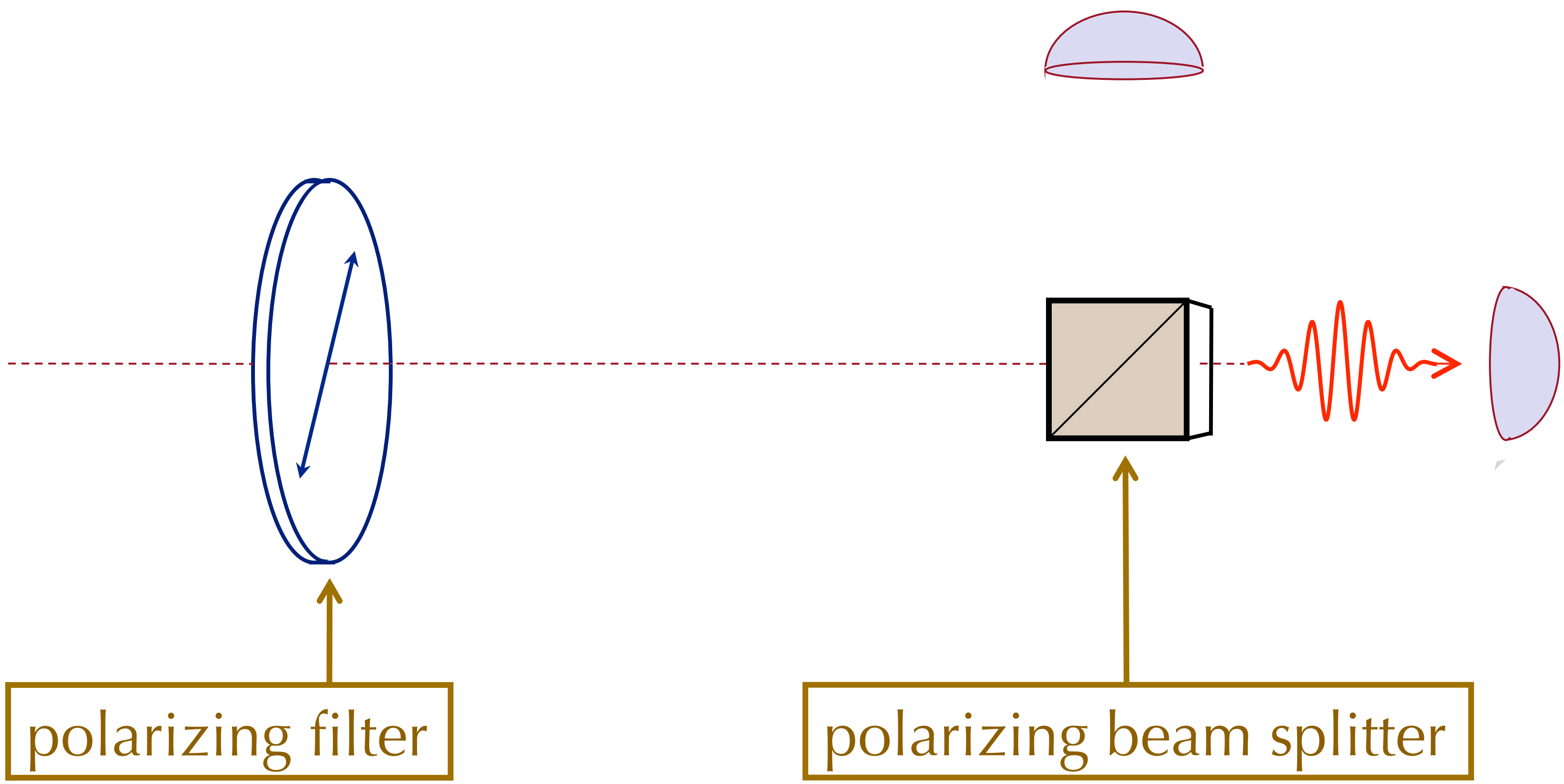
Photon polarization



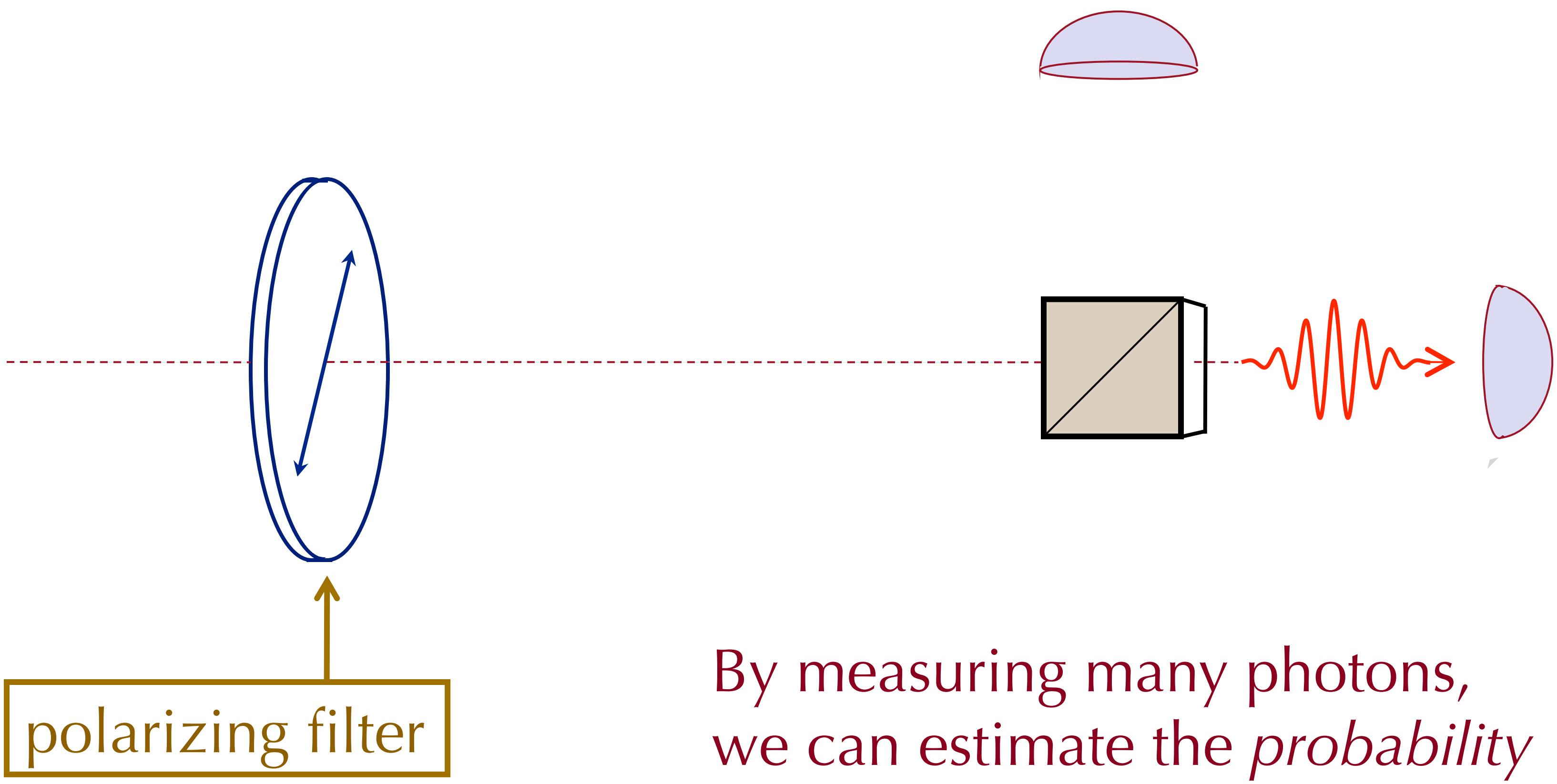
Measuring photon polarization



Measuring photon polarization



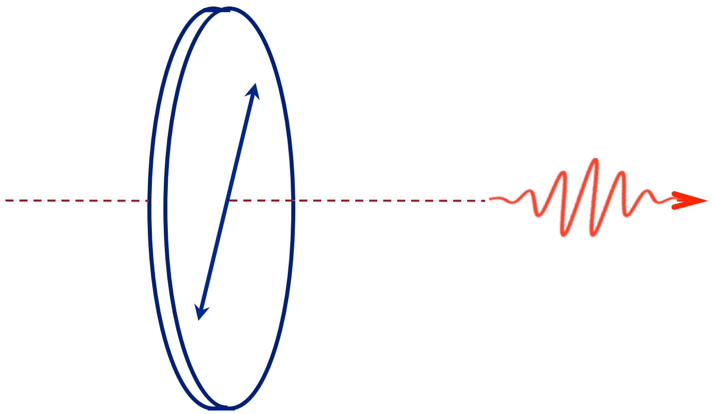
Measuring photon polarization

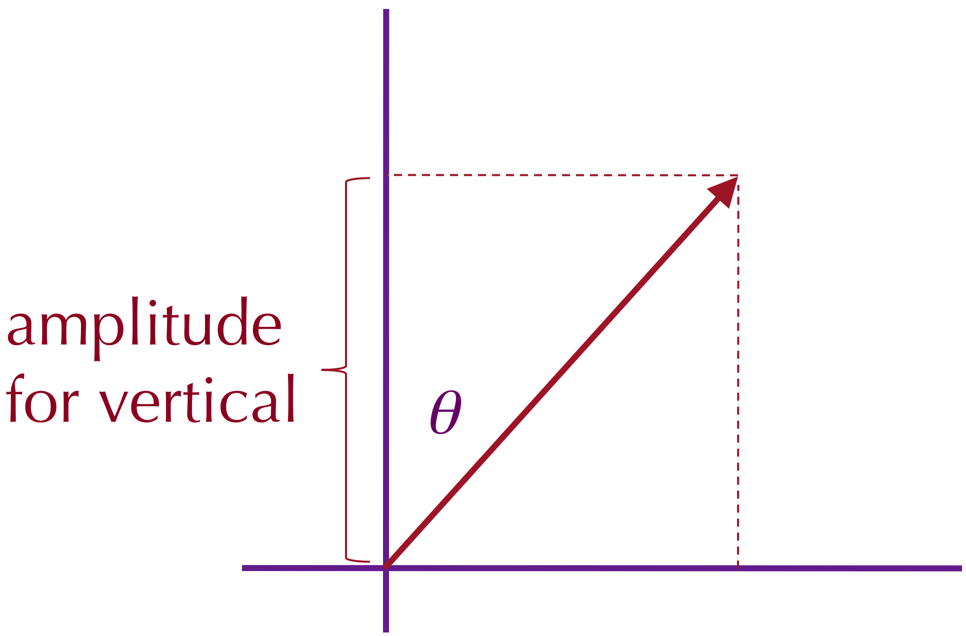


By measuring many photons,
we can estimate the *probability*
