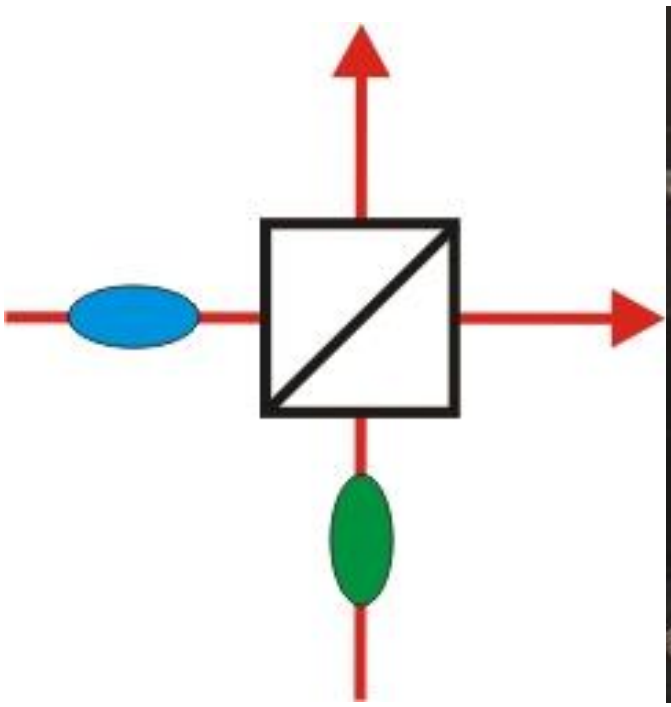
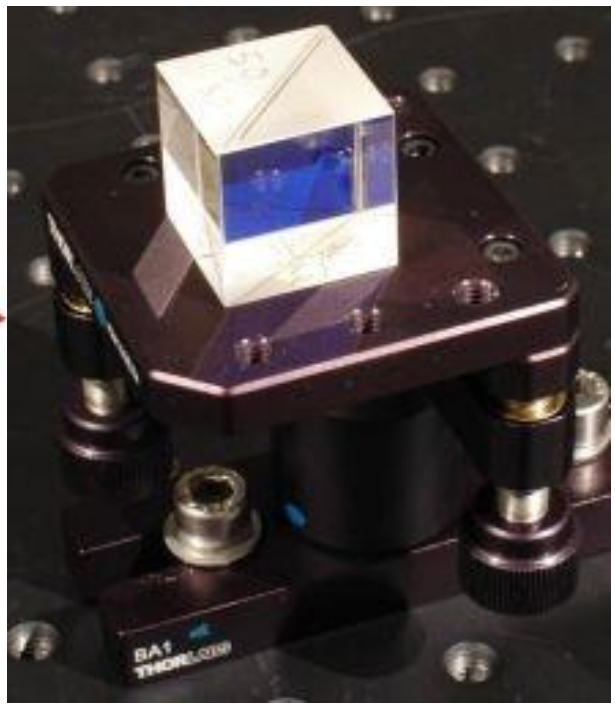
4.14 Zweiphotoneninterferenz

Versuchsanleitung zum Fortgeschrittenenpraktikum
Abteilung A
Version 1.5

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Physik
Institut für Angewandte Physik
AG Nichtlineare Optik / Quantenoptik





Inhaltsverzeichnis

Versuchsanleitung 3

Anhang

The Hong-Ou-Mandel interferometer in the undergraduate laboratory 14

Measurement of Subpicosecond Time Intervals between Two Photons by Interference 19

Notes on Hong-Ou-Mandel interference 22

Observation of a “quantum eraser”: A revival of coherence in a two-photon interference experiment 28

Introductory Quantum Optics 38

Advanced Quantum Mechanics 67

Optical Waveguides and Fibers 77

Fiber Optics: How Fused Fiber Optic Couplers Work 82

Two-photon Interferometer 84

Single-photon generation and detection 86

Entanglement Demonstrator 95

Understanding Correlated-Photon Based Efficiency Calibration of Photon Counting Detectors 101

Vorbereitung

- **Quantenphysikalische Beschreibung:** Fockzustände, Bosonen und Fermionen, Erzeugungs– und Vernichtungsoperatoren, Strahlteiler
- **Grundlagen der Optik:** Interferenzfilter, Doppelbrechung, Wellenplatten
- **Nichtlineare Frequenzkonversion:** Spontane parametrische Fluoreszenz, Phasenanpassung
- **Zweiphotoneninterferenz:** Zeitliche Kohärenz, Unterscheidbarkeit von Photonen, Mach-Zehnder-Interferometer
- **Einzelphotonenmessung:** Lawinenphotodioden, Koinzidenzmessungen
- **Glasfasertechniken:** Ein- und Multimodenfasern, polarisationserhaltende Fasern, Strahlteilerfaser

Setzen Sie sich mit der im Anhang dieser Anleitung gegeben Literatur auseinander und überlegen Sie sich **vor** Versuchsbeginn welche Größen gemessen werden müssen. **Erstellen Sie einen Messplan**, der sämtliche zu messenden Größen inkl. Fehlerangaben (!) jedes Aufgabenteils enthält.
Bearbeiten Sie die Aufgaben zur Vorbereitung.

Literatur

- D. Meschede: *„Optik, Licht und Laser"*, Vieweg + Teubner, 2008
- M. Young: *„Optik, Laser, Wellenleiter"*, Springer, 1997
- W. Demtröder: *„Experimentalphysik 2, Elektrizität und Optik"*, Springer, 2009
- L. E. Ballentine: *„Quantum Mechanics, A Modern Development"*, World Scientific, 1998
- F. Schwabl: *„Advanced Quantum Mechanics"*, Springer (2000)
- C. Gerry und P. Knight: *„Introductory Quantum Optics"*, Cambridge University Press, 2005 (siehe Anhang)

Außerdem findet sich ein interaktives Experiment unter folgendem Link (Stand 2019).
http://www.didaktik.physik.uni-erlangen.de/quantumlab/index.html?/quantumlab/Hong/index.html

Einleitung

Nichtlineare optische Prozesse haben eine große Bedeutung in der modernen Optik. Durch die spontane parametrische Fluoreszenz (engl. *spontaneous parametric down-conversion*, SPDC) lassen sich (verschränkte) Photonenpaare erzeugen. Damit kann eine Vielzahl von Quanteneffekten experimentell beobachten werden, welche den Vorstellungen der klassischen Physik widersprechen. Ein Experiment von Hong, Ou und Mandel hat als erstes die Interferenzfähigkeit der beiden Photonen demonstriert. Dabei trifft jeweils ein einzelnes Photon auf die zwei Eingänge (a und b) eines Strahlteilers. Dort werden sie mit jeweils 50 % Wahrscheinlichkeit in einen der beiden Ausgänge (c und d) transmittiert oder reflektiert. Die quantenphysikalische Wellenfunktion eines Photons am Eingang a bzw. Eingang b kann dabei formal beschrieben werden mit

$$\hat{a}^\dagger|0\rangle_a|0\rangle_b = |1\rangle_a|0\rangle_b \equiv |1,0\rangle_{ab} \equiv |1\rangle_a \quad \text{bzw.} \quad \hat{b}^\dagger|0\rangle_a|0\rangle_b = |1\rangle_b \ .$$

Die Operatoren $\hat{a}^\dagger$ bzw. $\hat{b}^\dagger$ wirken auf die Vakuumzustände $|0\rangle_a$ bzw. $|0\rangle_b$ mit jeweils null Photonen und erzeugen die Einzelphotonenzustände $|1\rangle_a$ bzw. $|1\rangle_b$. Für Erzeugungsoperatoren von Photonen gilt $(\hat{a}^\dagger)^n|0\rangle = \sqrt{n!}\,|n\rangle$.
Für die Wirkung des Strahlteilers auf die einzelnen Photonen ergibt sich

$$|1\rangle_a \xrightarrow{Strahlteiler} \frac{1}{\sqrt{2}}\left(\hat{c}^\dagger + i\hat{d}^\dagger\right)|0\rangle_c|0\rangle_d = \frac{1}{\sqrt{2}}\left(|1\rangle_c|0\rangle_d + i|0\rangle_c|1\rangle_d\right)$$

$$\text{und} \quad |1\rangle_b \xrightarrow{Strahlteiler} \frac{1}{\sqrt{2}}\left(i|1\rangle_c|0\rangle_d + |0\rangle_c|1\rangle_d\right) .$$

Dabei sind $\hat{c}^\dagger$ und $\hat{d}^\dagger$ die Erzeugungsoperatoren, welche auf die Vakuumzustände $|0\rangle_c$ und $|0\rangle_d$ der Ausgänge wirken. Für die reflektierten Komponenten gibt es eine Phasenverschiebung von $\pi/2$ in der komplexen Wellenfunktion und die einzelnen Photonen treten jeweils mit gleicher Wahrscheinlichkeit in nur einem der beiden Ausgänge auf. Für ein Photonenpaar $|1\rangle_a|1\rangle_b$ an beiden Eingängen des Strahlteilers erhält man

$$|1\rangle_a|1\rangle_b \xrightarrow{Strahlteiler} \frac{1}{2}\left(\hat{c}^\dagger + i\hat{d}^\dagger\right)\left(i\hat{c}^\dagger + \hat{d}^\dagger\right)|0\rangle_c|0\rangle_d$$

$$= \frac{1}{2}\Big(\underbrace{i\hat{c}^{\dagger 2}}_{I.} + \underbrace{\hat{c}^\dagger\hat{d}^\dagger}_{II.} - \underbrace{\hat{d}^\dagger\hat{c}^\dagger}_{III.} + \underbrace{i\hat{d}^{\dagger 2}}_{IV.}\Big)|0\rangle_c|0\rangle_d \ .$$

Es entstehen vier Terme die den vier möglichen Wegen der Photonen entsprechen (siehe Abbildung 1).

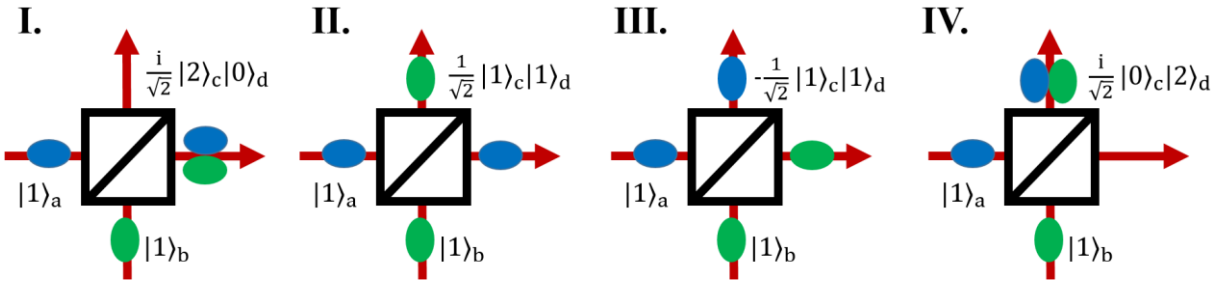


Abbildung 1: Die vier Möglichkeiten für zwei Photonen einen Strahlteiler zu verlassen.

Für ununterscheidbare Photonen gilt $[\hat{c}^\dagger,\ \hat{d}^\dagger] = 0$ und es ergibt sich

$$|1\rangle_a|1\rangle_b \xrightarrow{Strahlteiler} \frac{1}{2}\left(i\hat{c}^{\dagger 2} + i\hat{d}^{\dagger 2}\right)|0\rangle_c|0\rangle_d = \frac{i}{\sqrt{2}}\left(|2\rangle_c|0\rangle_d + |0\rangle_c|2\rangle_d\right).$$

Die Möglichkeiten II. und III. treten also nicht auf. Es werden daher immer nur zwei Photonen an einem der beiden Ausgänge des Strahlteilers gemessen.

In diesem Versuch wird mit einer Photonenquelle und einer Strahlteilerfaser ein Hong-Ou-Mandel Interferometer aufgebaut und die Koinzidenzrate gemessen. Aus der beobachteten Interferenz sollen die Unterscheidbarkeit der Photonen und die Kohärenzlänge bestimmt werden. Außerdem wird der optische Wegunterschied einer Wellenplatte gemessen und die Konversionseffizienz der spontanen parametrischen Fluoreszenz bestimmt.

Experimenteller Aufbau

Der Aufbau des Versuchs besteht aus drei Komponenten:

1. Photonenpaarquelle
2. Hong-Ou-Mandel Interferometer
3. Elektronische Kontroll- und Ausleseeinheit

1. Photonenpaarquelle

Der Aufbau der Photonenquelle ist in Abbildung 2 dargestellt. Als Pumplaser dient eine Laserdiode, welche sich in einem Gehäuse mit einem Anschluss für die Stromversorgung (1) befindet.

> WICHTIG: *Plötzliche Stromänderungen und Spannungsentladungen können die Laserdiode zerstören. Die Stromversorgung muss daher **immer** angeschlossen sein! Außerdem soll die Laserdiode nicht berührt werden!*

Das emittierte Licht mit etwa 405 nm Wellenlänge wird durch Spiegel umgelenkt und mit verschiedenen Linsen kollimiert. Doppelbrechende Kompensationskristalle (2) kompensieren eine zeitliche Unterscheidbarkeit zwischen vertikal und horizontal polarisierten Photonen für weitere, hier nicht relevante, Versuche. In dem nichtlinearen BBO-Kristall (3) findet eine spontane parametrische Fluoreszenz des Typ I statt. Dadurch entstehen mit einer geringen Wahrscheinlichkeit Photonenpaare bei einer Wellenlänge von etwa 810 nm, welche unter einem bestimmten Öffnungswinkel austreten. Zum Schutz gegen die Laserstrahlung befindet sich eine Schutzabdeckung über der Lichtquelle. Ein Filter blockt den Strahl des Pumplasers und lässt nur die einzelnen frequenzkonvertierten Photonenpaare nach außen. Diese werden mit Spiegeln umgelenkt und jeweils in eine polarisationserhaltende Einmodenfaser eingekoppelt (5). Davor können noch Bandpassfilter (4) in den Strahlgang gestellt werden. Außerdem ist das Aufsetzen von Langpassfiltern auf die Einkoppler möglich.

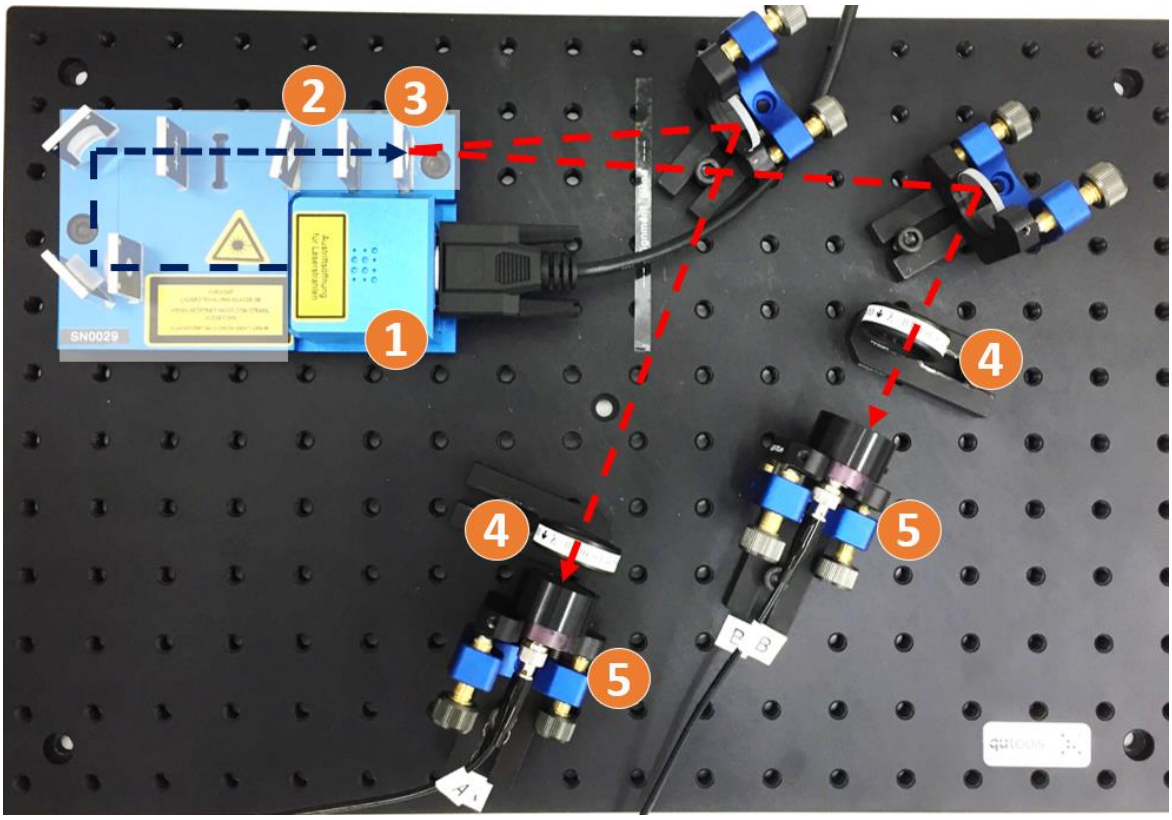


Abbildung 2: Aufbau für die Erzeugung von Photonenpaaren:
(1) Laser, (2) Kompensationskristalle, (3) BBO-Kristall, (4) Bandpassfilter,
(5) Optische Fasern mit Einkopplung und Langpassfiltern.

WICHTIG: *Beim Umgang mit den optischen Fasern dürfen die Endflächen nicht berührt werden, da diese sonst beschädigt werden könnten. Insbesondere muss auf das korrekte Anschließen der Fasern (siehe Abbildung 3) geachtet werden. Außerdem sollen die Schutzkappen aus Kunststoff verwendet werden, wenn die Fasern nicht angeschlossen sind.*

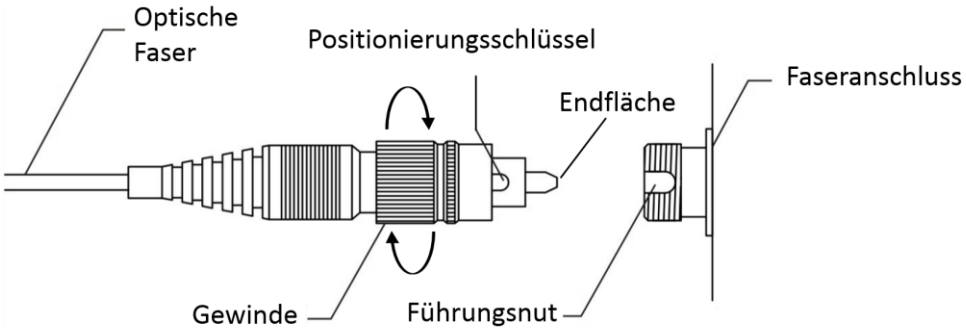


Abbildung 3: Zum Anschluss der Faser muss der Positionierungsschlüssel des Fasersteckers mit der Führungsnut des Faseranschlusses übereinstimmen. Danach kann über Drehung des Gewindes der Stecker festgeschraubt werden.

2. Hong-Ou-Mandel Interferometer

Abbildung 4 zeigt den Aufbau des Interferometers. Die Photonen aus den beiden mit A und B gekennzeichneten Fasern werden ausgekoppelt (4) und legen eine freie Wegstrecke zurück bevor sie wieder in Fasern eingekoppelt (5) werden. Diese Länge ist für einen der Wege fest (2). Für den anderen Weg (1) kann die Länge mit einer Mikrometerschraube (3) mit einer Genauigkeit von wenigen µm verstellt werden. Die eingekoppelten Photonen durchlaufen eine Strahlteilerfaser (6) mit einem Verzweigungsverhältnis von 50 % und können dann an den beiden Ausgängen gezählt werden.

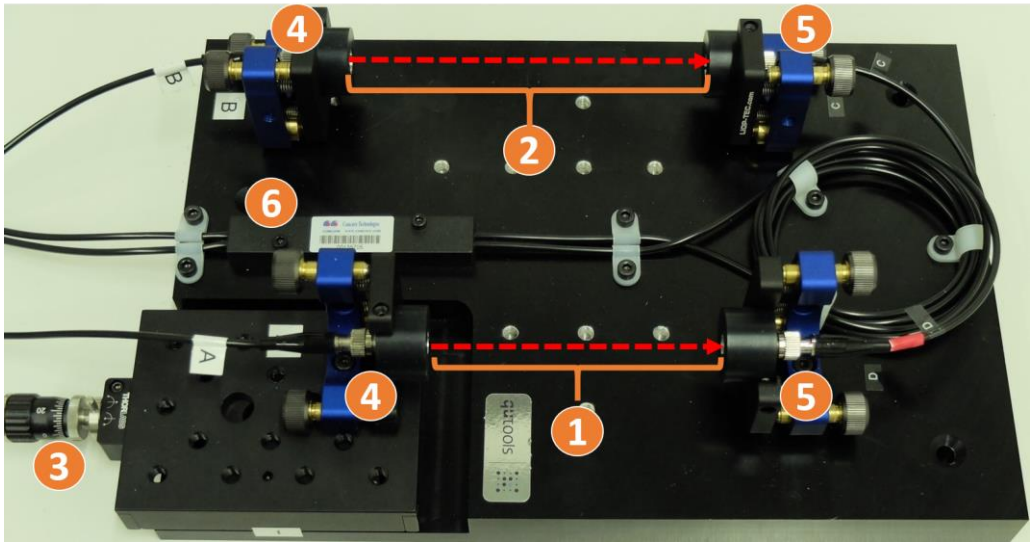


Abbildung 4: Optischer Aufbau des Hong-Ou-Mandel-Interferometers:
(1) Variable Weglänge, (2) Feste Weglänge, (3) Mikrometerschraube,
(4) Faserauskoppler, (5) Fasereinkoppler, (6) Strahlteilerfaser.

3. Elektronische Auslese- und Kontrolleinheit

Das Frontdisplay der Kontrolleinheit ist in Abbildung 5 gezeigt. Rechts befindet sich die Detektionseinheit (3), welche zwei Lawinenphotodioden für den Nachweis einzelner Photonen besitzt. Hier müssen die optischen Fasern angeschlossen werden. Mit dem Schlüsselschalter (1) wird die Kontrolleinheit eingeschaltet. Mit dem Schalter Laser on/off (2) wird der Laser ein- und ausgeschaltet. Die Steuerung und Anzeige aller Funktionen erfolgt über ein Touch Panel. Mit dem Drehgeber (4) können Werte wie Laserstrom und Integrationszeit verändert werden. Über das oberste Symbol der Menüspalte (5) werden Warnungen angezeigt und die Kontrolloptionen (6) für den Laser erreicht.

WICHTIG: *Den Schlüsselschalter **niemals** ausschalten, während der Laser an ist.*

Vor dem Ein- und Ausschalten des Lasers ist der Laserstrom auf null zurückzustellen.

*Der Laserstrom sollte immer nur **langsam** verändert werden.*

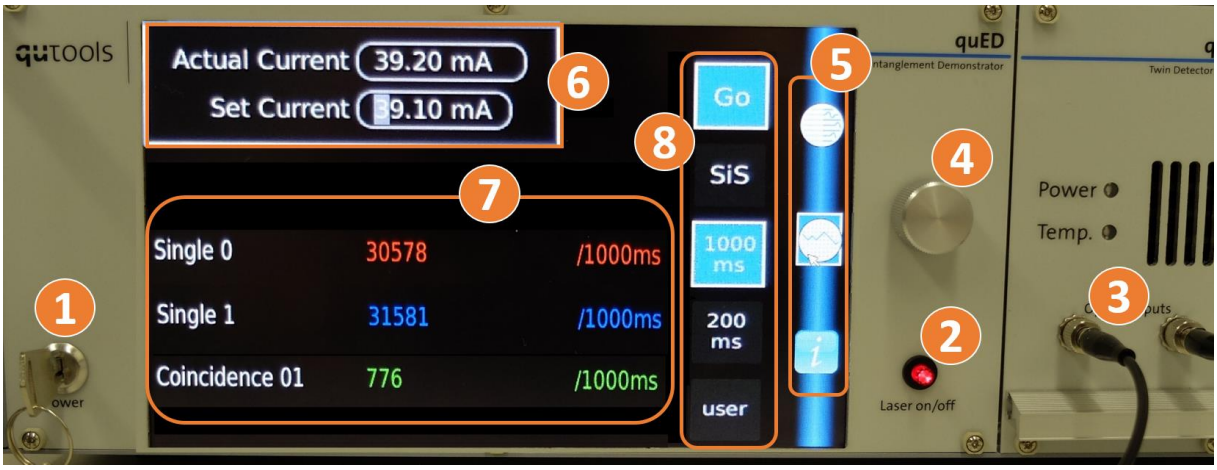


Abbildung 5: Elektronische Auslese- und Kontrolleinheit: (1) Schlüsselschalter, (2) Ein-/Ausschalter Laser, (3) Detektormodul, (4) Drehgeber, (5) Menüspalte, (6) Laserstrom, (7) Zählraten, (8) Messoptionen.

Über das mittlere Menüsymbol wird die Anzeige für die Zählraten der Einzelphotonen und der Koinzidenzen (7) ausgewählt. Als Messoptionen (8) gibt es die Möglichkeit zwischen kontinuierlicher Messung (*Go*) und Einzelmessung (*SiS*). Außerdem kann die Integrationszeit eingestellt werden.