of the vertical outcome.

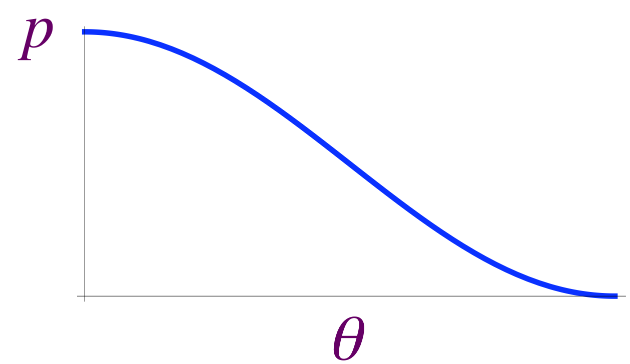
This tells us about the angle.

The standard account of probability *vs* angle

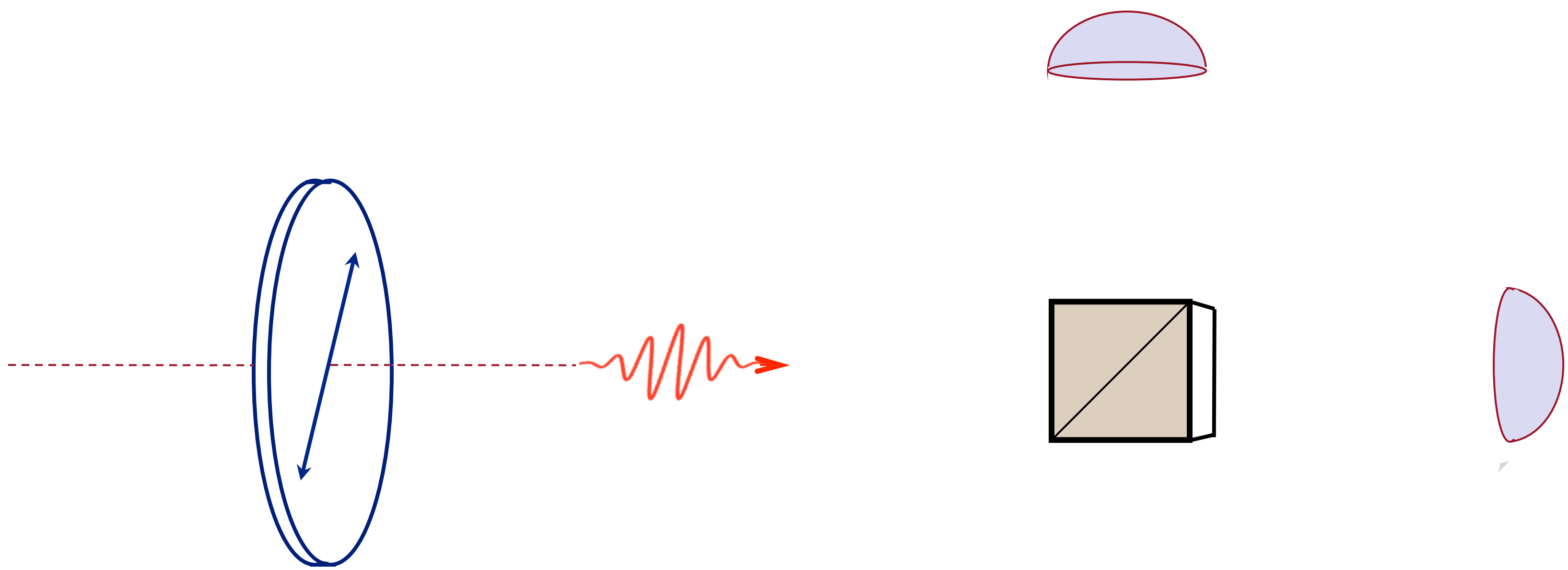




Squaring the amplitude gives the probability:



A completely different explanation for that curve: Optimal information transfer?



The angle varies continuously.
But the measurement is probabilistic with only two possible outcomes.

Is the communication optimal?

A Communication Puzzle

Alice is going to think of a number θ between 0 and $\pi/2$.

A Communication Puzzle

Alice is going to think of a number θ between 0 and $\pi/2$.

She will construct a coin, with her number encoded in the probability of heads.

A Communication Puzzle

Alice is going to think of a number θ between 0 and $\pi/2$.

She will construct a coin, with her number encoded in the probability of heads. She will send the coin to Bob.

A Communication Puzzle

Alice is going to think of a number θ between 0 and $\pi/2$.

She will construct a coin, with her number encoded in the probability of heads. She will send the coin to Bob.

To find θ, Bob will flip the coin...

A Communication Puzzle

Alice is going to think of a number θ between 0 and $\pi/2$.

She will construct a coin, with her number encoded in the probability of heads. She will send the coin to Bob.

To find θ, Bob will flip the coin, but it self-destructs after one flip.

The Goal: Find the optimal encoding $p(\theta)$

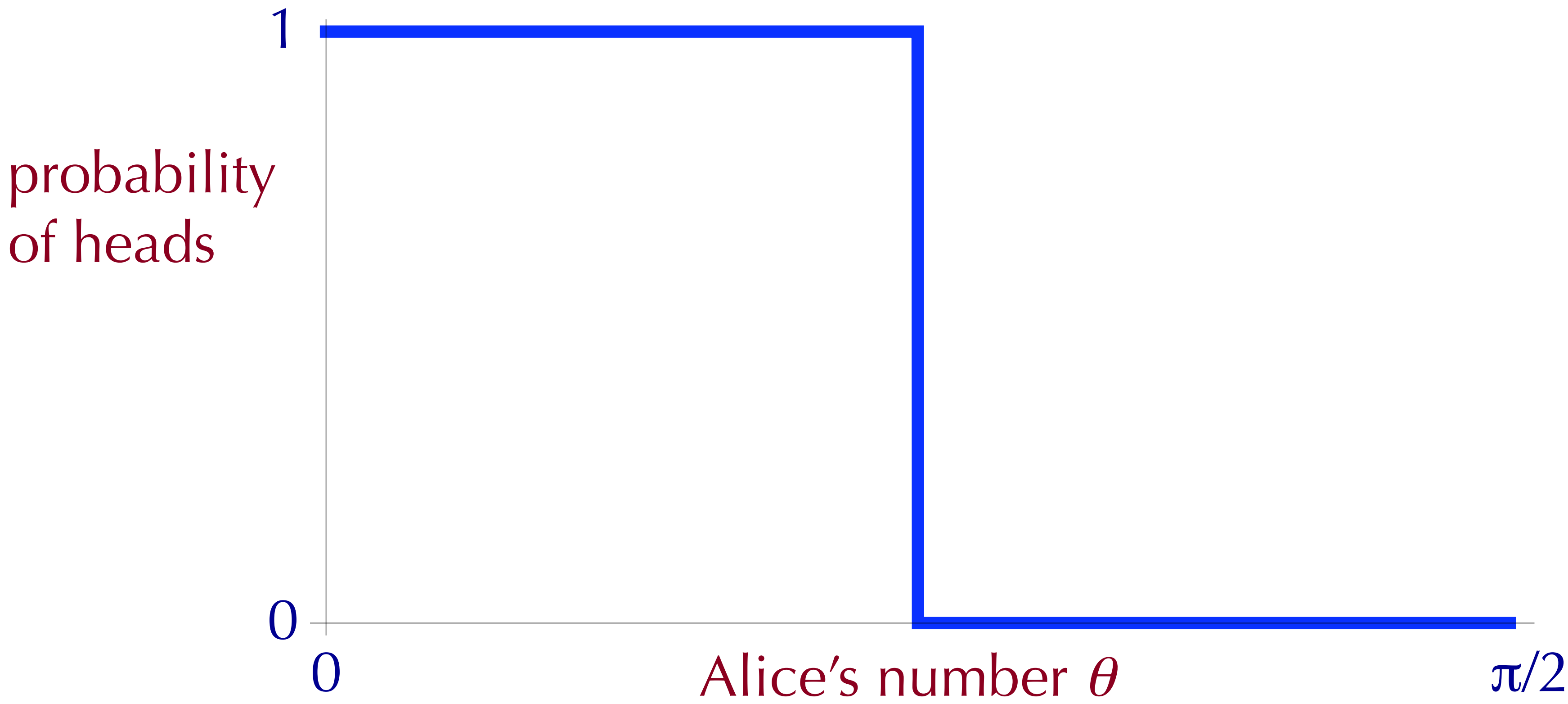
Maximize the mutual information:

$$I(\theta : n) = \left[-\int P(\theta) \log P(\theta) d\theta \right] - \left\langle -\int P(\theta|n) \log P(\theta|n) d\theta \right\rangle_n$$

Here n is the number of heads Bob sees (n = 0 or 1),
and θ is distributed *uniformly* between 0 and π/2.

An Optimal Encoding (1 flip)

(Information-maximizing for a uniform *a priori* distribution.)