Versuchsdurchführung & Auswertung

WICHTIG: *Dokumentieren Sie immer **alle** Messergebnisse der einzelnen Aufgaben! Die Messdaten sind am Ende des Versuchs vom Betreuer unterzeichnen zu lassen.*

*Die während der Versuchsdurchführung aufgenommenen Messwerte sind **im Original** in die Auswertung einzufügen! Trennen Sie die Auswertung der Messwerte von der Versuchsdurchführung! Es muss **nachvollziehbar** sein, wie die Auswertungsergebnisse aus den Messdaten erhalten wurden!*

Aufgaben zur Vorbereitung:

1. Quantenphysikalische Interpretation

Bei der Erklärung des Hong-Ou-Mandel Effekts wird manchmal argumentiert, dass zwei einzelne Photonen miteinander interferieren würden. Korrigieren Sie diese Aussage in Übereinstimmung mit der Quantenphysik.

2. Hong-Ou-Mandel Interferenz mit Elektronen

Interferenzexperimente können nicht nur mit Photonen durchgeführt werden, sondern auch mit anderen Quantenobjekten. Was würde sich beispielsweise verändern, wenn in diesem Experiment Elektronen verwendet würden? Beschreiben Sie qualitativ die zu erwartende Interferenzstruktur und begründen Sie diese mathematisch.

3. Kontrast der Hong-Ou-Mandel Interferenz

Berechnen Sie den Kontrast der Hong-Ou-Mandel Interferenz für zwei identische Photonen, wenn die Polarisation eines der beiden Photonen mit einer Wellenplatte um einen Winkel α gedreht wird.

Messaufgaben:

1. Inbetriebnahme und Messung der Konversionseffizienz

Nehmen Sie die Photonenquelle **unter Anleitung des Betreuers** in Betrieb. Schließen Sie dazu die Ausgänge der Fasern A und B an die Detektionseinheit an und vergewissern Sie sich, dass das Kabel für die Stromversorgung des Lasers korrekt angeschlossen ist. Schalten Sie dann das Kontrollmodul an und warten Sie einige Minuten bis die Lämpchen für Stromversorgung und Temperatur an der Detektionseinheit verlöschen. Danach schalten Sie den Laser ein und erhöhen den Laserstrom **langsam** auf **39.1 mA**. Überprüfen Sie anschließend die Justage der Spiegel und Fasereinkoppler **vorsichtig** darauf, dass an beiden Detektoren eine ähnlich hohe Zählrate gemessen wird und maximieren Sie die Koinzidenzrate. Bei der Justage dürfen die Zählraten **zu keinem Zeitpunkt komplett auf die Untergrundrate absinken**, da sonst ein Rückstellen auf den vorherigen Justagezustand, auch für den Betreuer, extrem aufwendig ist.

Nehmen Sie nun die Einzelphotonenzählrate für beide Kanäle sowie die Koinzidenzrate bei einer Integrationszeit von 10 s auf. Führen Sie weiterhin eine Untergrundmessung mit geblockten Stahlen durch.

2. Justage des Interferometers

WICHTIG: *Direkte Einstrahlung des Laserlichts in die Einzelphotonenzähler könnte die Detektoren zerstören. Deshalb darf der **zweite Ausgang** des Strahlteilers während der Justage mit dem Justierlaser **nicht** an die Detektionseinheit **angeschlossen** sein.*

Abbildung 6: Justierlaser mit abgedecktem Faseranschluss links.

Stellen Sie mit dem Justierlaser (Abbildung 6) sicher, dass Licht von beiden Armen des Interferometers in den Strahlteiler eingekoppelt wird. Schließen Sie dazu wie in Abbildung 7 dargestellt den Justierlaser an einen der beiden Ausgänge des Strahlteilers ((6) in Abbildung 4) an und verbinden Sie den anderen Ausgang jeweils mit einem der durch A oder B gekennzeichneten Faserauskoppler ((4) in Abbildung 4). Beobachten Sie mit einem Stück Papier die Position der Strahlen und versuchen Sie diese auf der gesamten freien Wegstrecke des Interferometers zu überlagern. Dazu können die Schrauben für horizontale und vertikale Einstellung an den Aus- und Einkopplern ((4) und (5) in Abbildung 4) verwendet werden. Dann wiederholen Sie dies für den anderen Arm des Interferometers.

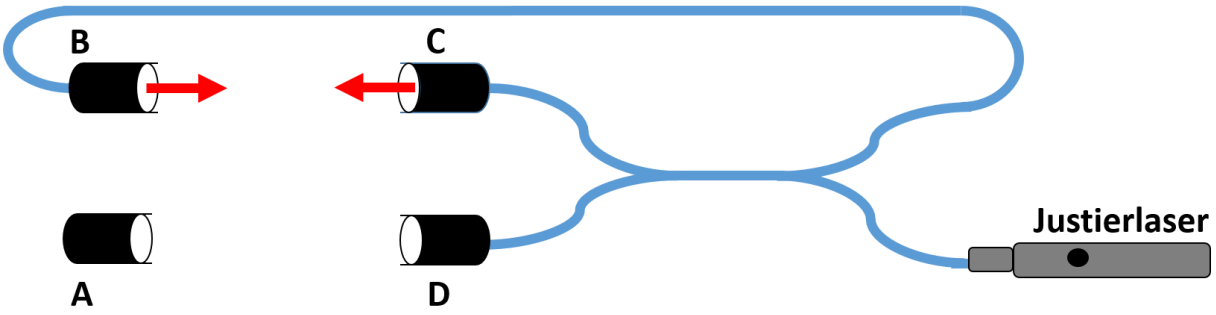


Abbildung 7: Justage eines Arms des Interferometers. An den Auskopplern C und D entstehen schwache Reflexe, die in entgegengesetzter Richtung durch die Strahlteilerfaser laufen und bei B ausgekoppelt werden.

Danach werden die Fasern (A bzw. B) von der Photonenquelle an den entsprechenden Faserauskoppler des Interferometers (A bzw. B) angeschlossen. Beobachten Sie am anderen Ende der Fasern ((5) in Abbildung 2), ob das Licht des Justierlasers die Fasern verlässt (entfernen Sie hierfür vorsichtig die Langpassfilter) und optimieren Sie dies gegebenenfalls durch leichtes Verstellen der Schrauben von Aus- und Einkopplern des Interferometers ((4) und (5) in Abbildung 4).
Anschließend entfernen Sie den Justierlaser und verbinden die Ausgänge des Interferometers mit dem Zählmodul. Blockieren Sie jeweils einen der beiden Strahlen und beobachten Sie die Einzelphotonenzählraten. Optimieren Sie diese nochmal durch Verstellen der Schrauben der Faserkoppler ((4) und (5) in Abbildung 4) bis im Vergleich zu den in Aufgabe 1 gemessenen Zählraten mindestens eine Einkoppeleffizienz von 60 % erreicht ist.

(Die Winkeleinstellungen von Ein- und Auskoppler sind miteinander verknüpft. Jede Änderung der vertikalen oder horizontalen Ausrichtung des Auskopplers führt zu einem Abfall der Zählrate, was durch eine entsprechende Änderung in der gleichen Richtung für den Einkoppler ausgeglichen werden muss, damit die Rate wieder ein Maximum erreicht. Durch iteratives Optimieren in beiden Richtungen sollte die beste Einstellung gefunden werden können.)

3. Beobachtung der Hong-Ou-Mandel Interferenz

- Zum Versuchsaufbau gehört eine λ/2-Platte, welche sich in einer rotierbaren Halterung befindet. Bestimmen Sie mit dem Polarisator die Lage der optischen Achsen der λ/2- Platte und stellen Sie diese unter einem Winkel von $\boldsymbol{\Theta = 0°}$ zwischen der horizontalen Richtung und einer gewählten optischen Achse in den Strahlengang eines der beiden Photonenpfade.

- Setzen Sie die zwei Bandpassfilter in die Strahlengänge ein. Durch leichtes Verkippen kann die Zentralwellenlänge der Filter auf die Wellenlänge der Photonen abgestimmt werden, indem der Verkippungswinkel auf hohe Koinzidenzrate optimiert wird.
  Beobachten Sie die Koinzidenzzählrate, während Sie die freie Weglänge mit der Grobeinstellung der Mikrometerschraube verändern. Sobald es einen deutlichen Einbruch (<50%) der Koinzidenzrate gibt, ziehen Sie die Feststellschraube an und bestimmen mit der Feineinstellung den Bereich, in dem das Minimum der Koinzidenzrate sichtbar ist. Setzen Sie die Integrationszeit auf **10 s** und messen Sie die Koinzidenzrate im Einzelmodus in Abhängigkeit der relativen Wegänderung alle **2,5**

µm innerhalb eines Bereichs, der das Minimum großzügig abdeckt. Notieren Sie einmal während dieser Messung die Einzelphotonenzählraten.

- Entfernen Sie die Bandpassfilter und stellen Sie die λ/2-Platte auf $\boldsymbol{\Theta = 0°}$ zwischen gewählter optischer Achse und horizontaler Richtung. Bestimmen Sie wieder die Lage des Minimums der Koinzidenzrate und vermessen Sie dieses nun in Schritten von **0,5 µm**. Notieren Sie auch hier einmal die Einzelphotonenzählraten. (Eventuell müssen Sie den Bereich für das Minimum erneut suchen, benutzten Sie dazu jedoch nur die Feineinstellung der Mikrometerschraube und notieren Sie die Positionen, so dass Sie die relativen Wegänderungen zwischen den einzelnen Messreihen nachher vergleichen können.)
- Wiederholen Sie diese Messung unter einem Winkel der von $\boldsymbol{\Theta = 20°}$, $\boldsymbol{\Theta = 40°}$ und $\boldsymbol{\Theta = 90°}$ zwischen der horizontalen Richtung und der zuvor gewählten optischen Achse der λ/2- Platte. Notieren Sie dabei jeweils einmal für jede Messreihe die Einzelphotonenzählraten.
- Drehen Sie die λ/2-Platte von $\boldsymbol{\Theta = 0°}$ bis **180°** in Schritten von 10° und führen Sie jeweils eine Koinzidenzmessung an der Position des Minimums durch, welche Sie zuvor jeweils grob bestimmen. Die maximale Koinzidenzrate außerhalb der Interferenz-struktur und die Einzelphotonenzählraten sollen ebenfalls notiert werden.

Auswertung:

1. Konversionseffizienz der spontanen parametrischen Fluoreszenz

Die Ausgangsleistung des Pumplasers beträgt $P_{405\,\mathrm{nm}} = (14{,}28 \pm 0{,}04)$ mW. Berechnen Sie die Konversionseffizienz für den Prozess der spontanen parametrischen Fluoreszenz und vergleichen Sie mit Literaturwerten.

2. Hong-Ou-Mandel Interferenz

Tragen Sie die gemessenen Interferenzstrukturen grafisch auf und passen Sie jeweils eine geeignete Funktion an die Messdaten an. Die Koinzidenzen werden innerhalb eines Zeitfensters von 30 ns registriert. Machen Sie aus den notierten Einzelphotonenraten eine Abschätzung für die zufälligen Koinzidenzen und korrigieren Sie die Zählraten entsprechend. Die gemessen Zählraten sollten einer Poissonverteilung unterliegen. Zeichnen Sie dementsprechend die Fehlerbalken ein.

3. Kohärenzlänge der Photonen

Vergleichen Sie die unter einem Winkel von $\Theta = 0°$ für die λ/2- Platte mit und ohne Bandpassfilter gemessenen Interferenzstrukturen in dem Sie diese gemeinsam grafisch darstellen. Verwenden Sie dazu eine geeignete Normierung. Bestimmen Sie jeweils die Kohärenzlänge und setzten Sie diese in Beziehung mit den spektralen Eigenschaften durch Verwendung der Bandpassfilter. Wie groß ist die spektrale volle Halbwertsbreite ohne Bandpassfilter?

4. Unterscheidbarkeit der Photonen

Tragen Sie die mit $\Theta = 0°$, $\Theta = 20°$ und $\Theta = 40°$ gemessenen Interferenzstrukturen zusammen auf und diskutieren Sie die Unterschiede. Gleichen Sie dabei den absoluten Gangunterschied zwischen den einzelnen Messreihen aus in dem Sie die Lage der Minima auf den Ursprung auf der horizontalen Achse legen und benutzen Sie eine Normierung. Bestimmen Sie jeweils den Kontrast des Minimums für jede der drei Messreihen.

Berechnen Sie außerdem den Kontrast aus den notierten maximalen und minimalen Koinzidenzraten für verschiedene Winkeleinstellungen der Wellenplatte. Stellen Sie die Abhängigkeit aller ermittelten Kontraste gemeinsam grafisch dar und passen Sie eine geeignete Funktion an die Messdaten an. Vergleichen Sie dies mit dem zuvor berechneten theoretischen Ergebnis.

5. Optischer Wegunterschied der Wellenplatte

Zeichen Sie die die HOM-Interferenz mit $\Theta = 0°$ und $\Theta = 90°$ gemeinsam in ein Diagramm ein und bestimmen Sie daraus den optischen Wegunterschied zwischen den beiden optischen Achsen der $\lambda/2$-Platte. Um welche Art von Wellenplatte handelt sich demnach? Diskutieren Sie die Unterschiede zu anderen Wellenplatten.

Wichtige Punkte zum Laserschutz

Ganz allgemein gilt: Im Umgang mit Lasern ist der gesunde Menschenverstand nicht zu ersetzen! Einige spezielle Hinweise werden im Folgenden angeführt.

1. Die Laserschutzvorschriften sind immer zu beachten.
2. Halten Sie Ihren Kopf niemals auf Strahlhöhe.
3. Die Justierbrille immer aufsetzen.
4. Schauen Sie nie direkt in Strahl – auch nicht mit Justierbrille!
5. Achtung: Praktisch alle Laser für Laboranwendungen sind mindestens Klasse 3, also von vornherein für die Augen gefährlich, ggf. auch für die Haut – evtl. auch hierfür Schutzmaßnahmen ergreifen. Zur Justage kann der Laserstrahl mittels eines Stücks Papier sichtbar gemacht werden.
6. Auch Kameras besitzen eine Zerstörschwelle!
7. Spiegel und sonstige Komponenten nie in den ungeblockten Laserstrahl einbauen! Vor Einbau immer überlegen, in welche Richtung der Reflex geht! Diese Richtung zunächst blocken, bevor der Strahl wieder freigegeben wird.
8. Nie mit reflektierenden Werkzeugen im Strahlengang hantieren! Unkontrollierbare Reflexe! Vorsicht ist z.B. auch mit BNC-Steckern geboten, die in den Strahlengang gelangen könnten! Gleiches gilt auch für Uhren und Ringe. Diese vorsichtshalber ausziehen, wenn Sie mit den Händen im Strahlengang arbeiten.
9. Auch Leistungsmessgeräte können Reflexe verursachen! Unbeschichtete Silizium-Fotodioden reflektieren über 30 % des Lichtes!
10. Achtung im Umgang mit Strahlteilerwürfeln! Diese haben immer einen zweiten Ausgang! Ggf. abblocken!
11. Warnlampen bei Betrieb des Lasers anschalten und nach Beendigung der Arbeit wieder ausschalten.
12. Dafür sorgen, dass auch Dritte im Labor die richtigen Schutzbrillen tragen, oder sich außerhalb des Laserschutzbereiches befinden.
13. Filtergläser in Laserschutzbrillen dürfen grundsätzlich nicht aus- oder umgebaut werden!
14. In besonderem Maße auf Beistehende achten.
15. Optiken (Linsen, Spiegel etc.) nicht direkt mit den Fingern berühren!

Hiermit erkläre ich, dass ich die vorstehenden Punkte gelesen und verstanden habe. Ich bestätige, dass ich eine Einführung in den Umgang mit Lasern sowie eine arbeitsplatzbezogene Unterweisung erhalten habe.

Name:

Unterschrift: Datum:

IOP PUBLISHING EUROPEAN JOURNAL OF PHYSICS
Eur. J. Phys. **33** (2012) 1843–1850 doi:10.1088/0143-0807/33/6/1843

The Hong–Ou–Mandel interferometer in the undergraduate laboratory

Jorge Carvioto-Lagos[1], Gustavo Armendariz P[1], Víctor Velázquez[1], Enrique López-Moreno[1], M Grether[1] and E J Galvez[2]

[1] Facultad de Ciencias, Universidad Nacional Autónoma de México, 04510 México, DF, Mexico
[2] Department of Physics and Astronomy, Colgate University, Hamilton, NY 13346, USA

E-mail: mdgg@ciencias.unam.mx

Received 13 June 2012, in final form 8 August 2012
Published 11 October 2012
Online at stacks.iop.org/EJP/33/1843

Abstract
The Hong–Ou–Mandel interferometer is an optical device that allows us to prove the quantum nature of light experimentally via the quantum amplitude superposition of two indistinguishable photons. We have implemented this experiment as an advanced undergraduate laboratory experience. We were able to overcome well-known difficulties using techniques reported recently by Thomas *et al* (2009 *Rev. Sci. Instrum.* **80** 036101).

(Some figures may appear in colour only in the online journal)

1. Introduction

Recent advances in quantum information have led to a revival of quantum mechanics and its application in new technologies such as cryptogtraphy, teleportation and computation. New demonstrations with single photons have led to an appreciation of the fundamental properties of quantum mechanics and single-photon experiments are now included in quantum mechanics instruction. A series of these experiments developed for the undergraduate laboratory has successfully emphasized these fundamentals through direct experimentation. Experiments such as proof for the existence of the photon [1], the quantum eraser [2] and quantum non-realism and nonlocality [3] deliver results that embody almost 100 years of discussion on the quantum nature of fundamental particles and light.

This paper presents the implementation of a landmark experiment in quantum optics in the undergraduate laboratory: the Hong–Ou–Mandel (HOM) interference experiment [4]. Many preceding investigations (see Pfleegor and Mandel [5–8]) aimed to show how Dirac's statement [9] about the quantum origin of interference is valid. This says that a photon can interfere only with itself, but two different photons do not produce interference. It is now known that interference of single photons exists [10], and it can be proven even in an undergraduate laboratory [2]. However, Pfleegor and Mandel showed that two photons from different sources

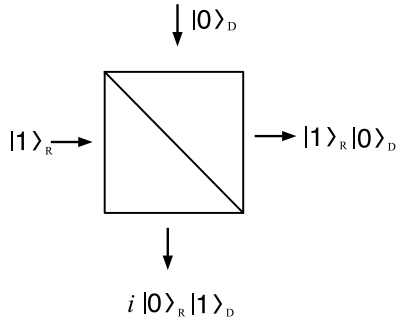


Figure 1. The possible outcomes when an individual photon is incident on a beam splitter. If the beam splitter is 50 : 50, then the probability of the photon being transmitted or reflected is 50%.

can still interfere. The HOM experiment shows that the interference of two photons occurs when they are indistinguishable. Moreover, there is a subtlety that is worth emphasizing: two photons do not interfere in a classical way, but rather it is the quantum mechanical amplitudes for finding them that interfere.

We used this experiment in the context of our graduate and advanced undergraduate quantum optics courses. At the graduate level, it represents one of the possible special projects on a compulsory course (Laboratorio Avanzado) within the Master's degree program, where the student not only treats fundamental quantum principles in an experiment, but also manipulates apparatuses and uses more sophisticated devices and technologies.

This paper is divided into three sections. The first section gives a theoretical description of quantum interference for two photons in a 50 : 50 beam splitter, in particular we indicate that it is dangerous to say that interference occurs 'between' two photons, because it leads to the misconception that in classical interferometers photons would interfere with each other and not with themselves. The second section discusses the experiment and the final section discusses the results and how we used the experiment in the curriculum.

2. Quantum description of a beam splitter and the interference of two photons

We begin by considering a single-photon incident on a 50–50 beam splitter. If initially one photon travels to the right but not 'down', i.e. coming from above in figure 1, it can be described in the photon number representation by the Fock state, or photon occupation state:

$$|\psi_i\rangle = |1\rangle_R|0\rangle_D, \tag{1}$$

where the subindices R and D refer to the mode of the light traveling to the right and down, respectively. In this problem the mode of the light where none is traveling 'down' may appear trivial or unnecessary, but the use of the 'vacuum' state $|0\rangle_D$ is an integral part of the quantized radiation field formalism. After the beam splitter, the quantum state of the light is in a superposition of traveling in two directions

$$|\psi_f\rangle = \frac{1}{\sqrt{2}}(|1\rangle_R|0\rangle_D + i|0\rangle_R|1\rangle_D), \tag{2}$$

A phase of $\pi/2$ is introduced in the state of the reflected light as is necessary to conserve energy [15]. From equation (2) the photon entering has a 50% probability of being transmitted, and a 50% probability of being reflected. If a detector is placed at each of the exits the

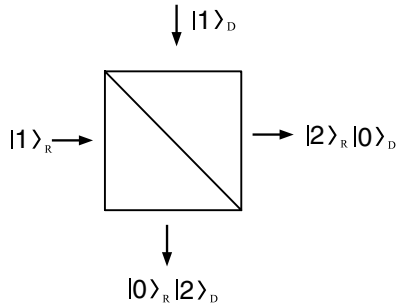


Figure 2. Two indistinguishable photons entering the beam splitter at the same time.

wavefunction collapses and there will be an anticorrelation, which means that either the photon is detected on transmission or the photon is detected on reflection [2, 10–12]. That a photon is in a superposition state of two states is mathematically simple but physically hard to explain to a beginning student. We can also say that the state of the photon is uncertain but balanced because it is equally possible to detect the photon coming out either way. Feynman is particularly insightful in his explanations of this superposition [13].

We have used the more intuitive Shrödinger picture to explain this phenomenon to the beginning student. A more rigorous and commonly used method in quantum optics uses the Heisenberg representation, photon-number creation and anihilation operators [14]. We will continue with our intuitive approach for the next case.

Let us now turn to the case of two photons entering the beam splitter from two separate ports, as shown in figure 2. A critical element of this case is the indistinguishability of the paths taken by the two photons. The pump photon produces two photons that can arrive at the detectors in two possible ways. The indistinguishability of the possibilities makes the quantum mechanical amplitudes interfere.

The state of the light can be represented by $|1\rangle_R|1\rangle_D$. After the beam splitter the state of the light is obtained by applying the beam splitter action to each occupied mode, and the state vector results, $|\psi_f\rangle$:

$$\frac{1}{2}(|1\rangle_R|0\rangle_D + i|0\rangle_R|1\rangle_D)\ (|1\rangle_D|0\rangle_R + i|0\rangle_D|1\rangle_R) \tag{3}$$

$$= \frac{i}{2}(|2\rangle_R|0\rangle_D + |0\rangle_R|2\rangle_D), \tag{4}$$

where we have made the condensation $|1\rangle_R|1\rangle_R = |2\rangle_R$, and similarly for D. The terms with $|1\rangle_R|1\rangle_D$ cancel out when the paths are indistinguishable. Thus, the interference is destructive. Experimentally, if we put detectors at the two output ports of the interferometer we would not get any coincidences because of the destructive interference. Conversely, when the photons arrive at distinguishable times the paths are distinguishable, and so the light is in an incoherent sum of two photons, each incident on the beam splitter, as in the single-photon case. That situation would lead to no interference and photons heading to the two detectors (and hence, coincidences) half of the time.

3. Experiment

The HOM experiment consists of generating a pair of photons and making them converge onto the input ports of a beam splitter. Detectors placed at the output ports of the beam

splitter detect coincidences. When the photons are indistinguishable no coincidences are recorded. The photons for the experiment come from a single source: photon pairs produced simultaneously by spontaneous parametric down conversion (SPDC). In order for the photons to be indistinguishable, they have to arrive at the beam splitter at the same time within the coherence time of the light, and be in the same polarization, spatial and momentum modes. The HOM experiment itself is notoriously difficult to recreate, especially in a teaching, non-research setting because of the challenge of making the momentum modes of the photons overlap. That is, the photons after the beam splitter must be collinear. As a consequence, alignment is difficult and tedious. Our efforts benefited from a recent paper that explains a simple way to reach the required alignment [16]. The arrangement also includes a way to vary the time delay between the two photons, so to shift from the case when they are distinguishable to the case when they are indistinguishable. When the latter is achieved one observes a drop in the coincidences, the famous 'HOM dip'. The light is collected by lens collimators attached to multimode fibers, which send the photons to avalanche photo diode (APD) single-photon detectors.

The HOM interferometer (HOMI) has three fundamental parts: the source of the pairs of photons, an interferometer and a photon detection system. The interferometer has a dual use: it is used as an interferometer (HOMI) in the experiments, and as a Mach–Zehnder interferometer (MZI) for the alignment of the optics.

The SPDC source used in the experiment is a type I beta–barium–borate (BBO) crystal, pumped by a 25 mW violet laser with a 405 nm wavelength. SPDC requires conservation of energy and momentum inside the crystal

$$\hbar\omega_p = \hbar\omega_1 + \hbar\omega_2, \tag{5}$$

and

$$\hbar\mathbf{k}_p = \hbar\mathbf{k}_1 + \hbar\mathbf{k}_2. \tag{6}$$

It is very important for the experiment that the laser beam be parallel to the optical table so that the pair of photons is detected in the same horizontal plane. This type of source produces pairs of individual photons with the same polarization and which follow symmetric trajectories about the pump laser beam direction, hereafter called the symmetry axis. In the degenerate case ($\omega_1 = \omega_2$), the two photons come out of the crystal at the same angle $\theta = \cos^{-1}(\mathbf{k}_p \cdot \mathbf{k}_i/(k_p k_i)$, $i = 1, 2$, but on opposite sides of the symmetry axis.