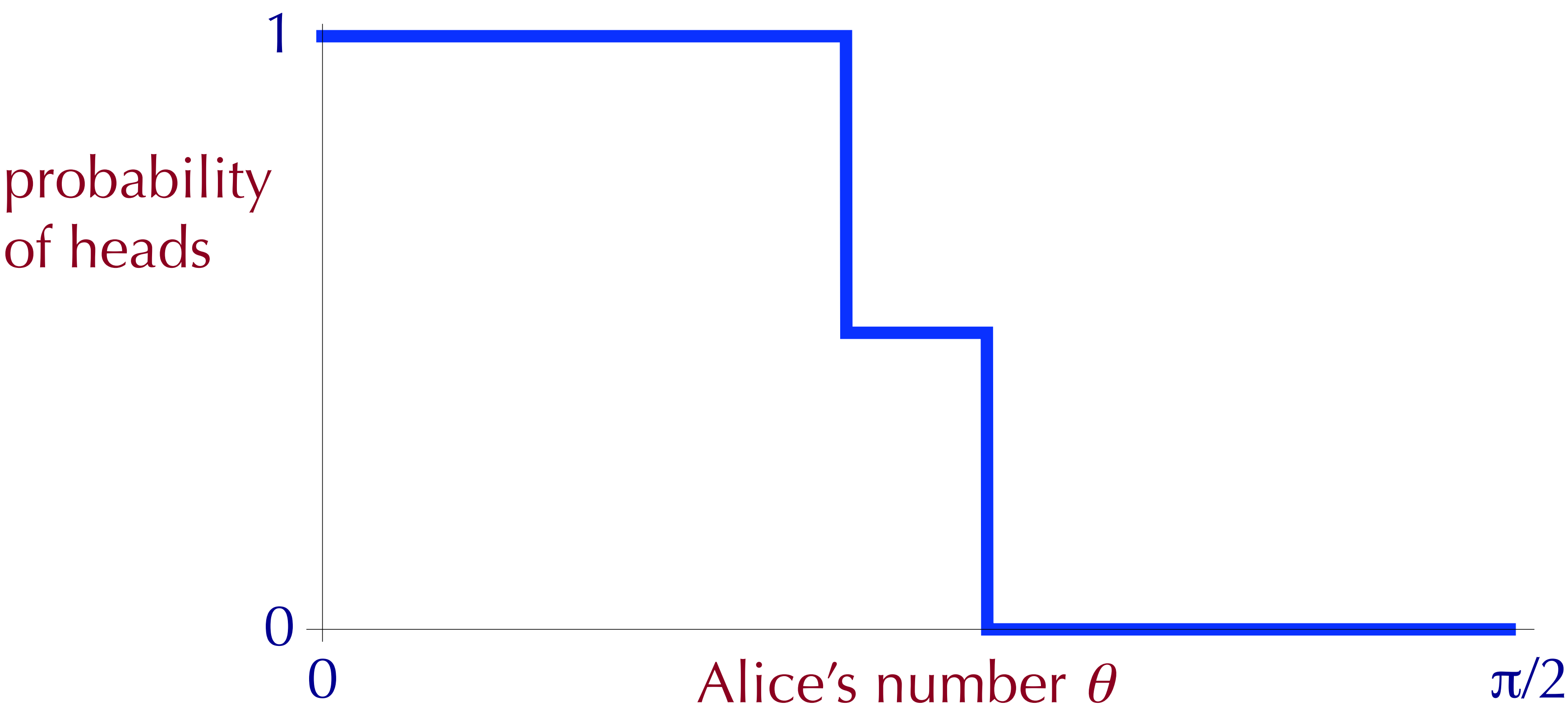


Modified Puzzle—Bob Gets ***Two*** Flips

The coin self-destructs after ***two*** flips.

(It's like sending two photons with the same polarization.)

An Optimal Encoding (2 flips)

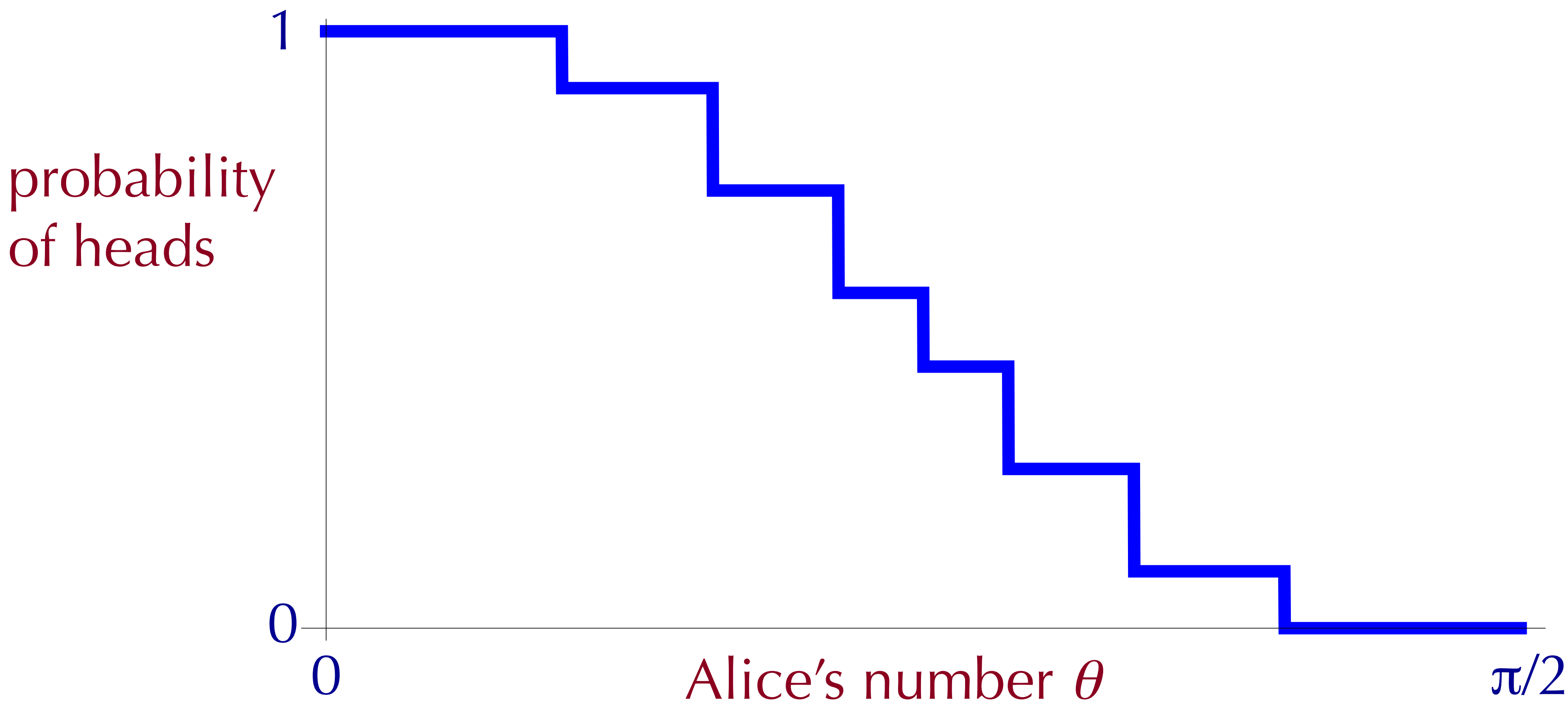


New Modification—Bob Gets ***25*** Flips

The coin self-destructs after ***25*** flips.

(It's like sending 25 photons with the same polarization.)

An Optimal Encoding (25 flips)