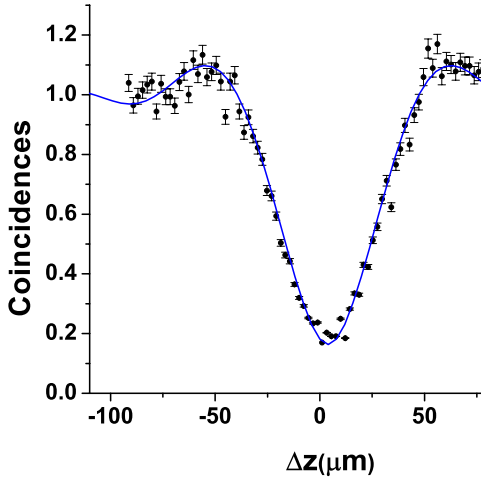


Figure 5. Normalized coincidences versus displacement in optical path (Δz). Each pace produced by the motor modifies the optical path by $2\,\mu$m. The average coincidence count rate outside the zone of interference is around 30^{-1}.

In the third stage of our experiment the individual photons went through a wide band filter (810 ± 10 nm) before entering the collimators connected to the APDs by an optical fiber. The detectors generated square pulses that were 25 ns long and 3.5 V in height. These were then sent into a photon counting board. This board can detect coincidences between four inputs [17] within a delay interval that can be adjusted to between 30 to 2500 ns. Before taking any data, the collectors should be aligned as before, so there is a maximum number of counts of individual photons. If the system has been aligned correctly, then the number of coincidences per second in a 30 ns window should be very small. This is because of the destructive HOM interference. If the optical path difference (Δz) is changed by moving the stepper-motor-driven stage, the paths may no longer interfere and the coincidence count will grow. There should be a minimum of coincidence counts when the destructive interference is taking place. The individual photon counts should be almost constant because the number of interfering photons is very small and is only altered by the pairs of photons that interfere. Figure 5 shows the behavior of the photon coincidences at the output ports of the beam splitter. The average count rate of coincidences far from the interference zone is 30 s^{-1}. A Gaussian function multiplied by a sinc function was fitted to the experimental data. The fitting function depends on the bandwidth of the down-converted photons and the difference in optical path [18].

Measurement of Subpicosecond Time Intervals between Two Photons by Interference

C. K. Hong, Z. Y. Ou, and L. Mandel
Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627
(Received 10 July 1987)

A fourth-order interference technique has been used to measure the time intervals between two photons, and by implication the length of the photon wave packet, produced in the process of parametric down-conversion. The width of the time-interval distribution, which is largely determined by an interference filter, is found to be about 100 fs, with an accuracy that could, in principle, be less than 1 fs.

PACS numbers: 42.50.Bs, 42.65.Re

The usual way to determine the duration of a short pulse of light is to superpose two similar pulses and to measure the overlap with a device having a nonlinear response.[1] The latter might, for example, make use of the process of harmonic generation in a nonlinear medium. Indeed, such a technique was recently used[2] to determine the coherence length of the light generated in the process of parametric down-conversion.[3] The coherence time was found to be of subpicosecond duration, as predicted theoretically.[4] It is, however, in the nature of the technique that it requires very intense light pulses and would be of no use for the measurement of single photons. On the other hand, if we are dealing with two photons and wish to determine the time interval between them, which has a dispersion governed by the length of the photon wave packet, we are usually limited by the resolving time of the photodetector to intervals of order 100 ps or longer.[5]

We wish to report an experiment in which the time interval between signal and idler photons, and by implication the length of a subpicosecond photon wave packet, produced in parametric down-conversion was measured. The technique is based on the interference of two two-photon probability amplitudes in two-photon detection, and is easily able to measure a time interval of 50 fs, with an accuracy that could be 1 fs or better.

An outline of the experiment is shown in Fig. 1. A coherent beam of light of frequency ω_0 from an argon-ion laser oscillating on the 351.1-nm line falls on an 8-cm-long nonlinear crystal of potassium dihydrogen phosphate, where some of the incident photons split into two lower-frequency signal and idler photons of frequencies ω_1 and ω_2, such that

$$\omega_0 = \omega_1 + \omega_2. \tag{1}$$

The two signal and idler photons are directed by mirrors M1 and M2 to pass through a beam splitter BS as shown, and the superposed beams interfere and are detected by photodetectors D1 and D2. We measure the rate at which photons are detected in coincidence, when the beam splitter is displaced from its symmetry position by various small distances $\pm c\,\delta\tau$. It should be emphasized that the signal and idler photons have no definite phase, and are therefore mutually incoherent, in the sense that they exhibit no second-order interference when brought together at detector D1 or D2. However, fourth-order interference effects occur, as demonstrated by the coincidence counting rate between D1 and D2.[6-8] The experiment has some similarities to another, recently reported, two-photon interference experiment in which fringes were observed and measured, but without the use of a beam splitter.[6]

Although the sum frequency $\omega_1+\omega_2$ is very well defined in the experiment, the individual down-shifted frequencies ω_1,ω_2 have large uncertainties, that, in practice, are largely determined by the pass bands of the interference filters IF inserted in the down-shifted beams, as shown in Fig. 1. These pass bands are of order 5×10^{12} Hz, corresponding to a coherence time for each photon of order 100 fs. Needless to say, the two-photon probability amplitudes at the detectors D1,D2 are expected to interfere only if they overlap to this accuracy or better. We start by examining how this interference arises.

Let us label the field modes on the input sides of the beam splitter by 01,02 and on the output sides by 1,2 and suppose first that the light is monochromatic. If we take the state at the input resulting from one degenerate down-conversion to be the two-photon Fock state $|1_{01},1_{02}\rangle$, then one can show from general arguments[7] that the state on the output side of the beam splitter is

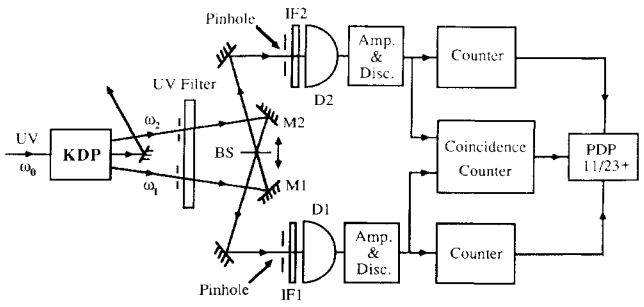


FIG. 1. Outline of the experimental setup.

given by

$$|\psi_{\text{out}}\rangle=(R-T)\,|1_1,1_2\rangle+i(2RT)^{1/2}\,|2_1,0_2\rangle+i(2RT)^{1/2}\,|0_1,2_2\rangle, \tag{2}$$

where R and T are the reflectivity and transmissivity of the beam splitter, with $R+T=1$. It follows that for a 50%:50% beam splitter with $R=\frac{1}{2}=T$, the first term is zero by virtue of the destructive interference of the corresponding two-photon probability amplitudes. No coincidences (other than accidentals) should therefore be registered by detectors D1 and D2.

In practice the down-shifted photons are never monochromatic. Let us represent the two-photon state produced by the potassium-dihydrogen-phosphate crystal by the linear superposition

$$|\psi\rangle=\int d\omega\,\phi(\omega_1,\omega_0-\omega_1)\,|\omega_1,\omega_0-\omega_1\rangle, \tag{3}$$

where $\phi(\omega_1,\omega_2)$ is some weight function which is peaked at $\omega_1=\frac{1}{2}\omega_0=\omega_2$. (A plausible example is given in Ref. 8.) We assume here that the directions of signal and idler photons are reasonably well defined by apertures, but the frequency spreads are substantial; in practice they are largely determined by the interference filters IF. Then the joint probability of the detection of photons at both detectors D1 and D2 at times t and $t+\tau$, respectively, is given by[9]

$$P_{12}(\tau)=K\langle\hat{E}_1^{(-)}(t)\hat{E}_2^{(-)}(t+\tau)\hat{E}_2^{(+)}(t+\tau)\hat{E}_1^{(+)}(t)\rangle, \tag{4}$$

where $\hat{E}_1^{(+)}(t)$ and $\hat{E}_2^{(+)}(t)$ are the positive-frequency parts of the fields at detectors D1 and D2, and K is a constant characteristic of the detectors. $\hat{E}_1^{(+)}(t)$ and $\hat{E}_2^{(+)}(t)$ are related to the fields $\hat{E}_{01}^{(+)}(t)$ and $\hat{E}_{02}^{(+)}(t)$ at the two mirrors M1 and M2 shown in Fig. 1. If $R+T=1$, then

$$\hat{E}_1^{(+)}(t)=\sqrt{T}\hat{E}_{01}^{(+)}(t-\tau_1)+i\sqrt{R}\hat{E}_{02}^{(+)}(t-\tau_1+\delta\tau), \tag{5}$$

$$\hat{E}_2^{(+)}(t)=\sqrt{T}\hat{E}_{02}^{(+)}(t-\tau_1)+i\sqrt{R}\hat{E}_{01}^{(+)}(t-\tau_1-\delta\tau). \tag{6}$$

Here τ_1 is the propagation time from mirror to detector, and $\pm c\,\delta\tau$ represents the small displacement of the beam splitter BS towards one or the other detector.

By combining Eqs. (3) to (6) we may readily show that the joint probability is

$$P_{12}(\tau)=K\,|G(0)|^2\{T^2|g(\tau)|^2+R^2|g(2\delta\tau-\tau)|^2-RT[g^*(\tau)g(2\delta\tau-\tau)+\text{c.c.}]\}, \tag{7}$$

where $G(\tau)$ is the Fourier transform of the weight function $\phi(\omega_0/2+\omega,\omega_0/2-\omega)$ with respect to ω,

$$G(\tau)=\int\phi(\omega_0/2+\omega,\omega_0/2-\omega)e^{-i\omega\tau}d\omega, \tag{8}$$

and $g(\tau)\equiv G(\tau)/G(0)$. This shows how $P_{12}(\tau)$ varies with the time interval τ. If $\phi(\omega_0/2+\omega,\omega_0/2-\omega)$ is real and symmetric in ω, as we assume, then $G(\tau)$ and $g(\tau)$ are both real and symmetric in τ.

In practice the coincidence measurement corresponds to an integration of the probability $P_{12}(\tau)$ with respect to τ over the coincidence resolving time of a few nanoseconds, but as this time is so much longer than the correlation time of $g(\tau)$ in the experiment, we may effectively integrate $P_{12}(\tau)$ over all τ. From Eq. (7) the expected number N_c of observed photon coincidences is then given by

$$N_c=C\left[R^2+T^2-2RT\frac{\int_{-\infty}^{\infty}g(\tau)g(\tau-2\delta\tau)d\tau}{\int_{-\infty}^{\infty}g^2(\tau)d\tau}\right], \tag{9}$$

where C is another constant. It follows from this relation that $N_c=C(R-T)^2$ when $\delta\tau=0$, which vanishes when $R=\frac{1}{2}=T$, whereas $N_c=C(T^2+R^2)$ when $\delta\tau$ appreciably exceeds the correlation time of $g(\tau)$. A plot of the number of coincidences N_c versus the displacement $\delta\tau$ therefore should exhibit a sharp dip near $\delta\tau=0$, of width determined by the length of the wave packet, or coherence time, of the down-shifted photons. The vanishing of the photon coincidence rate is a purely quantum-mechanical feature of the fourth-order interference, as has been shown.[6,8]

In the special case when $g(\omega_0/2+\omega,\omega_0/2-\omega)$ is Gaussian in ω with bandwidth $\Delta\omega$, then $g(\tau)$ has the Gaussian form

$$g(\tau)=e^{-(\Delta\omega\,\tau)^2/2}, \tag{10}$$

and Eq. (9) yields

$$N_c=C(T^2+R^2)\left[1-\frac{2RT}{R^2+T^2}e^{-(\Delta\omega\,\delta\tau)^2}\right]. \tag{11}$$

In the experiment the path difference $c\,\delta\tau$ was varied by our mounting the beam splitter on a translator and making very small displacements with a precisely calibrated micrometer. Still finer adjustments can be made with a piezoelectric transducer. The coincidence counting rate was measured by our feeding the amplified and standardized photomultiplier pulses to the start and the stop inputs of a time-to-digital converter, and recording the time interval distribution. Because of the transit-

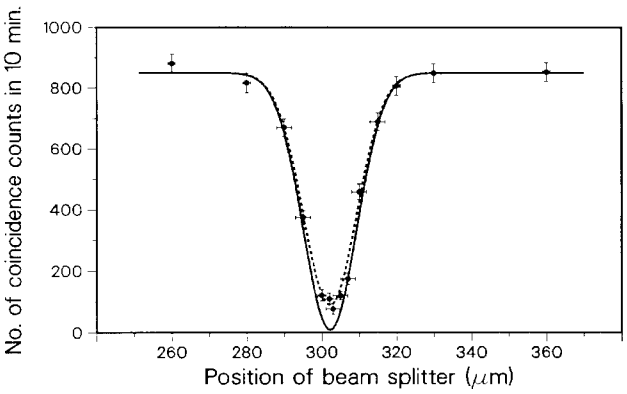


FIG. 2. The measured number of coincidences as a function of beam-splitter displacement $c\,\delta\tau$, superimposed on the solid theoretical curve derived from Eq. (11) with $R/T=0.95$, $\Delta\omega=3\times10^{13}$ rad s^{-1}. For the dashed curve the factor $2RT/(R^2+T^2)$ in Eq. (11) was multiplied by 0.9. The vertical error bars correspond to one standard deviation, whereas horizontal error bars are based on estimates of the measurement accuracy.

time spread of the photoelectric pulses and the slewing of the discriminator pulses, a range of time intervals centered on zero delay was obtained with a spread of several nanoseconds. For the purpose of the measurement, pulse pairs received within a 7.5-ns interval were treated as "coincident." Pulse pairs received within an interval of 35 to 80 ns were regarded as accidentals, and when scaled by the factor 7.5/45 provided a measure of the number of accidental coincidences that occur within any 7.5-ns interval.

The results of the experiment are presented in Fig. 2, which is a plot of the number of observed photon coincidences, after subtraction of accidentals, as a function of the displacement of the beam splitter. It will be seen that for a certain symmetric position of the beam splitter, the two-photon coincidence rate falls to a few percent of its value in the wings, by virtue of the destructive interference of the two two-photon probability amplitudes. The width of the dip in the coincidence rate provides a measure of the length of the photon wave packet. It is found to be about 16 μm at half height, corresponding to a time of about 50 fs, which should really be doubled to allow for the greater movement of the mirror image. This time is about what is expected from the passband of the interference filters.

Direct measurements of the beam-splitter reflectivity and transmissivity show that $R/T\approx0.95$, which makes the combination $2RT/(R^2+T^2)\approx0.999$, and implies that N_c should fall close to zero when $\delta\tau=0$. That it does not fall quite that far is probably due to a slight lack of overlap of the signal and idler fields admitted by the two pinholes, causing less than perfect destructive interference. The solid curve in Fig. 2 is based on Eq. (11) with $R/T=0.95$ and $\Delta\omega=3\times10^{13}$ rad/s $\approx5\times10^{12}$ Hz, if we identify $c\,\delta\tau$ with the beam-splitter displacement $(x-302.5)$ in micrometers. For the dashed curve the factor $2RT/(R^2+T^2)$ was multiplied by 0.9 to allow for less than perfect overlap of the signal and idler photons. It will be seen that, except for the minimum, Eq. (11) is obeyed quite well, corresponding to a coherence time of about 100 fs.

We have therefore succeeded in measuring subpicosecond time intervals between two photons, and by implication the length of the photon wave packet, by a fourth-order interference technique. Unlike second-order interference, this method does not require that path differences be kept constant to within a fraction of a wavelength. The method is applicable to other situations in which pairs of single photons are produced, but becomes less efficient for more intense pulses of light, because the "visibility" of the interference is then reduced and cannot exceed 50% at high intensities.[6] In principle, the resolution could be better than 1 μm in length or 1 fs in time.

This work was supported by the National Science Foundation and by the U.S. Office of Naval Research.

[1]See, for example, E. P. Ippen and C. V. Shank, in *Ultrashort Light Pulses,* edited by S. L. Shapiro (Springer-Verlag, Berlin, 1984), 2nd ed., p. 83.

[2]I. Abram, R. K. Raj, J. L. Oudar, and G. Dolique, Phys. Rev. Lett. **57**, 2516 (1986).

[3]D. C. Burnham and D. L. Weinberg, Phys. Rev. Lett. **25**, 84 (1970).

[4]C. K. Hong and L. Mandel, Phys. Rev. A **31**, 2409 (1985).

[5]S. Friberg, C. K. Hong, and L. Mandel, Phys. Rev. Lett. **54**, 2011 (1985).

[6]R. Ghosh and L. Mandel, Phys. Rev. Lett. **59**, 1903 (1987).

[7]Z. Y. Ou, C. K. Hong, and L. Mandel, to be published.

[8]R. Ghosh, C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. A **34**, 3962 (1986).

[9]R. J. Glauber, Phys. Rev. **130**, 2529 (1983), and **131**, 2766 (1963).

Notes on Hong-Ou-Mandel interference

Agata M. Brańczyk[1]
[1]*Department of Physics, University of Toronto**
(Dated: February 4, 2013)

Disclaimer: These notes were written mostly off the top of my head, so if you notice any mistakes or typos, please contact me and I will fix them.

Hong-Ou-Mandel interference occurs when two indistinguishable photons interfere on a beam-splitter as shown in FIG. 1. It is named after C. K. Hong, Z. Y. Ou and L. Mandel who experimentally verified the effect in 1987 [1].

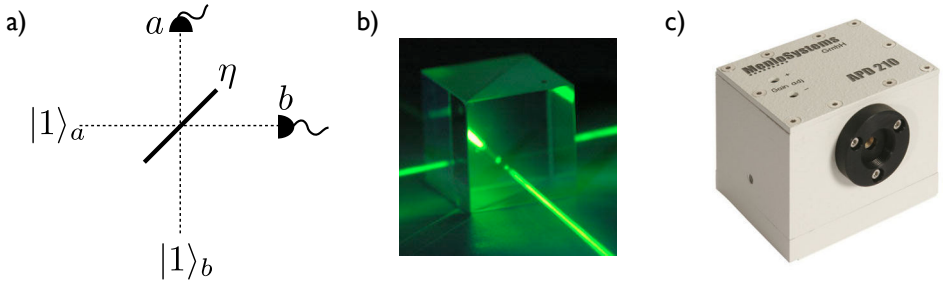


FIG. 1: a) Schematic of two photons incident on a beam splitter (BS) of reflectivity η, followed by single-photon detection in each output mode of the BS. b) Photo of real beam splitter cube (image: Beam Robotics). c) Photo of an avalanche photo diode (APD) detector, typically used in these types of experiments (image: Menlo Systems).

To model the effect, we begin by writing down a state consisting of two photons, one in each input mode of the beam splitter.

$$|\psi_{\text{IN}}\rangle = |1;i\rangle_a|1;j\rangle_b \tag{1}$$
$$= \hat{a}_i^\dagger \hat{b}_j^\dagger|0\rangle\,, \tag{2}$$

where a and b refer to the input modes and i and j describe a property of the photons that determines their distinguishability. These could refer to the polarization of the photon (e.g. horizontal or vertical), or the frequency of the photon, or the arrival time, or the transverse spatial mode (see e.g. [2]), or many other things. For the time being, we're going to keep it general and not make any assumptions about their level of distinguishability.

The effect of a beam splitter with reflectivity η can be modelled with a unitary $\hat{U}_{\text{BS}}$ that has the following action on the operators:

$$\hat{a}_i^\dagger \to i\sqrt{\eta}\hat{a}_i^\dagger + \sqrt{1-\eta}\hat{b}_i^\dagger; \tag{3}$$
$$\hat{b}_j^\dagger \to \sqrt{1-\eta}\hat{a}_j^\dagger + i\sqrt{\eta}\hat{b}_j^\dagger; \tag{4}$$

where the factor of i accounts for the phase acquired upon reflection from the BS. After passing through the BS, the output state is

$$|\psi_{\text{out}}\rangle = \hat{U}_{\text{BS}}|\psi_{\text{in}}\rangle \tag{5}$$
$$= (i\sqrt{\eta}\hat{a}_i^\dagger + \sqrt{1-\eta}\hat{b}_i^\dagger)(\sqrt{1-\eta}\hat{a}_j^\dagger + i\sqrt{\eta}\hat{b}_j^\dagger)|0\rangle \tag{6}$$
$$= \left(i\sqrt{\eta}\sqrt{1-\eta}\hat{a}_i^\dagger\hat{a}_j^\dagger + (1-\eta)\hat{a}_j^\dagger\hat{b}_i^\dagger - \eta\hat{a}_i^\dagger\hat{b}_j^\dagger + i\sqrt{\eta}\sqrt{1-\eta}\hat{b}_i^\dagger\hat{b}_j^\dagger\right)|0\rangle\,. \tag{7}$$

*Electronic address: branczyk@physics.utoronto.ca

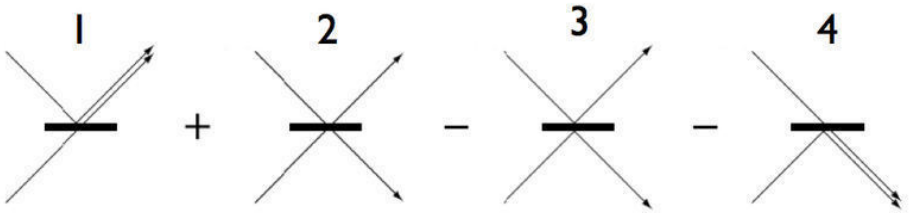


FIG. 2: Four different ways for two photons to interact with a beam splitter. Note that they used a different beam-splitter phase convention than we were using (image: Wikipedia)

We can draw a diagram representing the situation that lead to each of these four terms (FIG. 2).

Let's consider a beam splitter with reflectivity $\eta = 1/2$ and two photons of orthogonal polarizations; H and V. This means that the photons are *distinguishable*. In this scenario, the output state is

$$|\psi_{\text{out}}\rangle = \frac{1}{2}\Big(i\hat{a}_H^\dagger \hat{a}_V^\dagger + \hat{a}_V^\dagger \hat{b}_H^\dagger - \hat{a}_H^\dagger \hat{b}_V^\dagger + i\hat{b}_H^\dagger \hat{b}_V^\dagger\Big)|0\rangle \tag{8}$$

$$= \frac{i}{2}|1;H\rangle_a|1;V\rangle_a + \frac{1}{2}|1;V\rangle_a|1;H\rangle_b - \frac{1}{2}|1;H\rangle_a|1;V\rangle_b + \frac{i}{2}|1;H\rangle_b|1;V\rangle_b\,. \tag{9}$$

In HOM interference, we are interested in the coincidence probability of detecting one photon in each output port of the BS. We can see the only terms in Equation (9) where this happens are the two middle terms. Therefore, the coincidence probability of detecting one photon in each output mode is $p = 1/2$.

Now consider the same beam splitter, but two input photons with the same polarization mode H. This means that the photons are *in*distinguishable. In this scenario, the output state is

$$|\psi_{\text{out}}\rangle = \Big(i\hat{a}_H^\dagger \hat{a}_H^\dagger + \hat{a}_H^\dagger \hat{b}_H^\dagger - \hat{a}_H^\dagger \hat{b}_H^\dagger + i\hat{b}_H^\dagger \hat{b}_H^\dagger\Big)|0\rangle \tag{10}$$

$$= \Big(i\hat{a}_H^\dagger \hat{a}_H^\dagger + i\hat{b}_H^\dagger \hat{b}_H^\dagger\Big)|0\rangle \tag{11}$$

$$= \frac{1}{\sqrt{2}}|2;H\rangle_a + \frac{1}{\sqrt{2}}|2;H\rangle_b\,. \tag{12}$$

Note that the state is normalized because $(\hat{a}^\dagger)^n|0\rangle = \sqrt{n!}|n\rangle$. Here, the coincidence probability of detecting one photon in each output mode is $p = 0$. Incidentally, the state described in Equation (12) is a two-photon N00N state $|N\rangle|0\rangle + |0\rangle|N\rangle$ where $N = 2$ [3].

So we see that when two identical photons interfere on a symmetric beam splitter, the amplitudes for "both transmitted" and "both reflected" perfectly cancel out.

I. THE HOM DIP

In the previous section, we considered the situations where either two photons were completely distinguishable or completely indistinguishable. However, it is possible to tune the level of distinguishability. Consider a photon that has some sort of spectral bandwidth characterized by the spectral amplitude function $\phi(\omega)$. By controlling the time delays between these two photons, it is possible to tune their level of distinguishability. This is shown schematically in FIG. 3.

We can describe a photon with the spectral amplitude function $\phi(\omega)$ as follows:

$$|1;\phi\rangle = \int d\omega \phi(\omega)\hat{a}^\dagger(\omega)|0\rangle\,, \tag{13}$$

where $\hat{a}^\dagger(\omega)$ represents a creation operator acting on a single frequency mode ω. The state is normalized such that

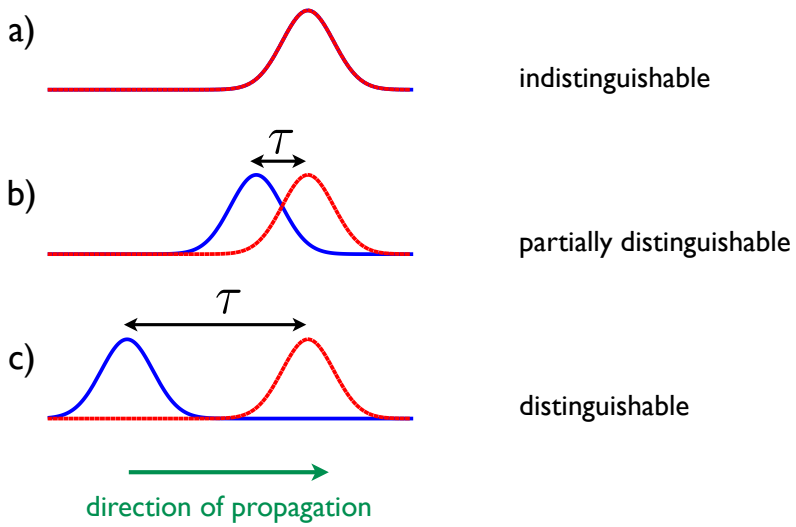


FIG. 3: Schematic demonstrating how changing the time delay between two photons with the same finite bandwidth affects their level of distinguishability.