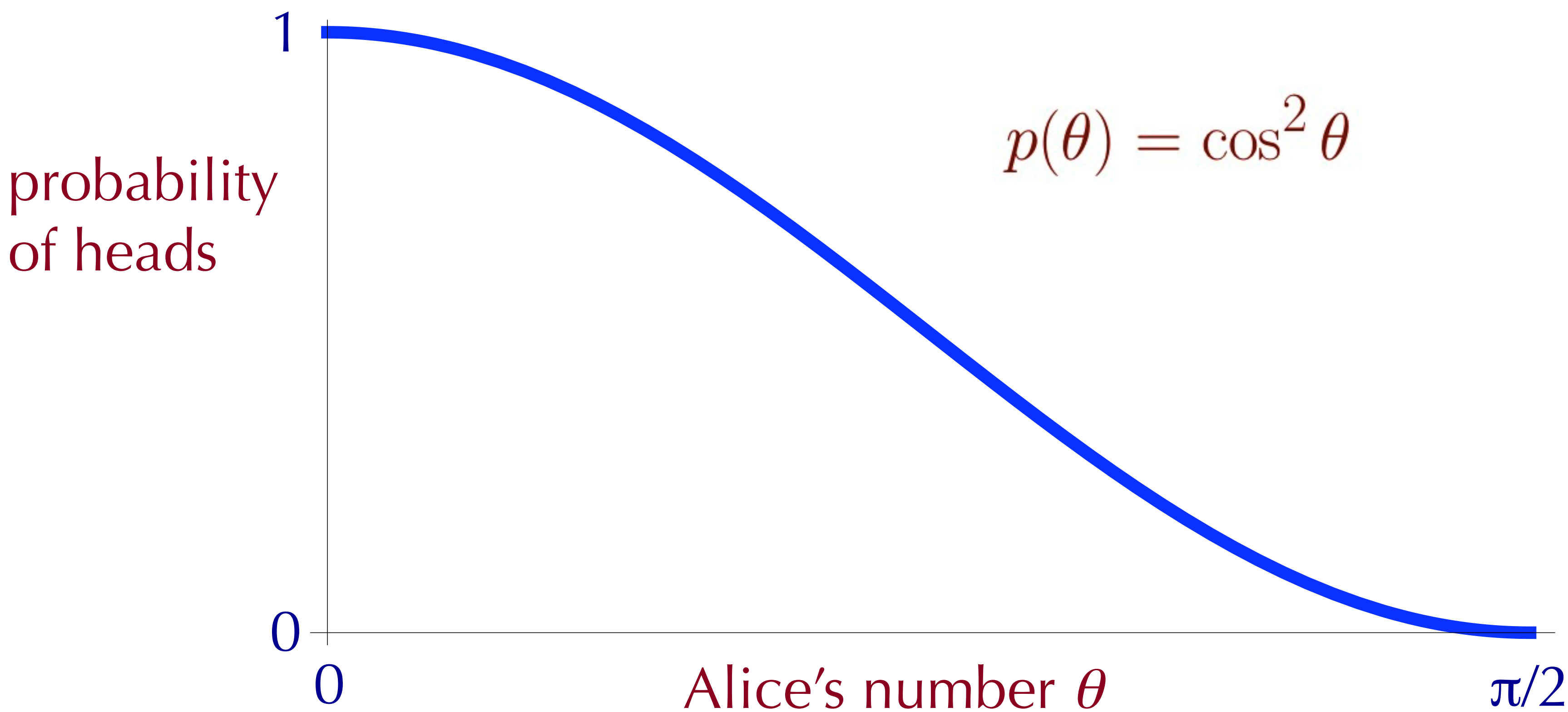
Taking the limit of an infinite number of flips

For any given encoding $p_{\text{heads}}(\theta)$, consider the following limit.

$$\lim_{N \to \infty} \left[I(n : \theta) - \log \sqrt{N} \right]$$

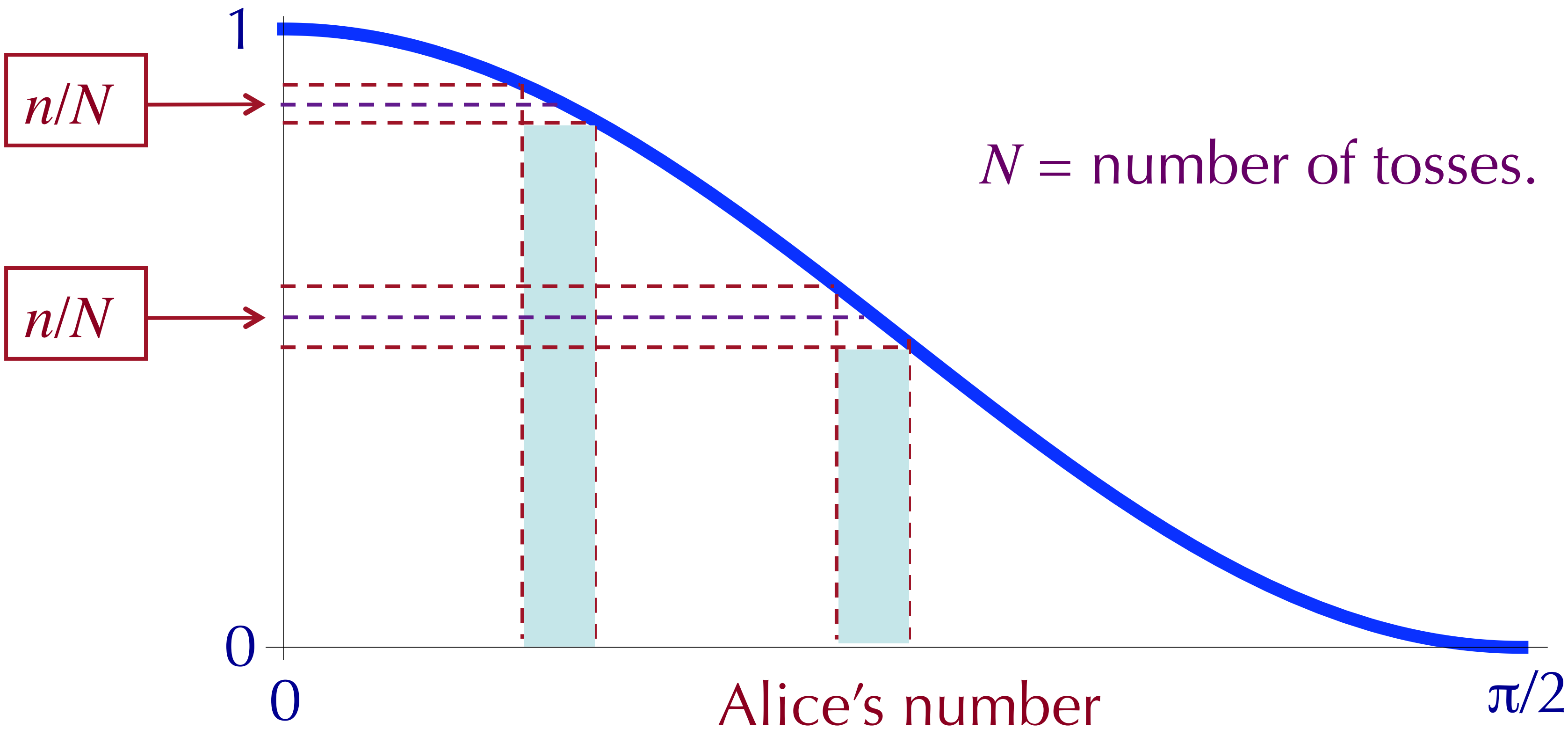
We ask what encodings maximize this limit.