$\int d\omega |\phi(\omega)|^2 = 1$. The two photon input state can therefore be written as

$$|\psi_{\text{in}}\rangle = |1;\phi\rangle_a |1;\phi\rangle_b \tag{14}$$

$$= \int d\omega_1 \phi(\omega_1)\hat{a}^\dagger(\omega_1) \int d\omega_2 \phi(\omega_2)\hat{b}^\dagger(\omega_2)|0\rangle \, . \tag{15}$$

We now want to introduce a time delay in one of the modes, as shown in FIG. 4. The time delay has the following action on the creation operators:

$$\hat{a}^\dagger(\omega) \to \hat{a}^\dagger(\omega)\mathrm{e}^{-i\omega\tau} \, . \tag{16}$$

We can therefore write the time delayed state as

$$|\psi_{\text{td}}\rangle = \int d\omega_1 \phi(\omega_1)\hat{a}^\dagger(\omega_1) \int d\omega_2 \phi(\omega_2)\hat{b}^\dagger(\omega_2)\mathrm{e}^{-i\omega_2\tau}|0\rangle \, . \tag{17}$$

We now want to send the photons through a beamsplitter. Let's keep it simple and restrict the reflectivity to $\eta = 1/2$.

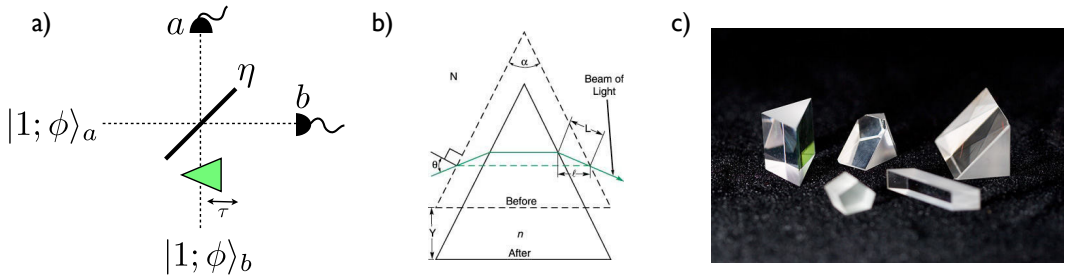


FIG. 4: a) Schematic demonstrating how a time delay is introduced in the form of a phase shift. Sliding the prism into and out of the photon path changes the effective distance that the photon has to travel therefore changing the time delay. b) Schematic demonstrating how the prism changes the effective path length (image: NASA). c) Photo of actual optical prisms (image: Creator Optics).

The action of the beam splitter is given by

$$\hat{a}^\dagger(\omega_1) \to i\frac{1}{\sqrt{2}}\hat{a}^\dagger(\omega_1) + \frac{1}{\sqrt{2}}\hat{b}^\dagger(\omega_1); \tag{18}$$

$$\hat{b}^\dagger(\omega_2) \to \frac{1}{\sqrt{2}}\hat{a}^\dagger(\omega_2) + i\frac{1}{\sqrt{2}}\hat{b}^\dagger(\omega_2); \tag{19}$$

where the BS acts on each frequency mode independently. This gives the output state:

$$|\psi_{\text{out}}\rangle = \frac{1}{2}\int d\omega_1\phi(\omega_1)\big(i\hat{a}^\dagger(\omega_1) + \hat{b}^\dagger(\omega_1)\big)\int d\omega_2\phi(\omega_2)\big(\hat{a}^\dagger(\omega_2) + i\hat{b}^\dagger(\omega_2)\big)\mathrm{e}^{-i\omega_2\tau}|0\rangle \tag{20}$$

$$= \frac{1}{2}\int d\omega_1\phi(\omega_1)\int d\omega_2\phi(\omega_2)\mathrm{e}^{-i\omega_2\tau}\big(i\hat{a}^\dagger(\omega_1)\hat{a}^\dagger(\omega_2) + \hat{a}^\dagger(\omega_2)\hat{b}^\dagger(\omega_1) - \hat{a}^\dagger(\omega_1)\hat{b}^\dagger(\omega_2) + i\hat{b}^\dagger(\omega_1)\hat{b}^\dagger(\omega_2)\big)|0\rangle\,. \tag{21}$$

In the previous section, it was simple to read off the coincidence probability from the state, but here, it is a bit more tricky so we should calculate it explicitly. We will model each detector as having a flat frequency response. The measurement operator describing the detector in mode $c = a, b$ is given by

$$\hat{M}_c = \int d\omega_c\hat{c}^\dagger(\omega_c)|0\rangle\langle 0|\hat{c}(\omega_c)\,. \tag{22}$$

The coincidence probability of getting one photon in each mode is given by

$$p_{ab} = \mathrm{Tr}[|\psi_{\text{out}}\rangle\langle\psi_{\text{out}}|\hat{M}_a \otimes \hat{M}_b] \tag{23}$$

$$= \langle\psi_{\text{out}}|\hat{M}_a \otimes \hat{M}_b|\psi_{\text{out}}\rangle \tag{24}$$

$$= \frac{1}{4}\int d\omega_1\phi^*(\omega_1)\int d\omega_2\phi^*(\omega_2)\mathrm{e}^{i\omega_2\tau}\langle 0|\big(i\hat{a}^\dagger(\omega_1)\hat{a}^\dagger(\omega_2) + \hat{a}^\dagger(\omega_2)\hat{b}^\dagger(\omega_1) - \hat{a}^\dagger(\omega_1)\hat{b}^\dagger(\omega_2) + i\hat{b}^\dagger(\omega_1)\hat{b}^\dagger(\omega_2)\big)$$
$$\times\int d\omega_a\int d\omega_b\hat{a}^\dagger(\omega_a)\hat{b}^\dagger(\omega_b)|0\rangle\langle 0|\hat{a}(\omega_a)\hat{b}(\omega_b) \tag{25}$$
$$\times\int d\omega_1'\phi(\omega_1')\int d\omega_2'\phi(\omega_2')\mathrm{e}^{-i\omega_2'\tau}\big(i\hat{a}^\dagger(\omega_1')\hat{a}^\dagger(\omega_2') + \hat{a}^\dagger(\omega_2')\hat{b}^\dagger(\omega_1') - \hat{a}^\dagger(\omega_1')\hat{b}^\dagger(\omega_2') + i\hat{b}^\dagger(\omega_1')\hat{b}^\dagger(\omega_2')\big)|0\rangle$$

$$= \frac{1}{4}\int d\omega_a\int d\omega_b\big(\phi^*(\omega_a)\phi^*(\omega_b)\mathrm{e}^{i\omega_a\tau} - \phi^*(\omega_a)\phi^*(\omega_b)\mathrm{e}^{i\omega_b\tau}\big)\big(\phi(\omega_a)\phi(\omega_b)\mathrm{e}^{-i\omega_a\tau} - \phi(\omega_a)\phi(\omega_b)\mathrm{e}^{-i\omega_b\tau}\big) \tag{26}$$

$$= \frac{1}{4}\int d\omega_a\int d\omega_b|\phi(\omega_a)\phi(\omega_b)|^2\big(\mathrm{e}^{i\omega_a\tau} - \mathrm{e}^{i\omega_b\tau}\big)\big(\mathrm{e}^{-i\omega_a\tau} - \mathrm{e}^{-i\omega_b\tau}\big) \tag{27}$$

$$= \frac{1}{4}\int d\omega_a\int d\omega_b|\phi(\omega_a)\phi(\omega_b)|^2\big(2 - \mathrm{e}^{i(\omega_b-\omega_a)\tau} - \mathrm{e}^{i(\omega_a-\omega_b)\tau}\big)\,. \tag{28}$$

Let's take a look at what p_{ab} looks like given photons with Gaussian spectral profiles, $\phi(\omega) = \frac{e^{-\frac{(\omega-\omega_0)^2}{4\sigma^2}}}{(2\pi)^{1/4}\sqrt{\sigma}}$. The normalization was chosen such that $\int d\omega|\phi(\omega)|^2 = 1$.

$$\begin{aligned} p_{ab} = {} & \frac{1}{4\pi\sigma^2}\int d\omega_a e^{-\frac{(\omega_a-\omega_0)^2}{2\sigma^2}}\int d\omega_b e^{-\frac{(\omega_b-\omega_0)^2}{2\sigma^2}} \\ & - \frac{1}{8\pi\sigma^2}\int d\omega_a e^{-\frac{(\omega_a-\omega_0)^2}{2\sigma^2}}\mathrm{e}^{-i\omega_a\tau}\int d\omega_b e^{-\frac{(\omega_b-\omega_0)^2}{2\sigma^2}}\mathrm{e}^{i\omega_b\tau} \\ & - \frac{1}{8\pi\sigma^2}\int d\omega_a e^{-\frac{(\omega_a-\omega_0)^2}{2\sigma^2}}\mathrm{e}^{i\omega_a\tau}\int d\omega_b e^{-\frac{(\omega_b-\omega_0)^2}{2\sigma^2}}\mathrm{e}^{-i\omega_b\tau} \end{aligned} \tag{29}$$

$$p_{ab} = \frac{1}{2} - \frac{1}{2}\mathrm{e}^{-\sigma^2\tau^2}\,. \tag{30}$$

This is plotted in FIG. 5 a) in blue.

Let's see what happens when we take $\phi(\omega) = \frac{\sqrt{\sigma}}{\sqrt{\pi}}\mathrm{sinc}(\sigma\omega)$. Why am I interested in this particular shape? One of the most common methods of generating single photons relies on the process of spontaneous parametric down conversion (SPDC). This is a nonlinear process where an intense laser beam is used to pump a nonlinear crystal. Every now-and-then, a photon from the pump beam spontaneously "splits" into two lower energy photons which satisfy energy

and momentum conservation with the pump photon. A pair of photons come out of the crystal, and conditional on detecting a photon in one mode, we know that we have a photon in the other mode. This is called "heralding". It turns out that the spectral profile of the down-converted photons is related to the Fourier transform of the non-linearity profile of the crystal, which for a constant nonlinearity, takes on a "top-hat". The Fourier transform of a "top hat" function is a sinc. Using a sinc spectral profile gives the coincidence probability

$$p_{ab} = \frac{((\tau - 2\sigma)\text{sgn}(\tau - 2\sigma) + (2\sigma + \tau)\text{sgn}(2\sigma + \tau) - 2\tau\text{sgn}(\tau))^2}{16\sigma^2}, \tag{31}$$

where

$$\text{sgn(x)} = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}. \tag{32}$$

This is plotted in FIG. 5 a) in red. Even though I said that heralded single photons generated via SPDC take on a sinc profile, it's a bit more complicated than that. I won't go into details about this now, but it is to do with spectral correlations and reduced purity. I am pointing this out so you don't get confused if you never see the red curve in FIG. 5 a) in any real experiments.

However, we can model a realistic experiment if we take the two photons generated via SPDC and rather than heralding on one of them, we send each one into one input mode of the beam splitter. The output state from an SPDC source is often highly correlated in frequency and we can write it as

$$|\psi_{\text{SPDC}}\rangle = \int d\omega_a \int d\omega_b f(\omega_a, \omega_b)\delta(\omega_a + \omega_b - \omega_p)\hat{a}^\dagger(\omega_a)\hat{b}^\dagger(\omega_b)|0\rangle \tag{33}$$

$$= \int d\omega' f'(\omega')\hat{a}^\dagger\Big(\frac{\omega_p}{2} + \omega'\Big)\hat{b}^\dagger\Big(\frac{\omega_p}{2} - \omega'\Big)|0\rangle \tag{34}$$

where ω_p is the frequency of the pump beam, $\omega' = \omega_a - \frac{\omega_p}{2}$, and I don't want to go into how $f(\omega_a, \omega_b)$ relates to $f'(\omega')$. The coincidence probability will be given by

$$p_{ab} = \frac{1}{4}\int d\omega' |f'(\omega')|^2 \big(2 - \text{e}^{-i2\omega'\tau} - \text{e}^{i2\omega'\tau}\big). \tag{35}$$

Typically, $f'(\omega')$ will be the sinc of a non-linear function of ω', but there are easily-accessible regimes where we can approximate it be the sinc of a linear function of ω': $f'(\omega') = \frac{\sqrt{\sigma}}{\sqrt{\pi}}\text{sinc}(\sigma\omega')$. This gives the coincidence probability

$$p_{ab} = \frac{1}{2} - \frac{1}{4\sigma}((\sigma - \tau)\text{sgn}(\sigma - \tau) + (\sigma + \tau)\text{sgn}(\sigma + \tau) - 2\tau\text{sgn}(\tau)). \tag{36}$$

This is plotted in FIG. 5 b) in red. When the down-converted photons are spectrally correlated as in Equation (34), the shape of the HOM dip will be convolution of the spectral function with itself. One place where you can see

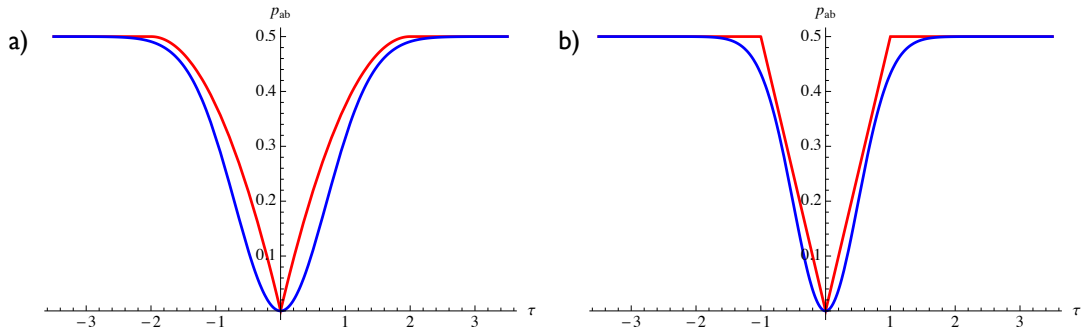


FIG. 5: a) Coincidence probability using $\sigma = 1$ for Equations (30) in blue and (31) in red. b) Coincidence probability using $\sigma = 1$ for Equations (37) in blue and (36) in red.

experimental data for an inverse triangle like this is in a recent paper of mine [4]. The original HOM paper [1] also used this type of source, but why didn't they see an inverse triangle? If you look at the figure in the paper, they have UV filters in front of the beam splitter. These filters would have had a Gaussian profile and been much more narrow than the original sinc function, resulting in photons with a Gaussian profile. The coincidence probability for a spectrally-correlated Gaussian function $f'(\omega') = \frac{e^{-\frac{(\omega'-\omega_0)^2}{4\sigma^2}}}{(2\pi)^{1/4}\sqrt{\sigma}}$ is

$$p_{ab} = \frac{1}{2} - \frac{1}{2}e^{-2\sigma^2\tau^2}\,. \tag{37}$$

This is plotted in FIG. 5 b) in blue.

One more fun fact. The visibility of the HOM dip, given by

$$V = \frac{\max - \min}{\max + \min}\,, \tag{38}$$

is related to the density operators of the two photons by

$$V = \mathrm{Tr}[\hat{\rho}_a\hat{\rho}_b]\,. \tag{39}$$

If the two photons are the same, i.e. $\hat{\rho}_a = \hat{\rho}_b$, then the visibility is equal to the purity of the photons

$$V = \mathrm{Tr}[\hat{\rho}_a^2] = P\,. \tag{40}$$

Remember above when I said that the situation of heralded photons is a bit complicated due to spectral correlations? Have a look at the two-photon state in Equation (34). This state is not only correlated in frequency, but it is also entangled in frequency. So if we use a detector with a flat (or even just broad) spectral response, to detect one of the photons, we will have to trace out over all frequencies to write down the state of the remaining photon. This is going to be mixed in frequency, and a mixed state is not pure. If you are interested in this, you can read about it in a paper I wrote a couple of years ago [5].

People go to a lot of effort to try and minimize these spectral correlations. One common way is to just filter the crap out of the photons. This has disadvantages like reducing your count rates. It also becomes problematic if you are trying to create higher photon number Fock states because the filters then introduce mixing in photon-number. Another method is to try and engineer the spectral function. See for example [6]. So let's say you have gone to all this effort to make a state that is less mixed and you want to check its purity. One way to do this is to prepare two of these states and then interfere them on a beam splitter and perform a HOM interference experiment. By looking at the visibility, you can determine how pure your states are.

---

[1] C. K. Hong, Z. Y. Ou, and L. Mandel, "Measurement of subpicosecond time intervals between two photons by interference," Phys. Rev. Lett. **59**, 2044 (1987).

[2] N. K. Langford *et al.*, "Measuring Entangled Qutrits and Their Use for Quantum Bit Commitment," Phys. Rev. Lett. **93**, 053601 (2004).

[3] H. Lee, P. Kok, and J. P. Dowling, "A quantum Rosetta stone for interferometry," Journal of Modern Optics **49**, 2325 (2002).

[4] A. M. Brańczyk *et al.*, "Engineered optical nonlinearity for quantum light sources," Opt. Express **19**, 55 (2011).

[5] A. M. Brańczyk *et al.*, "Optimized generation of heralded Fock states using parametric down-conversion," New Journal of Physics **12**, 063001 (2010).

[6] P. J. Mosley *et al.*, "Heralded Generation of Ultrafast Single Photons in Pure Quantum States," Phys. Rev. Lett. **100**, 133601 (2008).

Observation of a "quantum eraser": A revival of coherence in a two-photon interference experiment

Paul G. Kwiat, Aephraim M. Steinberg, and Raymond Y. Chiao
Department of Physics, University of California, Berkeley, California 94720
(Received 23 December 1991)

We have observed an effect known as a quantum eraser, using a setup similar to one previously employed to demonstrate a violation of Bell's inequalities. In this effect, an interfering system is first rendered incoherent by making the alternate Feynman paths which contribute to the overall process distinguishable; with our apparatus this is achieved by placing a half wave plate in one arm of a Hong-Ou-Mandel interferometer so as to rotate the polarization of the light in that arm by 90°. This adds information to the system, in that polarization is a new parameter which serves to label the path of a given photon, even after a recombining beam splitter. The quantum "eraser" removes this information from the state vector, *after* the output port of the interferometer, but in time to cause interference effects to reappear upon coincidence detection. For this purpose, we use two polarizers in front of our detectors. We present experimental results showing how the degree of erasure (which determines the visibility of the interference) depends on the relative orientation of the polarizers, along with theoretical curves. In addition, we show how this procedure may do more than merely erase, in that the act of "pasting together" two previously distinguishable paths can introduce a new relative phase between them.

PACS number(s): 03.65.Bz, 42.50.Wm, 42.50.Dv

I. INTRODUCTION

Interference is arguably the most fundamental effect in quantum mechanics, the Young's two-slit experiment being the canonical manifestation of complementarity. As discussed by Bohr in his classic dialogue with Einstein, if one tries to measure the momentum of the recoiling slits to determine which slit the particle (e.g., photon, electron, atom, etc.) traversed, then one will observe particle-like behavior, and no interference (wavelike behavior) will arise. In a variant of this example, Feynman proposed to "watch" the passage of an electron through a particular slit by placing a light source immediately after the slits and scattering photons off the electron [1]. Even if one does not observe the scattered light, the electron interference will be washed out (whenever the light is scattered sufficiently to carry unambiguous information about which slit was traversed).

This loss of interference is commonly interpreted as arising from uncontrollable, irreversible interactions of the interfering system with the environment, which often takes the form of a macroscopic apparatus [2]. The resulting measurement "reduces" the wave function of the interfering system, including any phase information carried by the particle, thereby eliminating the possibility of interference. In Feynman's example, scattering light off the electron changes its center-of-mass wave function in an uncontrollable manner, removing the phase coherence between the two paths. While it is true that many measurements are of this sort, there are situations where the measurement process need not be so uncontrollable. In these cases it is more helpful to view the loss of coherence as due to an entanglement of the system wave function with that of the measuring apparatus [3]. We will show below how this destroys interference. We can also understand all of these results in terms of Feynman's rules for calculating probabilities: (i) Probability amplitudes of *indistinguishable* paths are summed, then absolute squared, to yield the probability; this leads to interference terms. (ii) Probabilities of *distinguishable* paths are summed, yielding no interference. Thus it is the distinguishability of alternative paths which prevents interference. When information exists about which way *(welcher Weg)* the particle went, the paths are distinguishable, and no interference is possible. Interference may be regained, however, if one somehow manages to "erase" the distinguishing information. This is the central concept of the quantum eraser [4]. A nice review article on this and the related ideas of complementarity recently appeared in Ref. [5].

Scully *et al.* [6] discussed a simple experiment to see this effect, in which an atom is sent through a Stern-Gerlach interferometer [7]. Upon measurement of the atom's passage through one arm of the interferometer, the interference is made to vanish. This is true even if the measuring apparatus does not change the spin state of the atom, or affect the center-of-mass part of its wave function. Unfortunately, detailed calculations of the proposed experiment made clear that it would probably not be feasible in practice, due to the experimental difficulty of controlling the fields to the degree necessary to observe interference, even in the absence of a *welcher Weg* detector [8,9]. Another proposal using a two-slit type interference of neutrons, with micromaser cavities as *welcher Weg* detectors was also deemed very difficult [10]. To date, the most promising of the proposed experiments on particles involve the interference manifested in the quantum beat phenomenon [10,11]. However, in addition to also being rather difficult, though possibly feasible, these experiments suffer the conceptual disadvantage that there

are not actually spatially distinct paths as in the double-slit versions. A somewhat different scheme with photons using optical parametric amplifiers as *welcher Weg* detectors was also recently proposed [12]. However, only partial erasure is possible, and even then the practical obstacle of lack of near-unit efficiency photon detectors must be overcome.

As described below, we have performed a comparatively simple experiment involving the interference of photons, which we believe demonstrates all the salient features of the quantum-eraser phenomenon [13]. The *welcher Weg* information is stored in the polarization states of the photons, which are made distinguishable by means of a half wave plate. The erasure is performed by means of polarizers placed before the detectors. In Sec. II we briefly describe our setup and the two-photon light source in our system. The nonclassical interference effect we employ is reviewed in Sec. III, introducing the necessary quantum-field-theory formalism. The loss of this interference is investigated theoretically in Sec. IV, and experimental results are shown. A simple derivation of the quantum-eraser effect is presented in Sec. V, as are experimental results. We show that not only is it possible to recover interference, but also to change the form of the interference pattern. A comparison of our experiment with the various proposals is made in Sec. VI, along with a discussion of its relation to some Bell's inequalities experiments and other two-photon experiments. The main results are summarized in Sec. VII. Throughout we will try to explain the phenomena both at an intuitive level using Feynman's rules, and also at a more formal level, using the established quantum field-theoretic approach to photodetection and correlation.

II. EXPERIMENTAL SETUP

A schematic of our apparatus is shown in Fig. 1. Highly correlated pairs of photons are produced in a nonlinear crystal via the process of spontaneous parametric down-conversion of a uv pump beam, generated by an argon-ion laser. The $\chi^{(2)}$ nonlinear medium is a 10-cm-long potassium dihydrogen phosphate (KDP) crystal, with the direction of the optic axis cut at 53° with respect to the end faces, for type-I phase matching. The 351.1-nm pump photons are spontaneously "split" into conjugate photons (conventionally denoted "signal" and "idler"), which are horizontally polarized. With irises (2 mm diameter) and filters (10 nm bandwidth) at our detectors, we select out the nearly degenerate pairs at 702.2 nm. Each detector consists of an RCA C30902S avalanche photodiode (APD) cooled to −18 °C, whose output is fed into an EG&G-Ortec Model 583 constant fraction discriminator.

This particular light source has been well studied, and has been used previously in other investigations of fundamental quantum optical phenomena [14]. In one such configuration, the Hong-Ou-Mandel interferometer [Fig. 2(a)], the two correlated photons are brought back together by means of two mirrors, so that they impinge simultaneously on the surface of a translatable beam splitter [15]. We measure singles and coincidence rates at the output ports (using a Stanford Research Systems SR400 Gated Photon Counter). As explained in the following section, if the beam splitter is placed such that the two photons reach it essentially simultaneously (i.e., within their coherence times), interference will result, in such a way that both photons always exit the same port of the beam splitter. Thus a null in the coincidence rate appears as the path length of one of the arms is slowly scanned, even though the singles rates remain unchanged. The width of the dip [≈ 40 μm full width at half maximum (FWHM)] is determined by the filters in front of the detectors. In practice, it was preferable to vary the relative path length using an "optical trombone" in one arm of the interferometer, thus avoiding the problem of lateral walkoff associated with translating the recombining beam splitter directly. (One can show that dispersive effects of the trombone prism have essentially no effect on the interference dip [16].) Translation of the prism was effected by a Burleigh Inchworm piezoelectric motion system; a Heidenhain optical encoder yielded a position resolution of 0.1 μm.