An optimal encoding in the limit of infinitely many flips



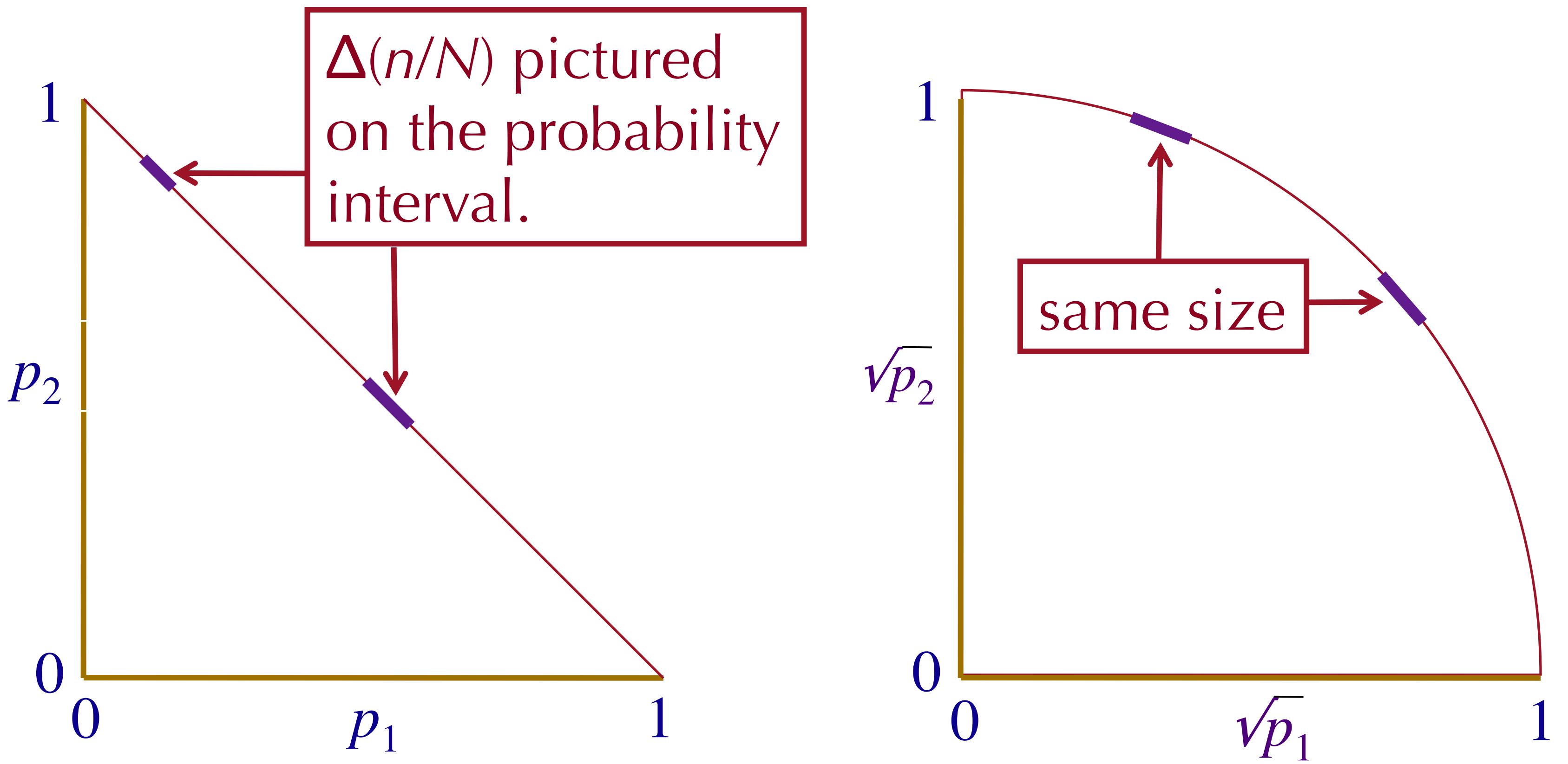
This is exactly what photons do!

Why this works: Wider deviation matches greater slope



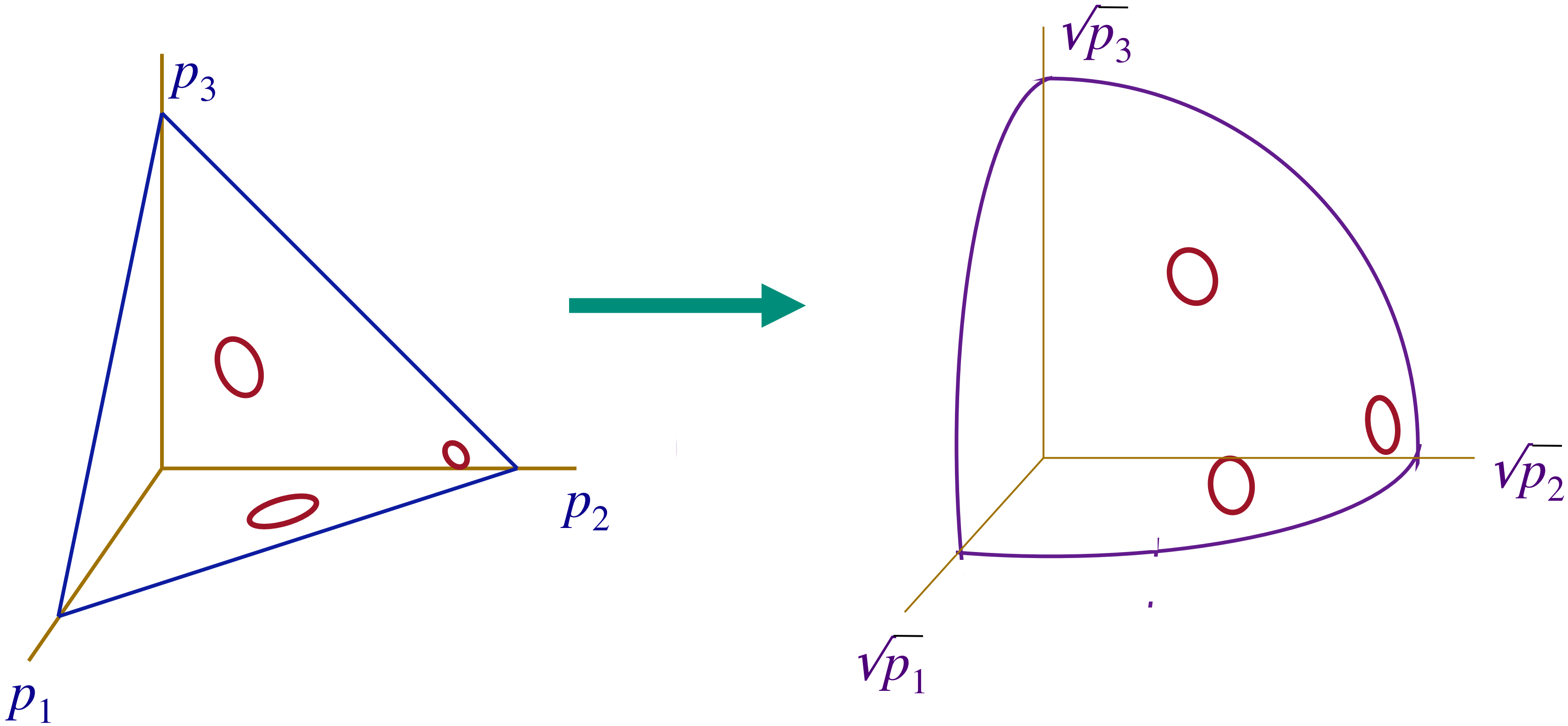
$$\Delta\left(\frac{n}{N}\right) = \sqrt{\frac{p(1-p)}{N}} \qquad \left|\frac{dp}{d\theta}\right| = 2\sqrt{p(1-p)}$$

Another way of seeing the same thing



Using square roots of probability equalizes the spread in the binomial distribution.

Same effect when there are more than two possible outcomes



In this sense square roots of probability arise naturally.