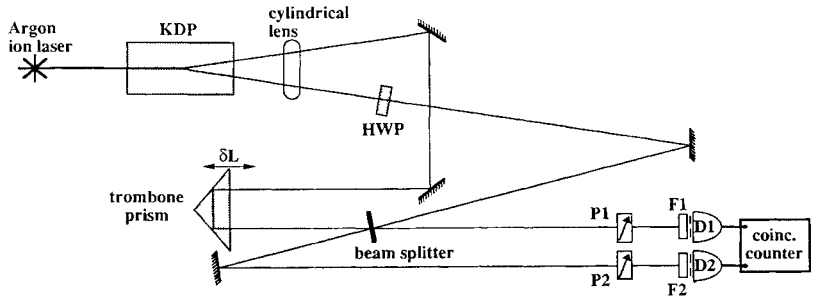


FIG. 1. Schematic of experiment to observe quantum eraser. D1 and D2 are avalanche photodiodes, P1 and P2 are polarizers, F1 and F2 are bandpass filters, and HWP is a half wave plate whose optic axis is at an angle $\phi/2$ to the horizontal polarization of the down-converted beams.

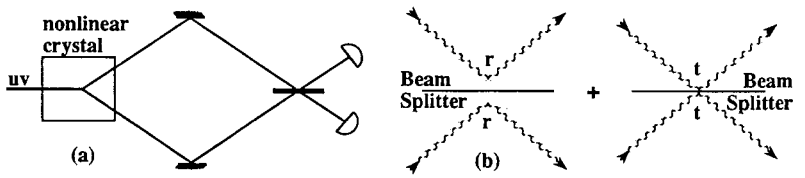


FIG. 2. (a) Simplified setup for a Hong-Ou-Mandel interferometer. (b) Feynman paths for coincidence detection.

III. HONG-OU-MANDEL INTERFERENCE

One can explain the coincidence null at zero path-length difference using the Feynman rules for calculating probabilities. The two indistinguishable processes [Fig. 2(b)] which lead to coincidence detection in the above setup are both photons being reflected at the beam splitter and both photons being transmitted. For simplicity, suppose we have a 50:50 beam splitter, and choose the amplitude transmission coefficient to be real. The Feynman amplitudes are then $r\times r$ and $t\times t$, and the probability of a coincidence detection is

$$P_c=|r\times r+t\times t|^2=\left|\frac{i}{\sqrt{2}}\times\frac{i}{\sqrt{2}}+\frac{1}{\sqrt{2}}\times\frac{1}{\sqrt{2}}\right|^2=0\ , \tag{1}$$

where the factors of i come from the phase shift upon reflection at a beam splitter. When the path-length difference is greater than the coherence length of the photons (i.e., when the photon wave packets no longer overlap at the beam splitter), there is no such cancellation effect and coincidence events occur one half of the time, since each photon *individually* has a 50% chance of being reflected or transmitted.

More formally, we write the wave function after the beam splitter in terms of Fock states:

$$\begin{aligned}|\psi\rangle_{\Delta x=0}&=\tfrac{1}{2}[\,|1_1 1_2\rangle+i|2_1 0_2\rangle+i|0_1 2_2\rangle-|1_1 1_2\rangle\,]\\&=\frac{i}{2}[\,|2_1 0_2\rangle+|0_1 2_2\rangle\,]\ ,\end{aligned} \tag{2}$$

where the subscripts denote the propagation modes to the two detectors, and the subscript of $|\psi\rangle$ indicates zero path-length difference. As discussed in earlier works [14–17], the conjugate photons actually have a relatively broadband frequency distribution, which is determined in practice by irises and filters in front of the detectors. However, since we operate near degeneracy, and since we are considering zero path-length difference, this generalization is an unnecessary complication for our purposes. According to the standard theory of photodetection and photon correlation [18], the coincidence counting rate is given by the fourth-order correlation function

$$\begin{aligned}P_c&\approx G^{(2)}(t_1,t_2;t_2,t_1)\\&=\langle\psi|\hat{E}_1^{(-)}(t_1)\hat{E}_2^{(-)}(t_2)\hat{E}_2^{(+)}(t_2)\hat{E}_1^{(+)}(t_1)|\psi\rangle\ ,\end{aligned} \tag{3}$$

where, omitting irrelevant normalization constants, the positive- and negative-frequency field operators are defined in terms of creation and annihilation operators as

$$\hat{E}_j^{(+)}(t_j)=\int d\omega\,\hat{a}_j(\omega)e^{-i\omega t_j}$$

and (4)

$$\hat{E}_j^{(-)}(t_j)=\int d\omega\,\hat{a}_j^{\dagger}(\omega)e^{+i\omega t_j}\quad(j=1,2)\ .$$

We have neglected polarization for the moment, which is justified because the photons are both horizontally polarized, and our detectors are polarization independent. At zero path-length difference the integrals over frequency contribute only an overall normalization factor [19]. We can then understand the essence of the measurement described by Eq. (3) by considering the reduced operator formed by creation and annihilation operators for the two detector modes: $\hat{P}_{c,\mathrm{red}}\equiv\hat{a}_1^{\dagger}\hat{a}_2^{\dagger}\hat{a}_2\hat{a}_1$. Clearly, $\hat{P}_{c,\mathrm{red}}$ gives zero when it operates on $|\psi\rangle_{\Delta x=0}$, as given in Eq. (2).

When the path-length difference is greater than the coherence length of the down-converted photons ($\Delta x\gg\tau_c$), the "transmission-transmission" and "reflection-reflection" coincidence possibilities are in principle distinguishable, so they do not interfere. In this limit, we find

$$\begin{aligned}P_c(\Delta x\gg\tau_c)&\approx\frac{1}{4}\langle 1_1 1_2|\hat{P}_{c,\mathrm{red}}|1_1 1_2\rangle\\&\quad+\frac{i^4}{4}\langle 1_1 1_2|\hat{P}_{c,\mathrm{red}}|1_1 1_2\rangle=\frac{1}{2}\ .\end{aligned} \tag{5}$$

(The reduction by a factor of 2 reflects the fact that we are only considering *coincidence* counts, not cases where both photons go to the same detector.)

This demonstrates the coincidence dip at zero path-length difference to the beam splitter. Note that the singles rate at either detector, given by $P_s\approx G^{(1)}(t_j;t_j)=\langle\psi|\hat{E}_j^{(-)}(t_j)\hat{E}_j^{(+)}(t_j)|\psi\rangle$, does not show this dependence on path-length difference. It has been shown that as long as the visibility of the coincidence dip is greater than 50%, no semiclassical field theory can account for the observed interference [20].

IV. LOSS OF INTERFERENCE

In the spirit of the Feynman two-slit experiment, we ask if one can perform a "measurement" on the photons which will yield which-way information. Of course, we could place an APD or photomultiplier directly in one of

the input arms of the interferometer, but then lack of coincidence is a trivial consequence. We consider instead what happens when a half wave plate at an angle $\phi/2$ to the horizontal is inserted into one input arm of the interferometer, as depicted in Fig. 1 (and we adjust the trombone to compensate for the optical path length of the wave plate). The polarization of the photon in that arm is then rotated by ϕ, making the two Feynman paths partially distinguishable, thereby reducing the amount of interference. The degree to which the interference is lost depends on the angle ϕ, and is calculated below. In the extreme case ($\phi/2=45°$) the polarization states of the two different photons reaching the beam splitter are orthogonal. The two paths are now *completely* distinguishable and the amplitudes are squared before being summed. The result: no interference. These effects are shown in Fig. 3(a).

In terms of our earlier formalism, we have entangled the number-state basis wave function with polarization information:

$$|\psi\rangle_{\Delta x=0}=\tfrac{1}{2}[|1_1 1_2\rangle\otimes|H_1(H+\phi)_2\rangle - |1_1 1_2\rangle\otimes|(H+\phi)_1 H_2\rangle]\,, \tag{6a}$$

where the notation H_j indicates that the photon reaching detector j is horizontally polarized, and $(H+\phi)_j$ indi-

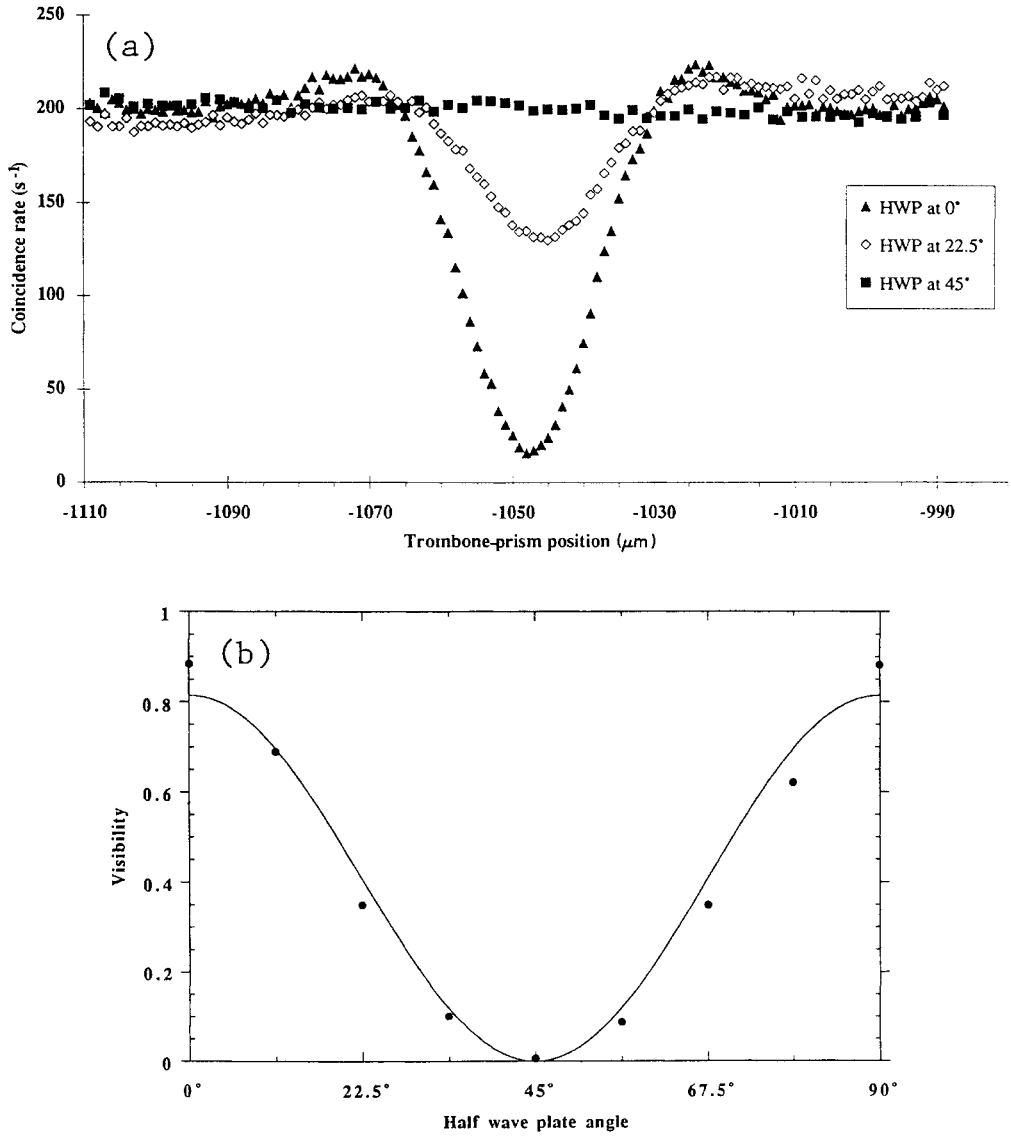


FIG. 3. (a) Profile of interference dip in coincidence rate for three wave plate orientations. (Accidental coincidences have been subtracted, and rates far from dip have been normalized to the same value.) Note that the interference effect is seen to vanish when the wave plate is at 45°, i.e., when the input ports to the beam splitter are made distinguishable. (b) Visibility as a function of half wave plate angle. The solid line is a fit to theory, with maximum visibility as the free parameter. The experimental points do not lie exactly on the same curve because slight fluctuations in alignment affect the visibility.

cates that the photon is polarized at an angle ϕ to the horizontal. We have already omitted the kets in which both photons go to the same detector, since we are interested here in coincidence rates. We introduce the following simplified notation:

$$|\psi\rangle_{\Delta x=0}=\tfrac{1}{2}[|1_1^H 1_2^{H+\phi}\rangle-|1_1^{H+\phi}1_2^H\rangle]\ . \tag{6b}$$

To account for polarization, the field operators may now be generalized to the vector operators $\hat{\mathbf{a}}_j^\dagger=\hat{a}_{j,H}^\dagger\boldsymbol{\epsilon}_{j,H}+\hat{a}_{j,V}^\dagger\boldsymbol{\epsilon}_{j,V}$ and $\hat{\mathbf{a}}_j=\hat{a}_{j,H}\boldsymbol{\epsilon}_{j,H}+\hat{a}_{j,V}\boldsymbol{\epsilon}_{j,V}$, where $\boldsymbol{\epsilon}_{j,H}$ and $\boldsymbol{\epsilon}_{j,V}$ are orthonormal polarization vectors (associated with detector index $j=1,2$) in the horizontal and vertical directions, respectively. The new reduced operator relevant for coincidence counting is then

$$\hat{P}'_{c,\mathrm{red}}\equiv\sum_{\lambda_1,\lambda_2=H,V}a^\dagger_{1,\lambda_1}a^\dagger_{2,\lambda_2}a_{2,\lambda_2}a_{1,\lambda_1}=\sum_{\lambda_1,\lambda_2=H,V}a^\dagger_{1,\lambda_1}a_{1,\lambda_1}a^\dagger_{2,\lambda_2}a_{2,\lambda_2}$$

$$=(\hat{\mathbf{a}}_1^\dagger\cdot\hat{\mathbf{a}}_1)(\hat{\mathbf{a}}_2^\dagger\cdot\hat{\mathbf{a}}_2)=(\hat{a}^\dagger_{1,H}\hat{a}_{1,H}+\hat{a}^\dagger_{1,V}\hat{a}_{1,V})(\hat{a}^\dagger_{2,H}\hat{a}_{2,H}+\hat{a}^\dagger_{2,V}\hat{a}_{2,V})\ . \tag{7}$$

Using the expansion $|1_j^{H+\phi}\rangle=|1_j^H\rangle\cos\phi+|1_j^V\rangle\sin\phi$, we find

$$P_c(0)\approx\langle\psi|\hat{P}'_{c,\mathrm{red}}|\psi\rangle_{\Delta x=0}=\tfrac{1}{2}\sin^2\phi\ . \tag{8}$$

When the path-length difference is greater than the photon coherence length, the calculation of the coincidence rate proceeds as before:

$$P_c(\Delta x\gg\tau_c)\approx\frac{1}{4}\langle 1_1^H 1_2^{H+\phi}|\hat{P}'_{c,\mathrm{red}}|1_1^H 1_2^{H+\phi}\rangle$$

$$+\frac{i^4}{4}\langle 1_1^{H+\phi}1_2^H|\hat{P}'_{c,\mathrm{red}}|1_1^{H+\phi}1_2^H\rangle=\frac{1}{2}\ . \tag{9}$$

The visibility of the dip, defined as $V=\{P_c(\Delta x\gg\tau_c)-P_c(\Delta x=0)\}/P_c(\Delta x\gg\tau_c)$, has the form $\cos^2\phi$. The experimental demonstration of this relationship is shown in Fig. 3(b). The lack of perfect visibility even at $\phi=0$ results from imperfect alignment of the system, so that the signal and idler modes leaving the beam splitter are already somewhat distinguishable, regardless of their polarization.

Following Scully, Shea, and McCullen [6], one can also approach the loss of interference in terms of the density matrix. When the photon-propagation states are entangled with the polarization states, the density matrix of the system is enlarged. It still represents a pure state, however, with the quantum coherence of the entanglement manifested in the off-diagonal matrix elements. The "collapse" to a mixed state occurs when we trace over the polarization degrees of freedom, i.e., when we detect the final propagation direction of the photons *irrespective of their polarization*. In this case the reduced density matrix has only diagonal elements, because the polarization states $|H_j\rangle$ and $|V_j\rangle$ (which are essentially the "environment" for our purposes) are orthogonal. This is precisely the method of decoherence recently discussed by Zurek, although he focused on an environment which was either "uncontrollable" or possessed a large number of degrees of freedom [21]. In either case, the process is effectively irreversible, which is certainly not the case in our experiment, as we shall see presently.

V. QUANTUM ERASER

The essence of the quantum eraser can be understood relatively easily now in Feynman's language of distinguishability. As we have seen, with the half wave plate at 45°($\phi=90°$) the two paths leading to coincidence detection ("reflection-reflection" and "transmission-transmission") are distinguishable; they leave the light in each port in a different polarization state. For this reason, their probabilities are to be added incoherently, and there is no interference term. What if one could erase the information carried by the polarization, thus making the final states indistinguishable? This is precisely what happens when one places polarizers oriented at 45° to the horizontal in both output ports of the interferometer. (See Fig. 1.)

Both paths can lead to coincidence detection, and to the same final state. Therefore their probability *amplitudes* are added, thus reviving the Hong-Ou-Mandel interference dip at equal path length. Note that the insertion of a polarizer in only one of the output ports is insufficient to erase the distinguishability of the final states, because the photon in the other port still possesses *welcher Weg* information. Hence the only effect of a single polarizer is to reduce both the singles and the coincidence count rate by half.

The editing accomplished with two polarizers is not limited to erasure, as can be motivated by the following observation. Regardless of the rest of the system, the light in port 2 can always be broken up into its orthogonal polarization components. But we just saw that with both P1 and P2 at 45°, the interference dip reappeared (albeit attenuated by a factor of 4). Furthermore, we argued that the coincidence profile with polarizer P1 at 45° and P2 not in place was a flat line. It is clear then that if P1 is placed at 45° and P2 is placed at −45°, instead of a dip, a peak centered at zero path-length difference will now appear. These theoretical results are presented in Fig. 4(a) and our data in Fig. 4(b). As shown below, this is merely a specific instance of a more general property of the two-photon state emitted by the interferometer. (It should be noted that the possibility of producing a *peak*

at zero path-length difference greatly aids the alignment process for the Hong-Ou-Mandel interferometer.)

We now present a simplified analysis, limiting ourselves to the case $\phi=90°$. A more complete calculation is presented in the Appendix. The output of the interferometer is given by a special case of the entangled state in Eq. (6b):

$$|\psi(0)\rangle=\tfrac{1}{2}[|1_1^H 1_2^V\rangle-|1_1^V 1_2^H\rangle]\,, \tag{10}$$

where we have again dropped terms which could not lead to coincidence counts. Detection of a photon at one port

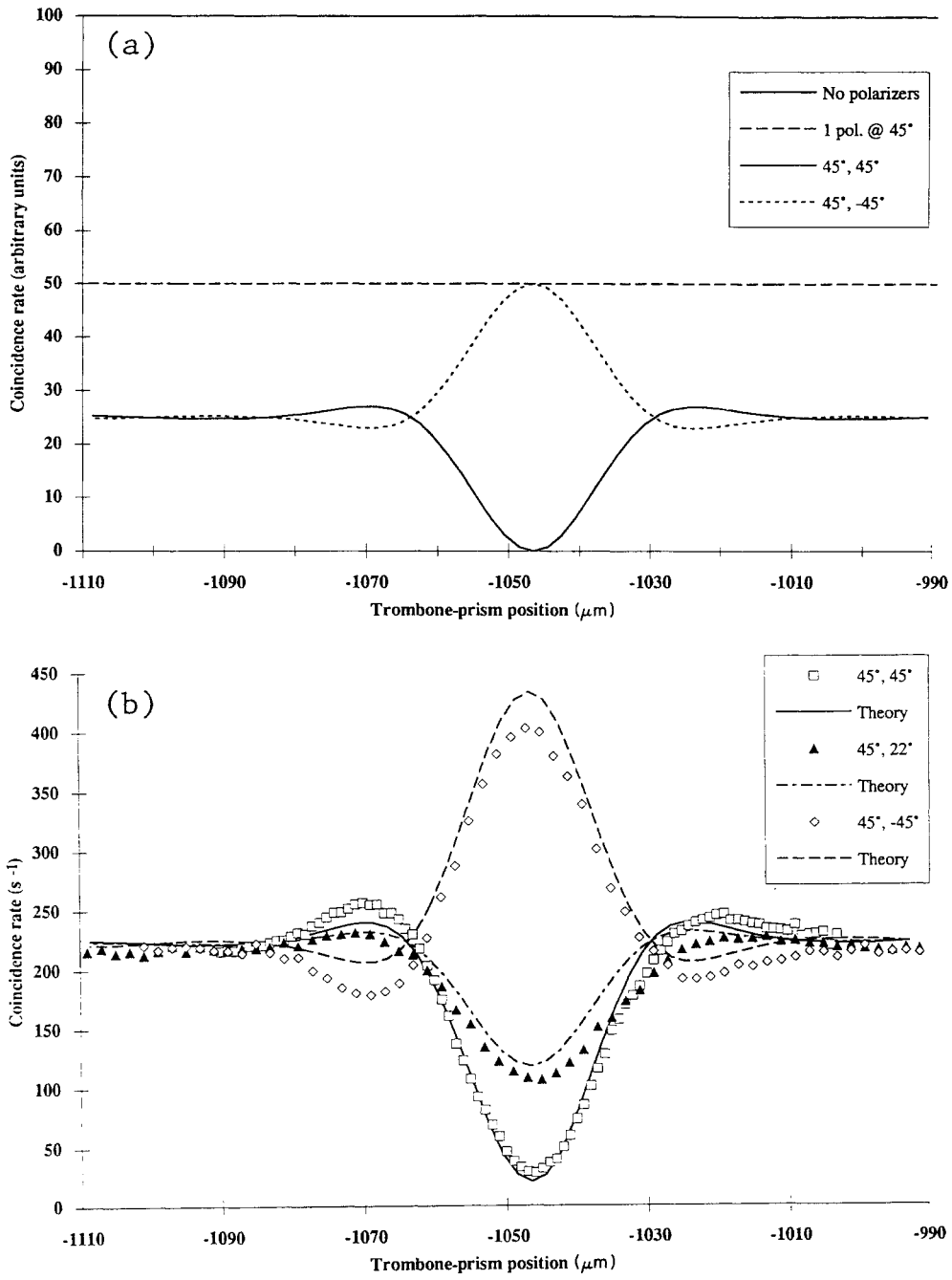


FIG. 4. (a) Theoretical curves showing how two polarizers at appropriately chosen angles can erase distinguishability, restoring an interference pattern. (b) Experimental data and scaled theoretical curves (adjusted to fit observed visibility of 91%) with polarizer 1 at 45° and polarizer 2 at various angles. Far from the dip, there is no interference and the angle is irrelevant. At the dip, the nonlocal collapse of the polarization of photon 2 causes us to observe sinusoidal variation as predicted in Eq. (13). [Normalization is the same as in Fig. 3(a).]

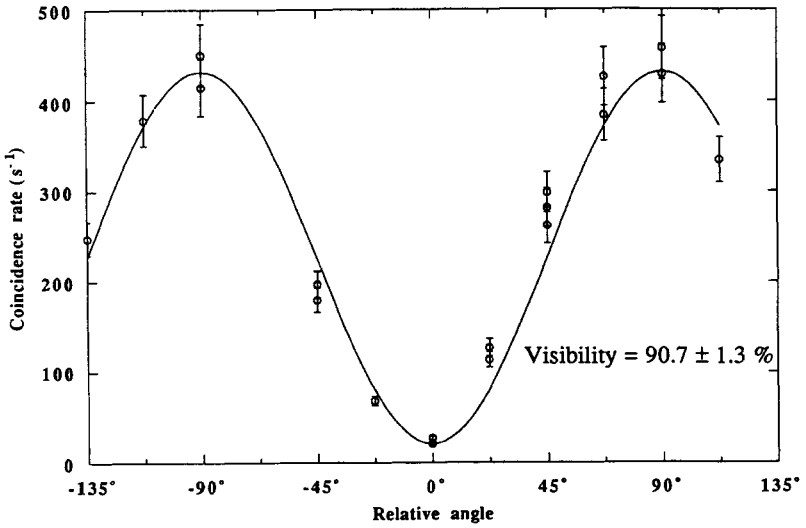


FIG. 5. Plot of coincidence rate vs relative angle of polarizers 1 and 2, corrected for accidentals. The smooth curve is a fit to theory, with visibility as a free parameter.

with no polarizer collapses the remote photon into a *mixed state* with half polarized horizontally and half polarized vertically. However, if a linear polarizer is placed at an angle θ_1 to the horizontal in output port 1, a detection event at detector 1 corresponds to a von Neumann projection in the subspace corresponding to that port onto the state vector $|\theta_1\rangle = (|1_1^H\rangle \cos\theta_1 + |1_1^V\rangle \sin\theta_1)$. We are left with a pure state for the conjugate photon:

$$\langle \theta_1|\psi\rangle_{\Delta x=0} = \tfrac{1}{2}(|1_2^V\rangle \cos\theta_1 - |1_2^H\rangle \sin\theta_1) \ . \tag{11}$$

Examining output port 2 with another polarizer, we observe that the light in this mode is polarized orthogonal to θ_1; the probability amplitude is

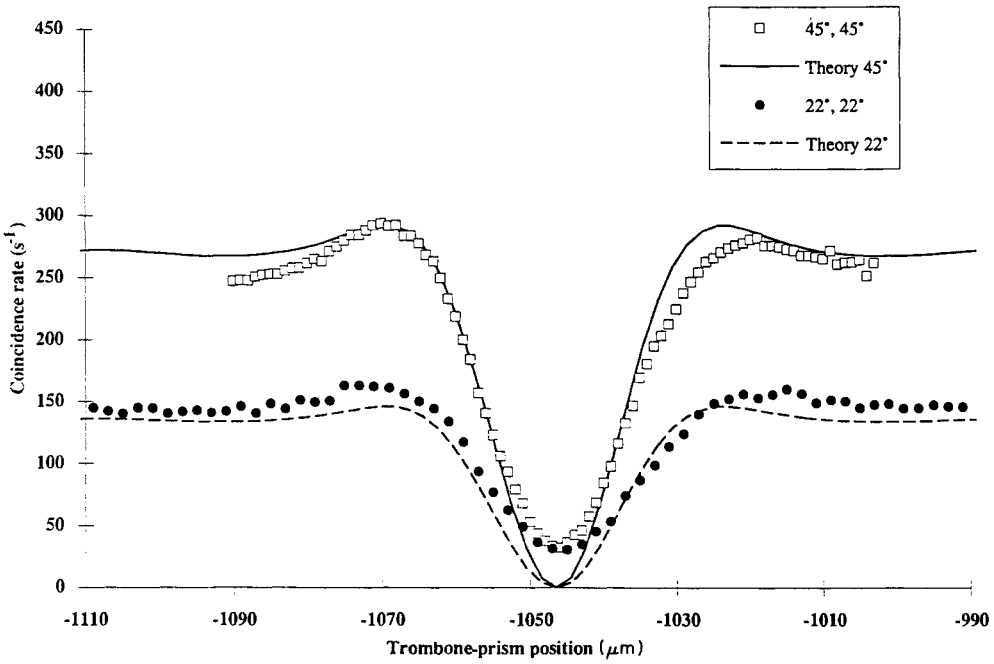


FIG. 6. Erasure also occurs, but in a somewhat different fashion, if the two polarizers are kept at the same angle and scanned towards 45°. With perfect visibility, the absolute angle would affect only the count rates far from the dip (clearly, at 0° or 90° no coincidences can ever be observed), and a total null would be observed at the dip because the two photons in the "singlet" state of Eq. (10) are always orthogonal. (The data are normalized to singles and corrected for accidentals.)

$$\langle\theta_2\theta_1|\psi\rangle_{\Delta x=0}=\tfrac{1}{2}(\langle 1_2^H|\cos\theta_2+\langle 1_2^V|\sin\theta_2)\times(|1_2^V\rangle\cos\theta_1-|1_2^H\rangle\sin\theta_1)=\tfrac{1}{2}(\sin\theta_2\cos\theta_1-\cos\theta_2\sin\theta_1)=\tfrac{1}{2}\sin(\theta_2-\theta_1)\ . \quad (12)$$

Thus

$$P_c(0)\approx|\langle\theta_1\theta_2|\psi\rangle_{\Delta x=0}|^2=\tfrac{1}{4}\sin^2(\theta_2-\theta_1)\ . \quad (13)$$

Hence the interference dip can be revived, but depending on the *relative* angle of the polarizers (see Fig. 5), it may be phase shifted and thus reappear as a peak, or any intermediate form. [Equation (13) may be recognized as a typical prediction of quantum theory when applied to certain tests of Bell's inequalities. This is hardly a coincidence, for the same nonlocal effect is responsible for both phenomena. We will discuss the relationship between our quantum-eraser experiment and similar Bell's inequalities experiments in Sec. VI.]

On the other hand, when the path-length difference is great compared with the coherence length of the photons, the probabilities for the two different paths leading to coincident detection add incoherently regardless of any polarizers:

$$P_c(\Delta x\gg\tau_c)\approx|\tfrac{1}{2}\langle\theta_1\theta_2|1_1^H1_2^V\rangle|^2+|\tfrac{1}{2}\langle\theta_1\theta_2|1_1^V1_2^H\rangle|^2=\tfrac{1}{4}\{\cos^2\theta_1\sin^2\theta_2+\sin^2\theta_1\cos^2\theta_2\}=\tfrac{1}{8}\{\sin^2(\theta_2-\theta_1)+\sin^2(\theta_2+\theta_1)\}\ . \quad (14)$$

This varies with absolute angle, in contrast to Eq. (13), since horizontal and vertical components act independently (see Fig. 6).

VI. DISCUSSION

The relationship between our experiment and other proposed quantum-eraser schemes is rather subtle. For comparison, we will focus on a particular proposal by Scully, Englert, and Walther [22], in which excited atoms are made to interfere in a two-slit–type geometry. A micromaser cavity is placed in each of the interfering paths, and prepared so that an atom passing through will decay with near certainty, leaving a photon in the cavity. For certain initial states of the cavity fields (i.e., number states), the extra photon from the decay constitutes *welcher Weg* information, and the *first*- order interference effect (fringes visible in singles detection of the atoms) is washed out. By allowing the cavity fields to subsequently interfere at a suitably placed detector, quantum erasure may be accomplished. However, as the authors stressed, this erasure is fundamentally a second-order phenomenon, in that the fringes can only be seen by *correlating* the photon counts with data stored elsewhere. In principle the decision of whether or not to erase could be postponed indefinitely, even beyond the time of detection of the atoms.

Our experiment differs somewhat from this proposal in that the basic Hong-Ou-Mandel interference effect is intrinsically a *second*-order, quantum-mechanical effect. That is, fringes are never visible in singles detection, and coincidence fringe visibility above 50% defies semiclassical explanation. While this puts the pre- and post-eraser fringes on the same footing, it has a disadvantage in that the distinguishing information we add and then erase is carried by the same photons which are to interfere. We maintain that this difference is not as fundamental as it may first appear. Although our photons themselves carry the information describing which trajectory they took, they do so only via their polarization vectors. We erase the information *after they have already left the interferometer,* and without affecting their center-of-mass wave function. In both Scully, Englert, and Walther's proposal and the present experiment, the measurement of which-way information consists of coupling the interfering particle's spatial wave function to the disjoint Hilbert space describing the *welcher Weg* detection system (e.g., micromaser cavities or photon polarization space). While this does not affect the spatial wave function, it does enlarge the Hilbert space in which it resides. It is the enlargement of the Hilbert space through entanglement, and subsequent reduction, which is the central feature of the quantum eraser.

It is useful to consider a slight gedanken variant of our experiment, which is in principle identical to it. We employ polarizing beam splitters, rather than simple polarizers, so that both polarizations may be detected. A computer then stores in one file the times of photon detection events (regardless of polarization), and in another file the polarizations of the detected photons. (Note that by making a polarization-insensitive quantum nondemolition measurement before the polarizers, one could delay the choice of polarizer orientations until after the coincident detection measurement.) Varying the orientation of the polarizing beam splitter then affects only the second file, and not the first; no interference is discernible in the first file until the data are correlated with that in the other file. As this may be performed long after the data are originally stored, we have a "delayed-choice" version of the quantum eraser.

Zajonc *et al.* have recently discussed two experiments in connection with the quantum eraser [23]. One of the experiments, while certainly a remarkable demonstration of complementarity, differs fundamentally from the quantum-eraser proposal in that it is entirely a first-order, not a second-order, interference effect. Detection events are never correlated with measurements on the "environment," and no delayed-choice version would be possible, even in principle. Their other result involves an interference effect which exists only in coincidence detection, as in our own experiment. The " 'delicate' change" which leads both to distinguishability and to erasure in their example is the removal and reinsertion of a beam splitter inside the interferometer [24]. In this sense, it is not a quantum eraser since it is the structure of the interferometer itself, and not just the structure of the detection scheme, which determines once and for all the presence or absence of interference fringes.

Some of the results presented here have been observed previously by other researchers, in the context of nonlocal correlations and Einstein-Podolsky-Rosen experi-

ments [25,26]. Our goal was to shift some of the focus to the phenomenon of quantum erasure, which is another striking aspect of quantum entanglement. The central element in many, if not most, tests of Bell's inequalities to date is the singlet state of the correlated photons [27]. Although our photons are not produced in such a state in the down-conversion process, it arises when their polarization states are entangled with their propagation modes (i.e., lower or upper arm) [28]. From this perspective, the quantum eraser and the Bell-type tests are just different approaches to investigating and understanding the character of the entangled states. One might then argue that some of the previous Bell-type tests *were* the first quantum-eraser results. However, we believe that it is important to demonstrate the loss of interference before reviving it, an aspect that, to our knowledge, has not been covered in any Bell-type experiments. Furthermore, the general goals of the two viewpoints differ. While the Bell's inequalities experiments seen to disprove the reality of local hidden variable models, the quantum eraser stresses the loss of coherence through entanglement with the "environment," and the possibility of recovering that coherence in certain circumstances.

VII. CONCLUSIONS

The quantum eraser offers an alternative perspective on interference and loss of quantum coherence in terms of (in)distinguishability of paths. Alternate paths are made distinguishable by correlating them to the "environment." Depending on how we reduce the resulting enlarged Hilbert space, we may opt to retain *welcher Weg* information and have no interference, or to reestablish indistinguishability and interference. We may make this choice long after the original interfering system has been detected, by correlating that data with the results of particular measurements on the environment with which the system was entangled. Of course, this demands that the coherence of the relevant environmental states be maintained.

Proposed experiments using atoms or neutrons, while intellectually engaging in principle, are at best very difficult in practice. We have seen that it is possible to demonstrate the essential features of the quantum eraser using a comparatively simple arrangement involving the correlated photons produced in spontaneous parametric down-conversion. The interference normally present when the two photons are superposed at a beam splitter was made to vanish when the alternate processes leading to coincidence counts were made distinguishable. For this purpose a half wave plate in one arm of the interferometer served to entangle the photon spatial wave function with the polarization subspace. Using polarizers at the output, it was then possible to restore interference, and even to alter its form.

One of the things the quantum eraser teaches us is that the state involved in interference is the *total* physical state, which in addition to photon spatial wave functions may include photon polarization, or even distant atoms with which the photons have interacted. In all realizations of the quantum eraser, the "magic" comes about through entangling the interfering system with some other degrees of freedom. The eraser "meddles" with the interference only via this entanglement, regardless of whether the extra information is stored in states of remote atoms or in the polarization components of the photonic wave functions. The process allows the introduction of an arbitrary phase between different components of the entangled state; in this sense, the phenomenon is better described as quantum *editing*.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research under Grant No. N00014-90-J-1259. We would like to thank Marlan Scully, Berthold-Georg Englert, Ivan Deutsch, and Yakir Aharonov for helpful discussions.

APPENDIX

As a preliminary to the general calculation of the quantum eraser effect, with arbitrary orientations of the half wave plate, polarizer P1, and polarizer P2, we write the effective projection operator for a polarizer at an angle θ to the horizontal, placed along the path corresponding to the propagation mode index j:

$$\hat{P}_{\mathrm{pol},j}(\theta)=|1_j^{H+\theta}\rangle\langle 1_j^{H+\theta}| =(|1_j^H\rangle\cos\theta+|1_j^V\rangle\sin\theta)(\langle 1_j^H|\cos\theta+\langle 1_j^V|\sin\theta)\ , \tag{A1}$$

where we are considering only the effect on single-photon states. It will be useful to consider the effect of this operator on a state of arbitrary polarization:

$$\begin{aligned}\hat{P}_{\mathrm{pol},j}(\theta)|1_j^{H+\phi}\rangle &=|1_j^{H+\theta}\rangle\langle 1_j^{H+\theta}|1_j^{H+\phi}\rangle \\ &=(|1_j^H\rangle\cos\theta+|1_j^V\rangle\sin\theta)(\langle 1_j^H|\cos\theta+\langle 1_j^V|\sin\theta)(|1_j^H\rangle\cos\phi+|1_j^V\rangle\sin\phi) \\ &=|1_j^H\rangle[\cos\theta(\cos\theta\cos\phi+\sin\theta\sin\phi)]+|1_j^V\rangle[\sin\theta(\cos\theta\cos\phi+\sin\theta\sin\phi)]\ .\end{aligned} \tag{A2}$$

For example, we can then examine the rate of single-event detection for a single-photon state, horizontally polarized, passed through a polarizer:

$$\begin{aligned}P_s\approx G^{(1)}(t;t)&=\langle\psi|\hat{\mathbf{E}}^{(-)}(t)\hat{\mathbf{E}}^{(+)}(t)|\psi\rangle \\ &\approx\langle 1^H|\hat{P}_{\mathrm{pol},j}(\theta)(\hat{a}_H^\dagger\hat{a}_H+\hat{a}_V^\dagger\hat{a}_V)\hat{P}_{\mathrm{pol},j}(\theta)|1^H\rangle=\cos^4\theta+\cos^2\theta\sin^2\theta=\cos^2\theta\ ,\end{aligned} \tag{A3}$$

which is the result expected from Malus's law.

We now show that using a single polarizer, before detector D1 for instance, is not enough to revive the interference dip. Using the reduced coincidence detection operator from Eq. (7), and the entangled state of the photons from (6b), we have

$$\begin{aligned}P_c(0)&\approx\langle\psi|\hat{P}_{\mathrm{pol},1}(\theta_1)\hat{P}_{c,\mathrm{red}}{}'\hat{P}_{\mathrm{pol},1}(\theta_1)|\psi\rangle_{\Delta x=0}\\&=\tfrac{1}{2}[\langle 1_1^H 1_2^{H+\phi}|-\langle 1_1^{H+\phi}1_2^H|]|1_1^{H+\theta_1}\rangle\langle 1_1^{H+\theta_1}|(\hat{a}^\dagger_{1,H}\hat{a}_{1,H}+\hat{a}^\dagger_{1,V}\hat{a}_{1,V})\\&\quad\times(\hat{a}^\dagger_{2,H}\hat{a}_{2,H}+\hat{a}^\dagger_{2,V}\hat{a}_{2,V})|1_1^{H+\theta_1}\rangle\langle 1_1^{H+\theta_1}|\tfrac{1}{2}[|1_1^H1_2^{H+\phi}\rangle-|1_1^{H+\phi}1_2^H\rangle]\\&=\tfrac{1}{4}\sin^2\phi(\cos^4\theta+2\cos^2\theta\sin^2\theta+\sin^4\theta)=\tfrac{1}{4}\sin^2\phi\ .\end{aligned}\tag{A4}$$

We see immediately that for $\phi=90°$, when the two paths are maximally distinguishable there is no null in coincidence for any orientation of a single polarizer at the output.

We turn now to the general case with two polarizers set at arbitrary angles θ_1 and θ_2.

$$\begin{aligned}P_c(0)&\approx\langle\psi|\hat{P}_{\mathrm{pol},1}(\theta_1)\hat{P}_{\mathrm{pol},2}(\theta_2)\hat{P}_{c,\mathrm{red}}{}'\hat{P}_{\mathrm{pol},2}(\theta_2)\hat{P}_{\mathrm{pol},1}(\theta_1)|\psi\rangle_{\Delta x=0}\\&=\tfrac{1}{2}[\langle 1_1^H1_2^{H+\phi}|-\langle 1_1^{H+\phi}1_2^H|](|1_1^{H+\theta_1}\rangle\langle 1_1^{H+\theta_1}|)(|1_2^{H+\theta_2}\rangle\langle 1_2^{H+\theta_2}|)(\hat{a}^\dagger_{1,H}\hat{a}_{1,H}+\hat{a}^\dagger_{1,V}\hat{a}_{1,V})\\&\quad\times(\hat{a}^\dagger_{2,H}\hat{a}_{2,H}+\hat{a}^\dagger_{2,V}\hat{a}_{2,V})(|1_2^{H+\theta_2}\rangle\langle 1_2^{H+\theta_2}|)(|1_1^{H+\theta_1}\rangle\langle 1_1^{H+\theta_1}|)\tfrac{1}{2}[|1_1^H1_2^{H+\phi}\rangle-|1_1^{H+\phi}1_2^H\rangle]\ .\end{aligned}$$

Using Eq. (A2), one can expand $|\tilde{\psi}\rangle_{\Delta x=0}=\hat{P}_{\mathrm{pol},2}(\theta_2)\hat{P}_{\mathrm{pol},1}(\theta_1)|\psi\rangle_{\Delta x=0}$. After simplifying algebra one finds

$$\begin{aligned}|\tilde{\psi}\rangle_{\Delta x=0}&=|1_1^H1_2^H\rangle\cos\theta_1\cos\theta_2\sin(\theta_2-\theta_1)\sin\phi+|1_1^V1_2^V\rangle\sin\theta_1\sin\theta_2\sin(\theta_2-\theta_1)\sin\phi\\&\quad+|1_1^H1_2^V\rangle\cos\theta_1\sin\theta_2\sin(\theta_2-\theta_1)\sin\phi+|1_1^V1_2^H\rangle\sin\theta_1\cos\theta_2\sin(\theta_2-\theta_1)\sin\phi\ .\end{aligned}$$

It then follows that

$$P_c(0)\approx\langle\tilde{\psi}|\hat{P}_{c,\mathrm{red}}{}'|\tilde{\psi}\rangle_{\Delta x=0}=\sin^2\phi\sin^2(\theta_2-\theta_1)\ ,\tag{A5}$$

which is the more general case of Eq. (13).

[1] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison Wesley, Reading, MA, 1965).

[2] A. Peres, Phys. Rev. D **22**, 879 (1980).

[3] A. Stern, Y. Aharonov, and Y. Imry, Phys. Rev. A **41**, 3436 (1990). Stern *et al.* have shown a general equivalence between this interpretation and one in which the environment "scrambles" the phase of the interfering system due to an uncertain interaction potential. We find the second picture less helpful in the experiments under consideration, as the relation of this potential to the experimental system is often rather subtle. Thus we restrict ourselves here to the first interpretation.

[4] M. Hillery and M. O. Scully, in *Quantum Optics, Experimental Gravitation, and Measurement Theory,* edited by P. Meystre *et al.* (Plenum, New York, 1983), pp. 65–85.

[5] M. O. Scully, B.-G. Englert, and H. Walther, Nature **351**, 111 (1991).

[6] M. O. Scully, R. Shea, and J. D. McCullen, Phys. Rep. **43**, 485 (1978).

[7] The atom is initially polarized in the $+\hat{\mathbf{x}}$ direction, using a Stern-Gerlach apparatus (SGA) along $+\hat{\mathbf{x}}$ as a filter. An SGA measuring $\langle S\rangle$ along $+\hat{\mathbf{z}}$ acts as the first beam splitter. Two SGA's (along $-\hat{\mathbf{z}}$) act as mirrors to deflect the spin-up atoms to the recombining beam splitter, another SGA along $+\hat{\mathbf{z}}$. Similarly, two other SGA's (along $-\hat{\mathbf{z}}$) act as mirrors to deflect the spin-down atoms to the recombining beam spitter. In principle, the SGA in the second half of the interferometer may be made to reverse the effects of the SGA in the first half; the atoms after the final SGA will then be polarized in the $+\hat{\mathbf{x}}$ direction again. This may be checked by an analyzer (a second SGA along $+\hat{\mathbf{x}}$) after the interferometer.

[8] B.-G. Englert, J. Schwinger, and M. O. Scully, Found. Phys. **18**, 1045 (1988).

[9] J. Schwinger, M. O. Scully, and B.-G. Englert, Z. Phys. D **10**, 135 (1988).

[10] M. O. Scully and H. Walther, Phys. Rev. A **39**, 5229 (1989).

[11] A. Zajonc, Phys. Lett. **96A**, 61 (1983).

[12] M. G. Raymer and S. Yang (unpublished).

[13] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, in *Proceedings of the Second Santa Fe Meeting on the Foundations of Quantum Mechanics,* edited by M. O. Scully *et al.* (World Scientific, Singapore, 1992).

[14] See, for instance, [20]; P. G. Kwiat and R. Y. Chiao, Phys. Rev. Lett. **66**, 588 (1991); C. K. Hong, Z. Y. Ou, and L. Mandel, in *Photons and Quantum Fluctuations,* Malvern Physics Series, edited by E. R. Pike *et al.* (Hilger, Bristol, 1988), p. 51–65, and references therein.

[15] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).

[16] A. M. Steinberg, P. G. Kwiat and R. Y. Chiao, Phys. Rev. Lett. **68**, 2421 (1992).

[17] R. Campos, B. Saleh, and M. Teich, Phys. Rev. A **42**, 4127 (1990).

[18] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).

Chapter 6
Beam splitters and interferometers

6.1 Experiments with single photons

Central to the entire discipline of quantum optics, as should be evident from the preceding chapters, is the concept of the photon. Yet it is perhaps worthwhile to pause and ask: what is the evidence for the existence of photons? Most of us first encounter the photon concept in the context of the photo-electric effect. As we showed in Chapter 5, the photo-electric effect is, in fact, used to indirectly detect the presence of photons – the photo-electrons being the entities counted. But it turns out that some aspects of the photo-electric effect can be explained without introducing the concept of the photon. In fact, one can go quite far with a semiclassical theory in which only the atoms are quantized with the field treated classically. But we hasten to say that, for a satisfactory explanation of *all* aspects of the photo-electric effect, the field *must* be quantized. As it happens, the other venerable "proof" of the existence of photons, the Compton effect, can also be explained without quantized fields.

In an attempt to obtain quantum effects with light, Taylor, in 1909 [1], obtained interference fringes in an experiment with an extremely weak source of light. His source was a gas flame and the emitted light was attenuated by means of screens made of smoked glass. The "double slit" in the experiment was, in fact, a needle whose shadow on a screen exhibited the fringes of a diffraction pattern when exposed to direct light from the source. But Taylor found that the fringes persisted upon attenuation of the source, even down to the lowest intensities where, one could naively conclude, on the basis of simple energy considerations, there was at most only one photon at a time between the source and the screen. Apparently, photons passing by the needle one at a time give rise to interference. Presumably, this is the origin of Dirac's famous remark [2] that "each photon interferes only with itself, interference between two photons does not occur". But we now know, as discussed in Chapter 5, that a thermal source, such as the gas flame used by Taylor, does not produce photons one-at-a-time but rather produces them in bunches. Hence it is naïve and wrong to use energy considerations alone to determine the number of photons between the source and the screen at any given time; there is a strong likelihood that there are two photons present, the

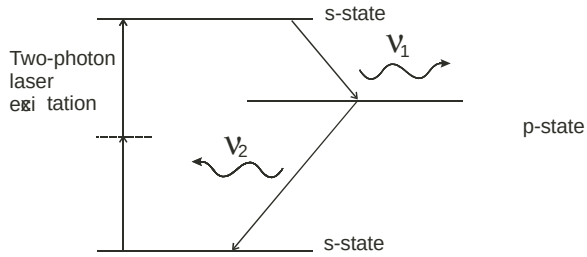


Fig. 6.1. Energy-level diagram of the single photon source for the experiments of Grangier *et al.* A calcium atom is irradiated by a laser, which excites the atom by two-photon absorption to a high-lying s-state. The atom then undergoes a cascade decay first to a p-state, emitting a photon of frequency ν_1, used as the trigger photon, and then to the original s-state emitting a photon of frequency ν_2.

photon bunching effect. A laser source produces photons randomly, so even when attenuated there is at least some chance that there may be more than one photon between the source and the screen. To get as close as possible to having a single photon between source and screen requires a source of antibunched photons. This in turn requires a source consisting of only a very few atoms, ideally a single atom.

Such a source was developed only relatively recently by Grangier *et al.* [3] originally for the purpose of a fundamental test of quantum mechanics, namely a search for violations of Bell's inequality, and then used to demonstrate the indivisibility of photons. This source consists of a beam of calcium atoms irradiated by laser light exciting the atoms to a high-lying s-state. The s-state undergoes a rapid decay to a p-state emitting a photon of frequency ν_1. Subsequently the atom rapidly undergoes another rapid decay, this time to the ground s-state, by emitting a second photon, this one having a frequency ν_2 (see Fig. 6.1). The photons, to conserve momentum, are emitted in opposite directions. In the experiment described in Reference [3], the first photon, detected by D_{trig}, was used as a "trigger" to alert a set of photo-detectors placed at the outputs of a 50:50 beam splitter upon which the second photon falls as illustrated in Fig. 6.2. The trigger tells the photo-detectors to expect a photon to emerge from the beam splitter by "gating" the detection electronics for a brief time interval. This eliminates spurious counts due to photons entering the detectors from irrelevant sources. The experimental setup, as pictured in Fig. 6.2, is such that only the particle nature of the photons will be manifested. That is, a single photon falling on the beam splitter would be either reflected into detector D_{ref} or transmitted into detector D_{tran}, i.e. it is a "which path" experiment and no interference effects are expected. There should be no simultaneous counts (the counts should anti-correlated) of reflected and transmitted photons and, because the beam splitter is 50:50, repeated runs of the experiment should result in each of the two detectors firing approximately 50% of the time. These expectations were confirmed by the investigators.

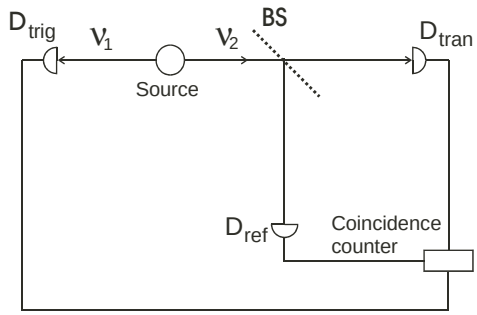


Fig. 6.2. Anti-correlation experiment of Grangier *et al.* Detection of the trigger photon alerts the coincidence counters to expect a photon to pass through the beam splitter. The beam splitter is 50:50.