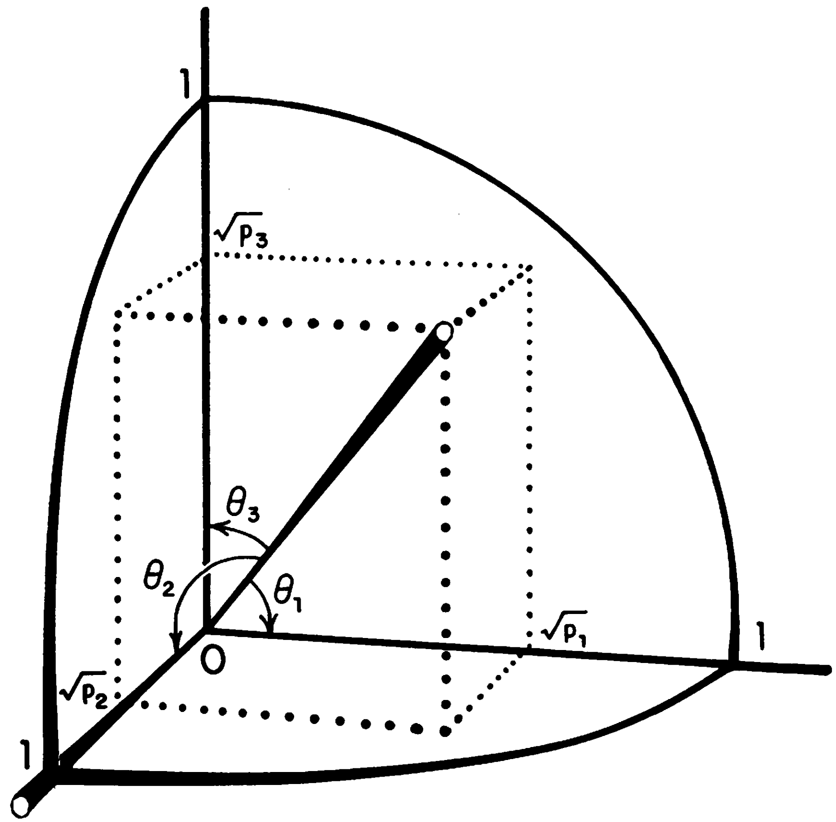


FIG. 4. Representation of a population, with gene frequencies p_1, p_2, p_3 at a single tri-allelic locus, on the octant of a sphere.

From Am. J. Human Genetics (1967).

A Good Story

In quantum theory, it's impossible to have a perfect correspondence between past and future (in measurement).

But the correspondence is as close as possible—information transfer is optimal—given the limitations of a probabilistic theory with discrete outcomes.

This fact might begin to explain why we have to use "square roots of probability."

But this good story is not true!

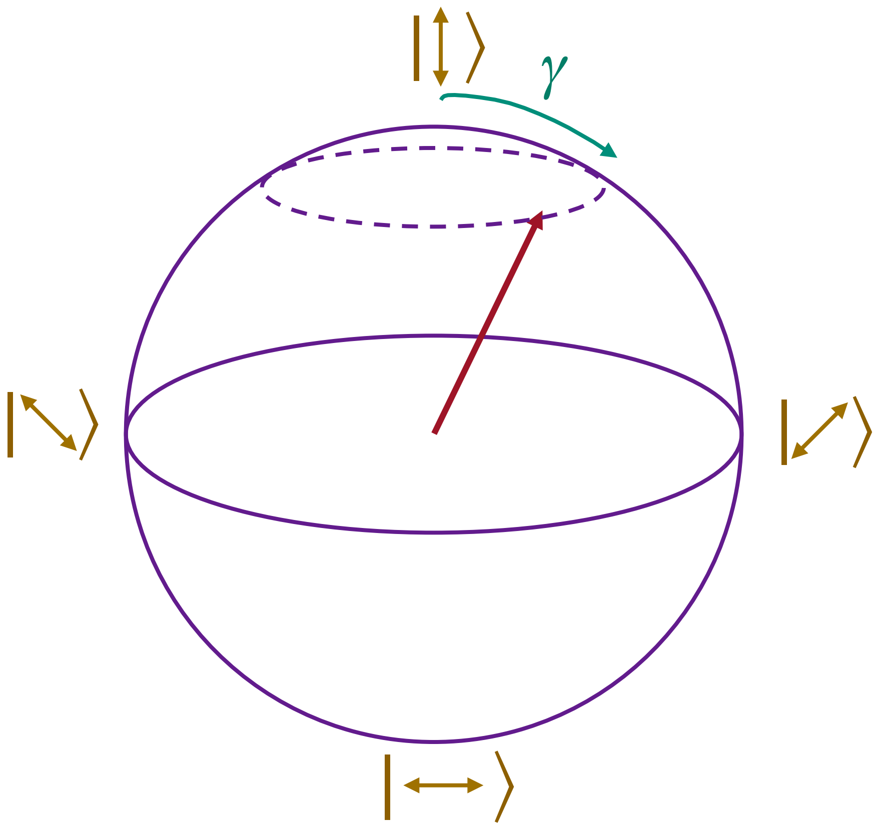
Why not?

But this good story is not true!

Why not?

Because probability amplitudes are complex.

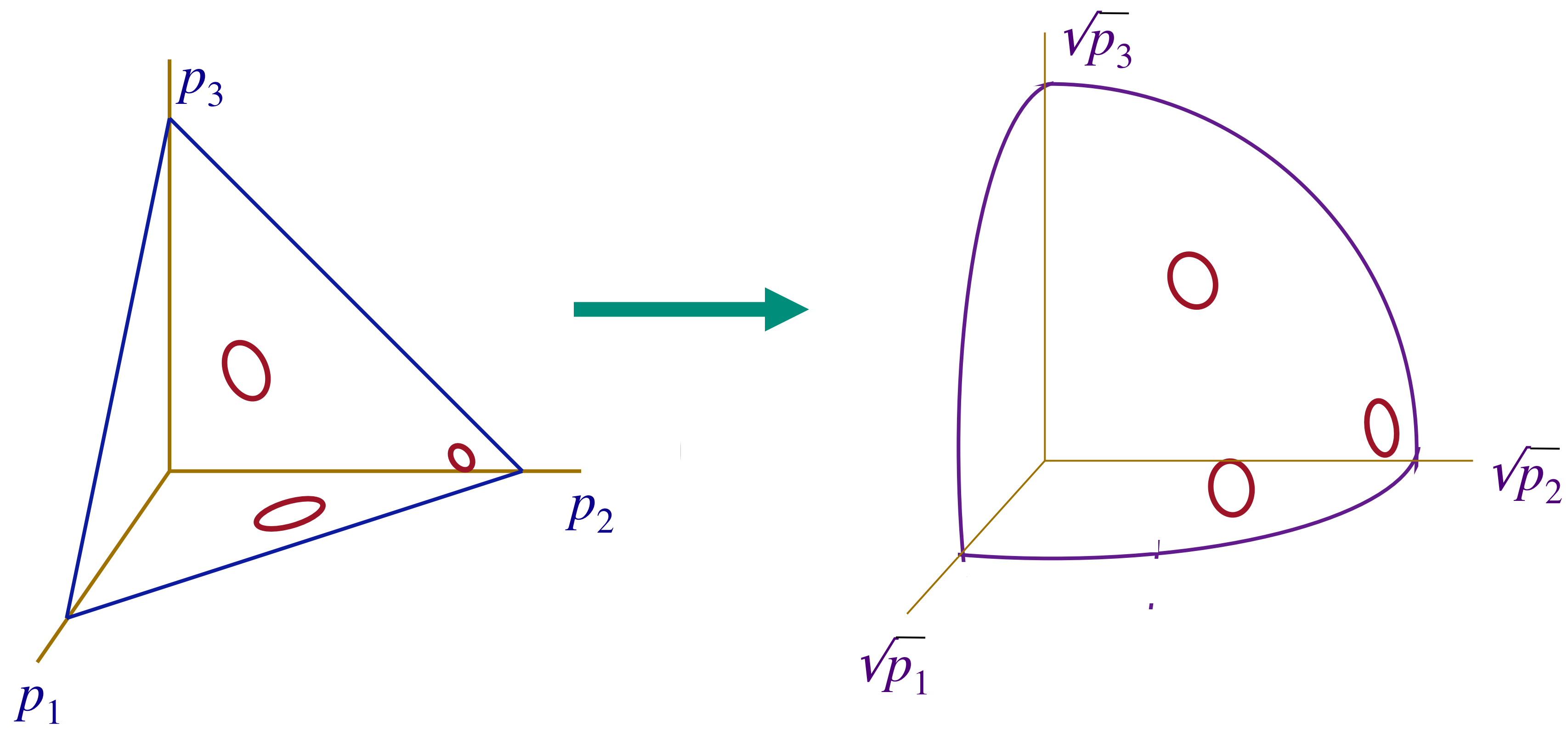
No information maximization in the *complex* theory.