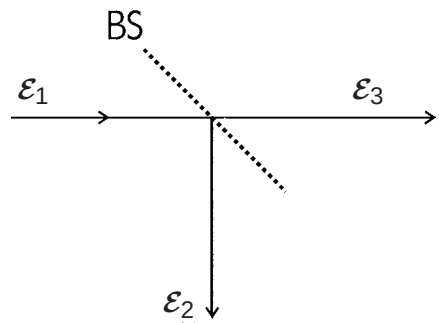


Fig. 6.3. Classical beam splitting. A classical field of amplitude $\mathcal{E}_1$ is split into fields of amplitudes $\mathcal{E}_2$ and $\mathcal{E}_3$.

6.2 Quantum mechanics of beam splitters

At this point we must pause and consider the beam splitter in a fully quantum-mechanical context. In the previous chapter, we used the notion of a beam splitter in a somewhat cavalier manner. We were able to get away with it because, for "classical"-like light beams, coherent and thermal beams, the quantum and classical treatments of beam splitters agree. But at the level of a single or few photons, the classical approach to beam splitting produces erroneous and quite misleading results.

To see how classical reasoning over beam splitting goes wrong, let us consider first a classical light field of complex amplitude $\mathcal{E}_1$ incident upon a lossless beam splitter as indicated in Fig. 6.3. $\mathcal{E}_2$ and $\mathcal{E}_3$ are the amplitudes of the reflected and transmitted beams respectively. If r and t are the (complex) reflectance and transmittance respectively of the beam splitter, then it follows that

$$\mathcal{E}_2 = r\mathcal{E}_1 \quad \text{and} \quad \mathcal{E}_3 = t\mathcal{E}_1. \tag{6.1}$$

For a 50:50 beam splitter we would have $|r| = |t| = 1/\sqrt{2}$. However, for the sake of generality, we do not impose this condition here. Since the beam splitter is assumed lossless, the intensity of the input beam should equal the sum of the intensities of the two output beams:

$$|\mathcal{E}_1|^2 = |\mathcal{E}_2|^2 + |\mathcal{E}_3|^2 \tag{6.2}$$

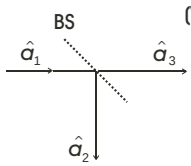


Fig. 6.4. Naïve, and incorrect, quantum mechanical depiction of a beam splitter.

which requires that

$$|r|^2 + |t|^2 = 1. \tag{6.3}$$

To treat the beam splitter quantum mechanically we might try replacing the classical complex field amplitudes $\mathcal{E}_i$ by a set of annihilation operators $\hat{a}_i$ ($i = 1, 2, 3$) as indicated in Fig. 6.4. In analogy with the classical case we might try setting

$$\hat{a}_2 = r\hat{a}_1 \quad \text{and} \quad \hat{a}_3 = t\hat{a}_1. \tag{6.4}$$

However, the operators of each of the fields are supposed to satisfy the commutation relations

$$\left[\hat{a}_i, \hat{a}_j^\dagger\right] = \delta_{ij}, \quad [\hat{a}_i, \hat{a}_j] = 0 = \left[\hat{a}_i^\dagger, \hat{a}_j^\dagger\right] \quad (i, j = 1, 2, 3), \tag{6.5}$$

but it is easy to see that for the operators of Eq. (6.4) we obtain

$$\begin{aligned} \left[\hat{a}_2, \hat{a}_2^\dagger\right] &= |r|^2 \left[\hat{a}_1, \hat{a}_1^\dagger\right] = |r|^2, \\ \left[\hat{a}_3, \hat{a}_3^\dagger\right] &= |t|^2 \left[\hat{a}_1, \hat{a}_1^\dagger\right] = |t|^2, \\ \left[\hat{a}_2, \hat{a}_3^\dagger\right] &= rt^* \neq 0, \text{ etc.} \end{aligned} \tag{6.6}$$

Thus the transformations in Eq. (6.4) do not preserve the commutation relations and therefore cannot provide the correct quantum description of a beam splitter. This conundrum is resolved as follows: in the classical picture of the beam splitter there is an unused "port" which, being empty of an input field, has no effect on the output beams. However, in the quantum-mechanical picture, the "unused" port still contains a quantized field mode albeit in the vacuum state and, as we have repeatedly seen, the fluctuations of the vacuum lead to important physical effects. The situation with the beam splitter is no exception. In Fig. 6.5 we indicate all the modes required for a proper quantum description of the beam splitter, $\hat{a}_0$ representing the field operator of the classically vacant input mode. Also indicated are two sets of transmittances and reflectances, allowing for the possibility of an asymmetric beam splitter. We now write the beam-splitter transformations for the field operators as

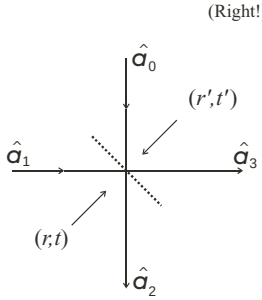


Fig. 6.5. The correct quantum-mechanical depiction of a beam splitter.

$$\hat{a}_2 = r\hat{a}_1 + t'\hat{a}_0, \quad \hat{a}_3 = t\hat{a}_1 + r'\hat{a}_0 \tag{6.7}$$

or collectively as

$$\begin{pmatrix} \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \begin{pmatrix} t' & r \\ r' & t \end{pmatrix} \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \end{pmatrix}. \tag{6.8}$$

It is easily seen that the commutation relations of Eq. (6.5) are satisfied as long as the following relations hold:

$$\left|r'\right| = |r|, |t| = \left|t'\right|, |r|^2 + |t|^2 = 1, \; r^*t' + r't'^* = 0, \text{ and } r^*t + r't'^* = 0. \tag{6.9}$$

These relations are known as the reciprocity relations and can also be derived on the basis of energy conservation.

Let us examine a couple of relevant examples. The phase shifts of the reflected and transmitted beams depend on the construction of the beam splitter [4]. If the beam splitter is constructed as a single dielectric layer, the reflected and transmitted beams will differ in phase by a factor of $\exp(\pm\, i\pi/2) = \pm\, i$. For a 50:50 beam splitter, assuming the reflected beam suffers a $\pi/2$ phase shift, the input and output modes are related according to

$$\hat{a}_2 = \frac{1}{\sqrt{2}}\,(\hat{a}_0 + i\hat{a}_1)\,, \quad \hat{a}_3 = \frac{1}{\sqrt{2}}\,(i\hat{a}_0 + \hat{a}_1)\,. \tag{6.10}$$

Since the transformation between input and output modes must be unitary, we may write Eq. (6.8) as

$$\begin{pmatrix} \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \hat{U}^\dagger \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \end{pmatrix} \hat{U} \tag{6.11}$$

where $\hat{U}$ is a unitary operator. This transformation constitutes a Heisenberg picture formulation of a beam splitter. For the specific transformation represented by Eq. (6.10), the operator $\hat{U}$ has the form

$$\hat{U} = \exp\left[i\,\frac{\pi}{4}\left(\hat{a}_0^\dagger \hat{a}_1 + \hat{a}_0 \hat{a}_1^\dagger\right)\right], \tag{6.12}$$

easily checked using the Baker–Hausdorf lemma in Eq. (6.11).

On the other hand, we may adopt the Schrödinger picture and ask the following question. For a given input state to the beam splitter, what is the output state? Remembering that all photon number states $|n\rangle$, hence any superposition or any statistical mixture of such states, may be constructed by the action of n powers of the creation operator on the vacuum, we may use Eqs. (6.7) and (6.8) to construct the output states from the action of the transformed creation operators on the vacuum states of the output modes, it being obvious that an input vacuum transforms to an output vacuum: $|0\rangle_0|0\rangle_1 \to |0\rangle_2|0\rangle_3$.

As an example, consider the single photon input state $|0\rangle_0|1\rangle_1$ which we may write as $\hat{a}_1^\dagger|0\rangle_0|0\rangle_1$. For the beam splitter described by Eq. (6.10) we find that $\hat{a}_1^\dagger = (i\,\hat{a}_2^\dagger + \hat{a}_3^\dagger)/\sqrt{2}$. Thus we may write, using that $|0\rangle_0|0\rangle_1 \xrightarrow{\text{BS}} |0\rangle_2|0\rangle_3$,

$$\begin{aligned} |0\rangle_0\,|1\rangle_1 &\xrightarrow{\text{BS}} \frac{1}{\sqrt{2}}\left(i\,\hat{a}_2^\dagger + \hat{a}_3^\dagger\right)|0\rangle_2\,|0\rangle_3 \\ &= \frac{1}{\sqrt{2}}\,(i\,|1\rangle_2\,|0\rangle_3 + |0\rangle_2\,|1\rangle_3)\,. \end{aligned} \tag{6.13}$$

This is an important result. It says that a single-photon incident at one of the input ports of the beam splitter, the other port containing only the vacuum, will be either transmitted or reflected with equal probability. Of course, this is precisely as we earlier claimed and explains why no coincident counts are to be expected with photon counters placed at the outputs of the beam splitter, as confirmed by the experiment of Grangier *et al.* [3]. Actually, because the beam splitter is well understood, the lack of coincident counts in the above experiment may be taken as an indication that the source is truly producing single-photon states. (Obviously,

the beam splitter is a *passive* device that neither creates nor destroys photons. There are, of course, *active* devices that convert one photon into two, for example, and we shall encounter these in the next chapter.)

One other point needs to be made about the output state of Eq. (6.13). It is an *entangled* state: it cannot be written as a simple product of states of the individual modes 2 and 3. The density operator (see Appendix A) for the (pure) state of Eq. (6.13) is

$$\begin{aligned}\hat{\rho}_{23} = \frac{1}{2}\{&|1\rangle_2\,|0\rangle_{32}\langle 1|_3\,\langle 0| + |0\rangle_2\,|1\rangle_{32}\,\langle 0|_3\,\langle 1| \\ &+ i\,|1\rangle_2\,|0\rangle_{32}\,\langle 0|_3\,\langle 1| - i\,|0\rangle_2\,|1\rangle_{32}\,\langle 1|_3\,\langle 0|\}\,. \end{aligned} \tag{6.14}$$

In placing detectors in the two output beams, we are measuring the full "coherence" as described by the state vector of Eq. (6.13) or equivalently the density operator of Eq. (6.14). Suppose, on the other hand, we make no measurement of, say, mode 3. Mode 2 is then described by the reduced density operator obtained by tracing over the states of the unmeasured mode (see Appendix A):

$$\begin{aligned}\hat{\rho}_2 = \mathrm{Tr}_3\hat{\rho}_{23} &= \sum_{n=0}^{\infty}{}_3\langle n|\hat{\rho}_{23}\,|n\rangle_3 \\ &= \frac{1}{2}\,(|0\rangle_{22}\,\langle 0| + |1\rangle_{22}\,\langle 1|)\,. \end{aligned} \tag{6.15}$$

This represents merely a statistical mixture, there being no "off-diagonal" coherence terms of the form $|0\rangle\langle 1|$ or $|1\rangle\langle 0|$. Thus placing a detector in only one of the output beams yields random results, 0 or 1, each 50% of the time, just as we would expect.

Before moving on to single-photon interference, let us consider two more examples of beam splitting. First we consider a coherent state, a classical-like state, rather the opposite of the highly nonclassical single-photon state, incident on the beam splitter with, again, only the vacuum in the other input port. That is, the initial state is $|0\rangle_0|\alpha\rangle_1 = \hat{D}_1(\alpha)|0\rangle_0|0\rangle_1$ where $\hat{D}_1(\alpha) = \exp(\alpha\hat{a}_1^\dagger - \alpha^*\hat{a}_1)$ is the displacement operator for mode 1. We may then, following the procedure above, obtain the output state according to

$$\begin{aligned}|0\rangle_0\,|\alpha\rangle_1 &\xrightarrow{\text{BS}} \exp\left[\frac{\alpha}{\sqrt{2}}\left(i\,\hat{a}_2^\dagger + \hat{a}_3^\dagger\right) - \frac{\alpha^*}{\sqrt{2}}\,(-i\,\hat{a}_2 + \hat{a}_3)\right]|0\rangle_2\,|0\rangle_3 \\ &= \exp\left[\left(\frac{i\,\alpha}{\sqrt{2}}\right)\hat{a}_2^\dagger - \left(\frac{-i\,\alpha^*}{\sqrt{2}}\right)\hat{a}_2\right]\exp\left[\left(\frac{\alpha}{\sqrt{2}}\right)\hat{a}_3^\dagger - \left(\frac{\alpha^*}{\sqrt{2}}\right)\hat{a}_3\right]|0\rangle_2\,|0\rangle_3 \\ &= \left|\frac{i\,\alpha}{\sqrt{2}}\right\rangle_2\left|\frac{\alpha}{\sqrt{2}}\right\rangle_3. \end{aligned} \tag{6.16}$$

Evidently we obtain the result expected for a classical light wave where the incident intensity is evenly divided between the two output beams, e.g. half the incident average photon number, $|\alpha|^2/2$, emerges in each beam. We also naturally obtain the phase shift $i = e^{i\pi/2}$ for the reflected wave, as expected. Finally, note that the output is not entangled.

Essentially everything about beam splitting with a coherent state is classical and in that regard it is worth inserting here a note of caution. The case of a single photon incident on the beam splitter is *not* obtainable as the limiting case of an incident coherent state, i.e. for $|\alpha|$ small. It is easy to check that the single-photon result of Eq. (6.13) is not obtainable in any way as a limiting case of Eq. (6.16). This is quite obvious without doing any calculations as the former is entangled whereas the latter is not. Entanglement cannot arise as a limiting case of a product state. The point is, as mentioned earlier in the chapter, that attempts at extrapolating low-field behavior from classical results are misleading and quite wrong.

As a last example with operator transformations, we return to the strictly quantum domain and consider the situation where single photons are simultaneously injected into the two input ports of our 50:50 beam splitter, the incident state being $|1\rangle_0|1\rangle_1 = \hat{a}_0^\dagger \hat{a}_1^\dagger |0\rangle_0|0\rangle_1$. Again, following the previous procedure with $\hat{a}_0^\dagger = (\hat{a}_2^\dagger + i\hat{a}_3^\dagger)/\sqrt{2}$ and $\hat{a}_1^\dagger = (i\hat{a}_2^\dagger + \hat{a}_3^\dagger)/\sqrt{2}$ we have

$$\begin{aligned} |1\rangle_0\,|1\rangle_1 &\xrightarrow{\text{BS}} \frac{1}{2}\left(\hat{a}_2^\dagger + i\,\hat{a}_3^\dagger\right)\left(i\,\hat{a}_2^\dagger + \hat{a}_3^\dagger\right)|0\rangle_2\,|0\rangle_3 \\ &= \frac{i}{2}\left(\hat{a}_2^\dagger\hat{a}_2^\dagger + \hat{a}_3^\dagger\hat{a}_3^\dagger\right)|0\rangle_2\,|0\rangle_3 \\ &= \frac{i}{\sqrt{2}}\left(|2\rangle_2\,|0\rangle_3 + |0\rangle_2\,|2\rangle_3\right). \end{aligned} \tag{6.17}$$

Apparently, the two photons emerge together such that photo-detectors placed in the output beams should not register simultaneous counts. But unlike the case of a single incident photon, the physical basis for obtaining no simultaneous counts is not a result of the particle-like nature of photons. Rather, it is caused by interference (a wave-like effect) between two possible ways of obtaining the (absent) output state $|1\rangle_2|1\rangle_3$: the process where both photons are transmitted (Fig. 6.6(a)) and the process where they are both reflected (Fig. 6.6(b)). Note the indistinguishability of the two processes for the output state $|1\rangle_2|1\rangle_3$. There is a simple and rather intuitive way of understanding this result. Recall Feynman's rule [5] for obtaining the probability for an outcome that can occur by several indistinguishable processes: one simply adds the probability *amplitudes* of all the processes and then calculates the square of the modulus. Assuming that our beam splitter is described by Eq. (6.10), the reflected photons each acquire an $e^{i\pi/2} = i$ phase shift. The amplitude for transmission for each photon is $A_\mathrm{T} = 1/\sqrt{2}$ and the amplitude for reflection for each is $A_\mathrm{R} = i/\sqrt{2}$. The amplitude that *both* photons are transmitted is $A_\mathrm{T} \cdot A_\mathrm{T}$ and that both are reflected is $A_\mathrm{R} \cdot A_\mathrm{R}$. Thus the probability of the photons emerging in both output beams is

$$P_{11} = |A_\mathrm{T} \cdot A_\mathrm{T} + A_\mathrm{R} \cdot A_\mathrm{R}|^2 = \left|\frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\frac{i}{\sqrt{2}}\right|^2 = 0. \tag{6.18}$$

An experimental demonstration of this effect was first performed by Hong, Ou, and Mandel [6] and is discussed in Chapter 9.

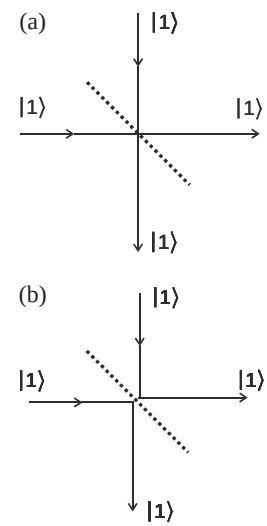


Fig. 6.6. Two indistinguishable processes with simultaneous single photon inputs. (a) Both photons are transmitted. (b) Both photons are reflected. These processes interfere destructively with each other.

Chapter 9
Optical test of quantum mechanics

Over the past three decades or so, experiments of the type called Gedanken have become real. Recall the Schrödinger quote from Chapter 8: ". . . we never experiment with just one electron or atom or (small) molecule." This is no longer true. We can do experiments involving single atoms or molecules and even on single photons, and thus it becomes possible to demonstrate that the "ridiculous consequences" alluded to by Schrödinger are, in fact, quite real. We have already discussed some examples of single-photon experiments in Chapter 6, and in Chapter 10 we shall discuss experiments performed with single atoms and single trapped ions. In the present chapter, we shall elaborate further on experimental tests of fundamentals of quantum mechanics involving a small number of photons. By fundamental tests we mean tests of quantum mechanics against the predictions of local realistic theories (i.e. hidden variable theories). Specifically, we discuss optical experiments demonstrating violations of Bell's inequalities, violations originally discussed by Bell in the context of two spin-one-half particles [1]. Such violations, if observed experimentally, falsify local realistic hidden-variable theories. Locality refers to the notion, familiar in classical physics, that there cannot be a causal relationship between events with space-like separations. That is, the events cannot be connected by any signal moving at, or less than, the speed of light; i.e. the events are outside the light-cone. But in quantum mechanics, it appears that nonlocal effects, effects seemly violating the classical notion of locality in a certain restricted sense, are possible. For example, the fact that a measurement on one part of an entangled system seems to instantaneously project the other part of the system into a particular state, even though the parts may be widely separated, is a nonlocal quantum effect. "Realism" in the context of hidden variable theories means that a quantum system has objectively definite attributes (quantum numbers) for all observables at all times. For example, a particle with spin $+\,{}^1\!/_2$ along the z-direction is known to be in a superposition state with spin $\pm\,{}^1\!/_2$ along the x-direction. The standard Copenhagen interpretation of quantum mechanics holds that the particle's spin along the x-direction is objectively indefinite, i.e. has no particular value of spin as a matter of principle until a measurement reduces the state vector to one of the possible states in the superposition. In a hidden-variable theory, the spin along the x-axis is assumed to be definite,

though unknown, and the experiment merely reveals the already existing state of the particle. But there are two types of hidden-variable theory, local and nonlocal. Nonlocal hidden-variable theories of the type considered by Bohm [2] reproduce the predictions of standard quantum mechanics. But they exhibit nonlocality, the feature of quantum mechanics that seems to have bothered Einstein even more than the apparent lack of realism. But local hidden-variable theories do make predictions different from those of standard quantum theory, and Bell's theorem provides a way to make such tests. In this chapter, we shall discuss optical tests of local hidden-variable theories using polarization-entangled photons.

In what follows we shall first discuss modern sources of paired and entangled single-photon states obtained from down-conversion processes (as opposed to the technique of cascaded emissions from atomic transitions as discussed in Chapter 6). We then review some one- and two-photon interference experiments, introduce the notion of the "quantum eraser" in this context, and then discuss an experiment on induced quantum coherence. Next, an experiment on photon tunneling exhibiting superluminal effects is discussed. Finally we describe two experiments on tests of Bell's inequalities, one involving polarization-entangled photons and a second (Franson's experiment) based on the time–energy uncertainty relation.

9.1 Photon sources: spontaneous parametric down-conversion

We have already discussed, in Chapter 7, the generation of nonclassical light via parametrically driven nonlinear media characterized by a second-order nonlinear susceptibility $\chi^{(2)}$. We focus on the nondegenerate case, and assuming the pump field to be quantized, the interaction Hamiltonian takes the form

$$\hat{H}_I \sim \chi^{(2)} \hat{a}_{\rm p} \hat{a}_{\rm s}^{\dagger} \hat{a}_{\rm i}^{\dagger} + H.c., \tag{9.1}$$

where we have altered our notation from Chapter 7 so that now $\hat{a}_{\rm p}$ is the annihilation operator of the pump beam and $\hat{a}_{\rm s}^{\dagger}$ and $\hat{a}_{\rm i}^{\dagger}$ are the creation operators of the "signal" and "idler" beams respectively. The denotations "signal" and "idler" appear for historical reasons and have no special significance, the choice of beam labels being somewhat arbitrary. In the simplest case, with the signal and idler beams initially in vacuum states, a single pump beam photon, typically in the ultraviolet spectral range, is converted into two optical photons, one in the signal beam, the other in the idler:

$$|1\rangle_{\rm p} |0\rangle_{\rm s} |0\rangle_{\rm i} \Rightarrow \hat{a}_{\rm p} \hat{a}_{\rm s}^{\dagger} \hat{a}_{\rm i}^{\dagger} |1\rangle_{\rm p} |0\rangle_{\rm s} |0\rangle_{\rm i} = |0\rangle_{\rm p} |1\rangle_{\rm s} |1\rangle_{\rm i} . \tag{9.2}$$

As the signal and idler modes are initially in vacuum states, the process is "spontaneous". Note that the photons produced in the signal and idler modes are assumed to be generated simultaneously. That this is the case was demonstrated many years ago by Burnham and Weinberg [3] who used coincidence counting with detectors arranged to satisfy momentum and energy conservation and to have equal

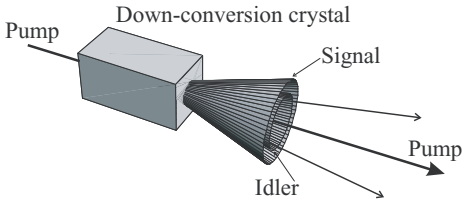


Fig. 9.1. Type I down-conversion. Photons from the pump beam are converted into signal and idler photons that emerge from the crystal along different directions. The signal and idler photons have identical polarization but orthogonal to that of the pump. The possible directions form a set of concentric cones. The light from the different cones is of different colors, typically orange near the center to a deep red at wider angles. The pump beam is in the ultraviolet.

time delays. The simultaneous production of signal and idler photons is key to the applications of such parametric sources to the fundamental test of quantum mechanics to be described. In order for the down-conversion process to go forward however, certain other conditions must be satisfied. Letting $\omega_{\rm p}$, $\omega_{\rm s}$, and $\omega_{\rm i}$ represent the frequencies of the pump, signal, and idler respectively, energy conservation requires that

$$\hbar\omega_{\rm p} = \hbar\omega_{\rm s} + \hbar\omega_{\rm i}. \tag{9.3}$$

Further, if $\mathbf{k}_{\rm p}$, $\mathbf{k}_{\rm s}$, and $\mathbf{k}_{\rm i}$ represent the respective wave numbers, then we must have, inside the crystal,

$$\hbar\mathbf{k}_{\rm p} \approx \hbar\mathbf{k}_{\rm s} + \hbar\mathbf{k}_{\rm i}, \tag{9.4}$$

where the $\approx$ sign appears as the result of an uncertainty given by the reciprocal of the length of the nonlinear medium [4]. Equations (9.3) and (9.4) are known as the "phase matching" conditions and they can be attained in certain types of nonlinear media, such as noncentrosymmetric crystals [5]. Only noncentrosymmetric crystals have a nonvanishing $\chi^{(2)}$. The most commonly used crystals are KDP (KD_2PO_4) and BBO (β-BaB_2O_4). The connection to nonlinear optics is given in Appendix D.

There are, in fact, two types of SPDC process. In type I, the signal and idler photons have the same polarization but these are orthogonal to that of the pump. The interaction Hamiltonian for this process is given by

$$\hat{H}_{\rm I} = \hbar\eta\,\hat{a}_{\rm s}^{\dagger}\hat{a}_{\rm i}^{\dagger} + H.c., \tag{9.5}$$

where the parametric approximation has been made and where $\eta \propto \chi^{(2)}\mathcal{E}_{\rm p}$,where $\mathcal{E}_{\rm p}$ is the amplitude of the classical coherent field. The phase-matching condition of Eq. (9.4) imposes a constraint such that the signal and idler photons (conjugate photons) must emerge from the crystal on opposite sides of concentric cones centered on the direction of the pump beam as shown in Fig. 9.1. Evidently there are an infinite number of ways of selecting the signal and idler beams. Examples

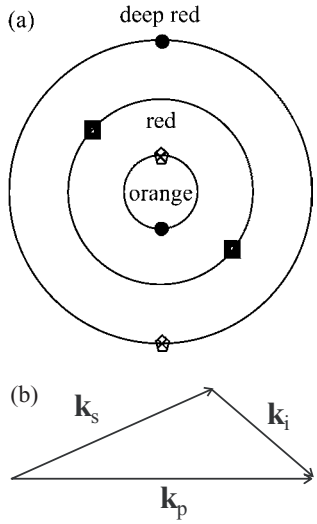


Fig. 9.2. (a) A cross-section of the cones of light emerging from a type I down-converter. Like symbols represent conjugate photons satisfying the phase-matching condition. Note that those on the middle circle are degenerate in frequency. (b) A graphical view of the phase-matching condition.