An orthogonal measurement completely misses a whole degree of freedom (phase).

$p_{\text{vertical}} = \cos^2(\gamma/2)$,

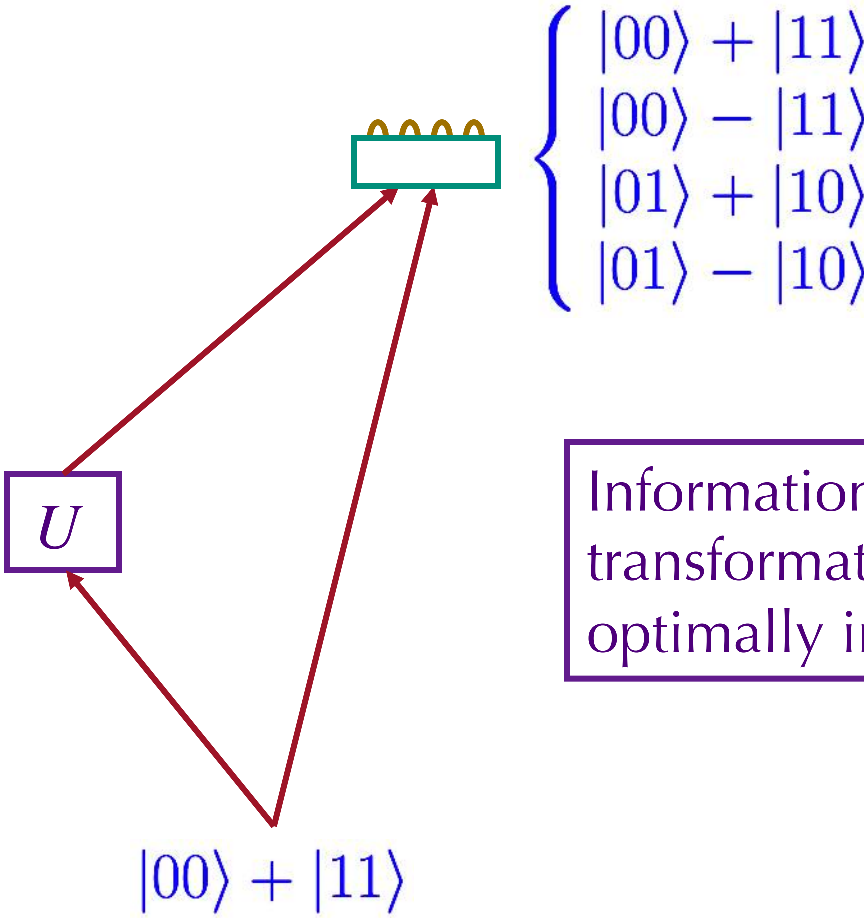
but γ is not uniformly distributed.



This picture does not suggest ***complex*** square roots of probability.

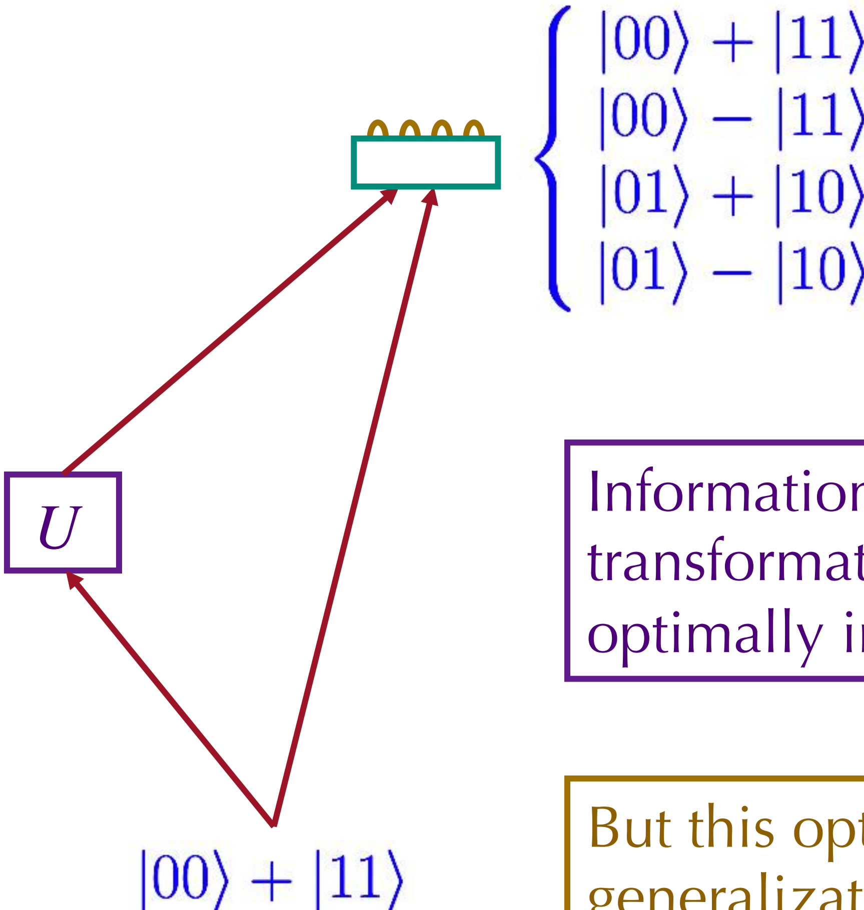
Very generally, information is transferred optimally (in our sense) in the ***real-amplitude*** variant of quantum theory but not in standard (complex) quantum theory.

Information about a unitary transformation?



Information about the special unitary transformation *U* ***is*** expressed optimally in the outcomes.

Information about a unitary transformation?



Information about the special unitary transformation *U* ***is*** expressed optimally in the outcomes.

But this optimization has no obvious generalization to higher dimension.

Conceivable answers to "Why complex amplitudes?"

- Want an uncertainty principle (Stueckelberg)
- Want "local tomography" (Hardy; Chiribella *et al;* Müller & Masanes *et al*; Dakić & Brukner; me)
- Want complementarity (Goyal)
- Want algebraic closure (many people)
- *Maybe* there's a ubit (Aleksandrova *et al*)