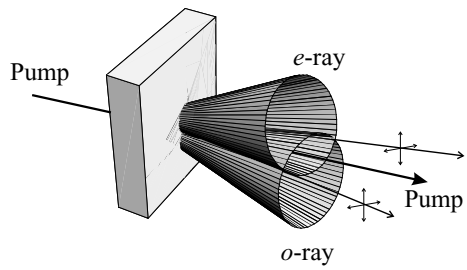


Fig. 9.3. Type II down-conversion. The signal and idler photons have orthogonal polarizations. Birefringence effects cause the photons to be emitted along two intersecting cones, one of the ordinary (*o*) ray and the extraordinary (*e*) ray.

are shown in Fig. 9.2. The Hamiltonian of Eq. (9.5) in practice represents a particular "post-selection" of the momenta of the output beams.

In type II down-conversion, the signal and idler photons have orthogonal polarizations. Because of birefringence effects, the generated photons are emitted along two cones, one for the ordinary (*o*) wave and another for the extraordinary (*e*) wave, as indicated in Fig. 9.3. The intersection of the cones provides a means for generating polarization-entangled states. We use the notation $|V\rangle$ and $|H\rangle$ to represent vertically and horizontally polarized *single-photon* states. For photons that emerge along the intersections of the cones, with photons from other parts of the cones being excluded by screens with pinholes in front of the intersection points, there will be an ambiguity as to whether the signal or idler photons will be vertically or horizontally polarized, as indicated in Fig. 9.4. The Hamiltonian describing this is given by

$$\hat{H}_{\rm I} = \hbar\eta\left(\hat{a}^{\dagger}_{\rm Vs}\hat{a}^{\dagger}_{\rm Hi} + \hat{a}^{\dagger}_{\rm Hs}\hat{a}^{\dagger}_{\rm Vi}\right) + H.c., \tag{9.6}$$

where the operators $\hat{a}^{\dagger}_{\rm Vs}$, $\hat{a}^{\dagger}_{\rm Hs}$, $\hat{a}^{\dagger}_{\rm Vi}$, and $\hat{a}^{\dagger}_{\rm Hi}$ are the creation operators for photons with vertical and horizontal polarization for the signal and idler beams,

respectively. Again, this Hamiltonian represents the post-selection obtained by placing a screen in front of the source with pinholes located just in the regions of the overlapping beams.

Let us take the initial state of the signal and idler modes to be represented by $|\Psi_0\rangle = |\{0\}\rangle$, which is the collective vacuum state for either type I or type II down-conversion. In either case, the state vector evolves according to

$$|\Psi(t)\rangle = \exp(-it\hat{H}_{\rm I}/\hbar)\,|\Psi_0\rangle\,, \tag{9.7}$$

which we expand, since $\hat{H}_{\rm I}$ has no explicit time dependence, as

$$|\Psi(t)\rangle \approx [1 - it\hat{H}_{\rm I}/\hbar + \tfrac{1}{2}(-it\hat{H}_{\rm I}/\hbar)^2]|\Psi_0\rangle \tag{9.8}$$

to second order in time. If we consider the type I SPDC then $|\Psi_0\rangle = |0\rangle_{\rm s}|0\rangle_{\rm i}$ and we have

$$|\Psi(t)\rangle = (1 - \mu^2/2)|0\rangle_{\rm s}|0\rangle_{\rm i} - i\mu|1\rangle_{\rm s}|1\rangle_{\rm i} \tag{9.9}$$

where $\mu = \eta t$. This state vector is normalized to first order in μ and we have dropped the term of order μ^2 containing the state $|2\rangle_{\rm s}|2\rangle_{\rm i}$. In the case of type II SPDC with the initial state $|\Psi_0\rangle = |0\rangle_{\rm Vs}|0\rangle_{\rm Hs}|0\rangle_{\rm Vi}|0\rangle_{\rm Hi}$, we have

$$\begin{aligned}|\Psi(t)\rangle &= (1 - \mu^2/2)|0\rangle_{\rm Vs}|0\rangle_{\rm Hs}|0\rangle_{\rm Vi}|0\rangle_{\rm Hi} \\ &\quad -i\mu\frac{1}{\sqrt{2}}(|1\rangle_{\rm Vs}|0\rangle_{\rm Hs}|0\rangle_{\rm Vi}|1\rangle_{\rm Hi} + |0\rangle_{\rm Vs}|1\rangle_{\rm Hs}|1\rangle_{\rm Vi}|0\rangle_{\rm Hi}).\end{aligned} \tag{9.10}$$

We now define the vertically and horizontally polarized vacuum and single photon states as $|0\rangle := |0\rangle_{\rm V}|0\rangle_{\rm H}$, $|V\rangle := |1\rangle_{\rm V}|0\rangle_{\rm H}$, and $|H\rangle := |0\rangle_{\rm V}|1\rangle_{\rm H}$, so that we may write

$$\begin{aligned}|\Psi(t)\rangle &= (1 - \mu^2/2)|0\rangle_{\rm s}|0\rangle_{\rm i} \\ &\quad -i\mu(|V\rangle_{\rm s}|H\rangle_{\rm i} + |H\rangle_{\rm s}|V\rangle_{\rm i}).\end{aligned} \tag{9.11}$$

The state in the second term, which when normalized reads

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|V\rangle_{\rm s}|H\rangle_{\rm i} + |H\rangle_{\rm s}|V\rangle_{\rm i}), \tag{9.12}$$

is one member out of a set of four states known as Bell states. The full set of Bell states is

$$|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|V\rangle_2 \pm |V\rangle_1|H\rangle_2), \tag{9.13}$$

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|H\rangle_2 \pm |V\rangle_1|V\rangle_2). \tag{9.14}$$

We shall discuss these states, and their implications, in Section 9.6.

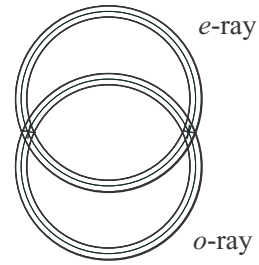


Fig. 9.4. The intersections of the cones for the *o*-ray and the *e*-ray are the sources for polarization entangled light. From these points it is not possible to tell from which beam a photon is obtained. The Hamiltonian in Eq. (9.6) describes the light taken from both intersections.

9.2 The Hong–Ou–Mandel interferometer

In Chapter 6 we discussed what happens when twin single-photon states are simultaneously incident at each input port of a 50:50 beam splitter: the photons

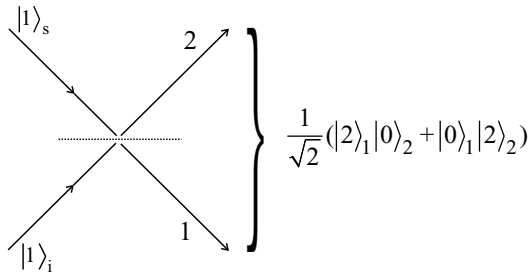


Fig. 9.5. Labeling of the output modes of a 50:50 beam splitter with inputs from the signal and idler beams of a type I down-converter. If single-photon states simultaneously fall onto the beam splitter, the output does not contain the term $|1\rangle_1|1\rangle_2$, for reasons discussed in Chapter 6.

emerge together in one output beam or the other; no single photons ever emerge in both beams. Recall that for the input state $|1\rangle_s|1\rangle_i$ one has after the beam splitter the state

$$|\psi_{BS}\rangle = \frac{1}{\sqrt{2}}(|2\rangle_1|0\rangle_2 + |0\rangle_1|2\rangle_2), \tag{9.15}$$

where we have labeled the output modes 1 and 2 in accordance with Fig. 9.5. Detectors placed at the outputs should never register simultaneous counts. In fact, one could take the lack of simultaneous counts as an indication that the photons are incident on the beam splitter simultaneously. The first demonstration of this effect was by Hong, Ou and Mandel (HOM) in a now-classic experiment performed in 1987 [6]. In fact, the experiment was designed to measure the time separation between the two photons striking a beam splitter. A sketch of their experiment is given in Fig. 9.6. The nonlinear crystal is pumped to produce, assuming type I down-conversion, twin single-photon states whose beams are then directed to the input ports of a 50:50 beam splitter. Photon detectors are placed at the outputs of the beam splitter and the count signals are fed into a correlator. Changing the position of the beam splitter causes a slight time delay between the times the photons fall onto the beam splitter. With a slight nonzero time delay, the photons independently reflect or transmit through the beam splitter causing both detectors sometimes to fire within a short time of each other. It can be shown [6] that the rate of coincident detections, R_{coin}, has the form

$$R_{coin} \sim \left[1 - e^{-(\Delta\omega)^2(\tau_s - \tau_i)^2}\right], \tag{9.16}$$

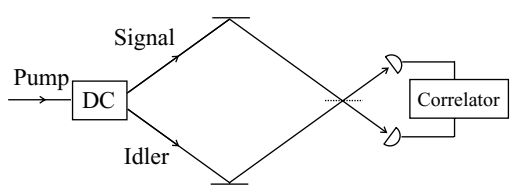


Fig. 9.6. The Hong–Ou–Mandel experiment. When the path lengths are equal, no coincident counts are detected.

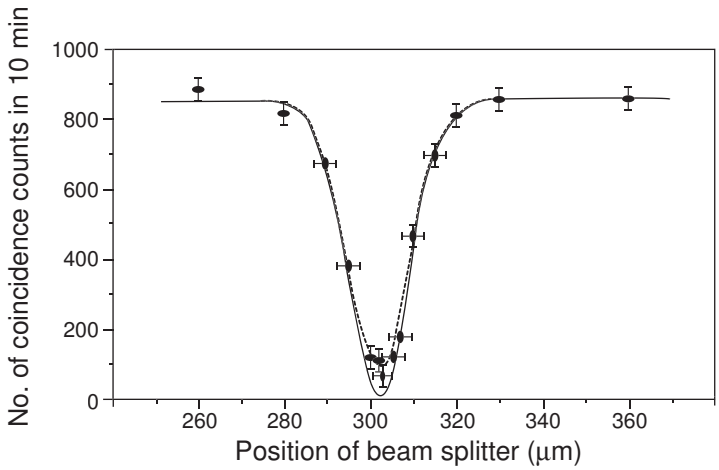


Fig. 9.7. Redrawn from the Hong–Ou–Mandel paper, the number of coincident counts in a 10 minute time interval as a function of the position of the beam splitter. The count rate does not go to zero exactly, owing to the fact that the beams do not perfectly overlap at the beam splitter. (Reprinted with permission.)

where $\Delta\omega$ is the bandwidth of the light and $c\tau_s$ and $c\tau_i$ are the distances that the signal and idler photons respectively travel from the down-converter to the beam splitter. The band width $\Delta\omega$ incorporates the reality that the signal and idler beams are not monochromatic, and its appearance in Eq. (9.16) results from the assumption that the spectral distribution is Gaussian. Obviously, for $\tau_s - \tau_i = 0$ we have $R_{coin} = 0$. The rate of coincidence counts rises to a maximum for $|\tau_s - \tau_i| \gg \tau_{corr}$, where $\tau_{corr} = 1/\Delta\omega$ is the correlation time of the photons. The correlation time is of the order of a few nanoseconds, which is hard to measure with conventional techniques as detectors commonly used do not often have short enough resolving times. But this correlation time can be measured with the HOM interferometer. The experimental results are plotted in Fig. 9.7, reprinted from Reference [6]. Plotted here is the number of counts over an interval of 10 min against the position of the beam splitter (essentially the time separation) with the solid line representing the theoretical prediction. The experimental data do not go exactly to the predicted minimum because it is not possible for the beams precisely to overlap at the beam splitter. From the distribution of the counts, the correlation time of the two photons can be determined to be $\sim$100 fs.

9.3 The quantum eraser

In the HOM experiment just described, the fact that the photons are indistinguishable is the key to understanding the results. Because type I SPDC is used as the source, both photons have the same polarization. They may have slightly different energies but the photon detectors are not really sensitive to the energy difference. We have not needed to specify the polarization direction of the photons; it was only important that they be the same. But now suppose, for the sake of definiteness, we take them to be horizontally polarized (see Fig. 9.8), denoting

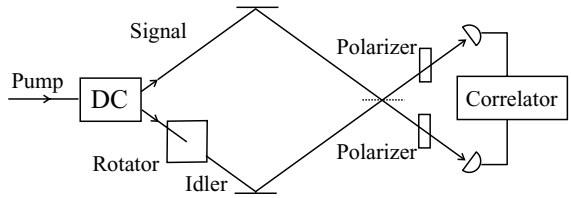


Fig. 9.8. A quantum erasure experiment. The rotator in the idler beam rotated the polarization of idler photons thus marking them and yielding path information. This destroys the quantum interference exhibited in the Hong–Ou–Mandel experiment. Inserting polarizers in the beam after the interferometer restores quantum interferences.

the twin single-photon states as $|H\rangle_s|H\rangle_i$. With such an input state to the beam splitter (assumed to not affect the polarization of the photons), the output state can be written as $i(|2H\rangle_1|0\rangle_2 + |0\rangle_1|2H\rangle_2)/\sqrt{2}$, where $|2H\rangle$ is a state with two horizontally polarized photons.

Suppose we now place a rotator in, say, the idler beam, such that its polarization is rotated by the angle θ to the horizontal, as indicated in Fig. 9.9. This transforms the idler polarization state, to

$$|\theta\rangle_i = |H\rangle_i \cos\theta + |V\rangle_i \sin\theta. \tag{9.17}$$

The polarization state vector for a photon with polarization orthogonal to this is given by

$$|\theta^\perp\rangle_i = -|H\rangle_i \sin\theta + |V\rangle_i \cos\theta. \tag{9.18}$$

The input state to the beam splitter is now

$$|H\rangle_s|\theta\rangle_i = \cos\theta|H\rangle_s|H\rangle_i + \sin\theta|H\rangle_s|V\rangle_i. \tag{9.19}$$

The output state will be

$$\begin{aligned}|\psi_{out}(\theta)\rangle = &\frac{i}{\sqrt{2}}\cos\theta(|2H\rangle_1|0\rangle_2 + |0\rangle_1|2H\rangle_2)\\ &+\frac{1}{\sqrt{2}}\sin\theta(|H\rangle_1|V\rangle_2 - |V\rangle_1|H\rangle_2),\end{aligned} \tag{9.20}$$

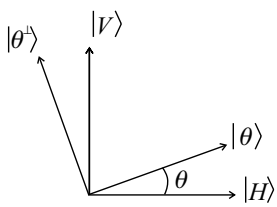


Fig. 9.9. The relationship between the $|\theta\rangle$, $|\theta^\perp\rangle$ and $|H\rangle$, $|V\rangle$ polarization vectors.

where the terms containing the single-photon states no longer cancel each other as the photons have different polarizations.* If the polarization of the idler beam is rotated all the way to the vertical, i.e. to $\theta = \pi/2$, then the output state will be

$$|\psi_{out}(\pi/2)\rangle = |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|V\rangle_2 - |V\rangle_1|H\rangle_2). \tag{9.21}$$

* Remember that to include the polarization state of a photon we really must expand the Hilbert space to write $|H\rangle_1 = |1\rangle_{H_1}|0\rangle_{V_1}$, etc., so that the last term will read $(|1\rangle_{H_1}|0\rangle_{V_1}|0\rangle_{H_2}|1\rangle_{V_2} - |0\rangle_{H_1}|1\rangle_{V_1}|1\rangle_{H_2}|0\rangle_{V_2})/\sqrt{2}$, which perhaps makes it more apparent why the single-photon states don't cancel as before.

There will now be *only* coincident counts in the detectors and neither detector, assuming 100% efficiency, will fire alone. The "marking" of one of the beams by the rotation of its polarization by 90° has the effect of removing photon indistinguishability and thus it is *possible* to determine the path taken by each of the photons before the beam splitter. Interestingly, in the scenario discussed, the photon polarization is never measured. The experimenter need not know the polarization of the photons; only the counts are measured. Evidently, the interference can be destroyed by the mere potential of obtaining which-path information, even if that information is never known to the experimenter. Now in the case of one-photon input to a beam splitter, one can observe the particle-like nature of the photon by placing detectors at the beam splitter outputs, or one can observe the wave-like nature, with interference effects, by using a Mach–Zehnder interferometer, as we have discussed in Chapter 6, providing an example of Bohr complementarity. Usually, it is the loss of coherence engendered by the availability of which-path information that is ascribed to the disappearance of interference. This "decoherence" supposedly renders the state vector into a statistical mixture where only probabilities, and not probability amplitudes, appear. But in the case of the two-photon interferometry experiment discussed here, there is no time, up to the point when the photons are detected, where we do not have a pure state.

Suppose now that a linear polarizer is placed at an angle θ_1 to the horizontal in front of the detector in output beam 1. The placing of a polarizer followed by photon detection constitutes a von Neumann projection (see Appendix B) onto the state vector

$$|\theta_1\rangle_1 = |H\rangle_1 \cos\theta_1 + |V\rangle_1 \sin\theta_1. \tag{9.22}$$

That is, only photons with polarization state $|\theta_1\rangle_1$ will be registered by the detector. The state of Eq. (9.21) is reduced to the pure state

$$\begin{aligned} |\psi, \theta_1\rangle &= \frac{|\theta_1\rangle\langle\theta_1 \mid \psi_{\text{out}}(\pi/2)\rangle}{\langle\psi_{\text{out}}(\pi/2) \mid \theta_1\rangle\langle\theta_1 \mid \psi_{\text{out}}(\pi/2)\rangle^{1/2}} \\ &= |\theta_1\rangle(|V\rangle_2 \cos\theta_1 - |H\rangle_2 \sin\theta_1) \end{aligned} \tag{9.23}$$

where the photon of mode 2 has a polarization orthogonal to that of the detected photon (see Eq. (9.18)). Suppose we similarly place a polarizer at the angle θ_2 to the horizontal in front of the detector for output beam 2. The probability that there will be coincident detections of a photon polarized at angle θ_1 and another at θ_2 is then

$$\begin{aligned} P_{\text{coin}} &= |\langle\theta_1|\langle\theta_2 \mid \psi_{\text{out}}(\pi/2)\rangle|^2 \\ &= \frac{1}{2}\sin^2(\theta_2 - \theta_1). \end{aligned} \tag{9.24}$$

From this result we see that the dip in the coincident count rate can be revived, depending on the relative angle of the polarizers. The effect of the polarizers placed just before the detectors is to erase the information encoded onto one

9.8 Applications of down-converted light to metrology without absolute standards

Finally in this chapter, we briefly discuss the application of down-converted light to problems of practical interest. These applications are possible because of the tight correlations of the photons that are spontaneously emitted during the down-conversion process as was originally shown by Burnham and Weinberg [3]. The virtue of using down-converted light is, as we shall see, that measurements can be performed to yield absolute results without the need of a calibrated standard.

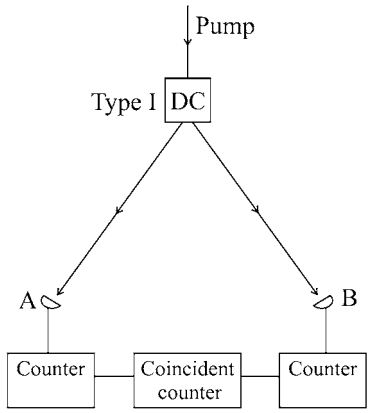


Fig. 9.15. Schematic for absolute calibration of a photon detector.

We first discuss the determination of the absolute calibration of a photon detector, essentially measuring its absolute quantum efficiency. A schematic of the measurement is given in Fig. 9.15. Two identical photon detectors are placed in the outputs of a type I down-converter. We regard detector A as the detector to be calibrated and detector B as the trigger, though which is which is quite arbitrary. Because the down-converter creates photons in pairs, coincident photons would be detected in both arms for ideal detectors. The measurement of the quantum efficiency is quite simple. Suppose that N is the number of pairs of photons produced by the down-converter. Whenever detector B fires, the experimenter checks to see if detector A also fires. Then the fraction of detections by B for which there is a coincident detection by A is the measured quantum efficiency of detector A. The method does not require knowledge of the quantum efficiency of detector B. If there is no detection by B, the experimenter simply disregards detector A altogether. More quantitatively, if we let the quantum efficiencies of detectors A and B be represented respectively by η_{A} and η_{B}, then the number of photons detected by each is $N_{\mathrm{A}} = \eta_{\mathrm{A}} N$ and $N_{\mathrm{B}} = \eta_{\mathrm{B}} N$. The number of coincident counts will be given by $N_{\mathrm{C}} = \eta_{\mathrm{A}}\eta_{\mathrm{B}} N = \eta_{\mathrm{A}} N_{\mathrm{B}}$. Thus $\eta_{\mathrm{A}} = N_{\mathrm{B}}/N_{\mathrm{C}}$, which is obviously independent of the quantum efficiency of detector B. The idea behind this sort of measurement is implicit in the work of Burnham and Weinberg [3]. Migdall *et al*. [31] at NIST performed a test of the method in 1995, comparing the method with a more-conventional method involving a standard, and found good agreement.

Appendix A

The density operator, entangled states, the Schmidt decomposition, and the von Neumann entropy

A.1 The density operator

Quantum-mechanical state vectors $|\psi\rangle$ convey the maximal amount of information about a system allowed by the laws of quantum mechanics. Typically, the information consists of quantum numbers associated with a set of commuting observables. Futhermore, if $|\psi_1\rangle$ and $|\psi_2\rangle$ are two possible quantum states then so is their coherent superposition

$$|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle \tag{A1}$$

if the coefficients c_1 and c_2 are known. If the states $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal ($\langle\psi_2 \mid \psi_1\rangle = 0$) then we must have $|c_1|^2 + |c_2|^2 = 1$. But there are frequently, in fact, more often than not, situations where the state vector is not precisely known. There are, for example, cases where the system of interest is interacting with some other system, possibly a very large system, e.g. a reservoir, with which it becomes entangled. It may be possible to write state vectors for the multicomponent system but not for the subsystem of interest. For example, for a system of two spin-1/2 particles with, the eigenstates of, say, the z-component of the spin denoted $|\uparrow\rangle$ for spin up and $|\downarrow\rangle$ for spin down, a possible state vector of the combined system is

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2\right], \tag{A2}$$

the so-called singlet state (total angular momentum zero), also known as one of the "Bell" states. Equation (A2) is an example of an entangled state. An entangled state cannot be factored, in any basis, into a product of states of the two subsystems, i.e.

$$|\psi\rangle \neq |\text{spin } 1\rangle\, |\text{spin } 2\rangle \quad \text{(for an entangled two-spin state)}\,. \tag{A3}$$

Entanglement is, apart from the superposition principle itself, an essential mystery of quantum mechanics, as was pointed out in 1935 by Schrödinger himself. Note though that entanglement follows from the superposition principle and is not something imposed on the theory. So Feynman's dictum that the superposition principle contains "the only mystery" is still correct.

Quantum states described by state vectors are said to be *pure* states. States that cannot be described by state vectors are said to be in *mixed* states. Mixed states are described by the density operator

$$\hat{\rho} = \sum_i |\psi_i\rangle\, p_i\, \langle\psi_i| = \sum_i p_i\, |\psi_i\rangle\, \langle\psi_i|, \tag{A4}$$

where the sum is over an ensemble (in the sense of statistical mechanics) where p_i is the probability of the system being in the ith state of the ensemble $|\psi_i\rangle$, where $\langle\psi_i \mid \psi_i\rangle = 1$. The probabilities satisfy the obvious relations

$$0 \le p_i \le 1, \quad \sum_i p_i = 1, \quad \sum_i p_i^2 \le 1. \tag{A5}$$

For the special case where all the p_i vanish except, say, the jth one, $p_i = \delta_{ij}$, we obtain

$$\hat{\rho} = |\psi_j\rangle\langle\psi_j|, \tag{A6}$$

the density operator for the pure state $|\psi_j\rangle$. Note that the density operator for this case is just the projection operator onto the state $|\psi_j\rangle$, and for the more general case of Eq. (A4), the density operator is a sum of the projection operators over the ensemble, weighted with the probabilities of each member of the ensemble.

We now introduce a complete, orthonormal, basis $\{|\varphi_n\rangle\}$ $(\sum_n |\varphi_n\rangle\langle\varphi_n| = \hat{I})$, eigenstates of some observable. Then for the ith member of the ensemble we may write

$$|\psi_i\rangle = \sum_n |\varphi_n\rangle\, \langle\varphi_n \mid \psi_i\rangle = \sum_n c_n^{(i)}\, |\varphi_n\rangle, \tag{A7}$$

where $c_n^{(i)} = \langle\varphi_n \mid \psi_i\rangle$. The matrix element of $\hat{\rho}$ between the n and n' eigenstates is

$$\langle\varphi_n\, |\hat{\rho}|\, \varphi_{n'}\rangle = \sum_i \langle\varphi_n|\psi_i\rangle\, p_i\, \langle\psi_i|\varphi_{n'}\rangle = \sum_i p_i c_n^{(i)} c_{n'}^{(i)*}. \tag{A8}$$

The quantities $\langle\varphi_n|\hat{\rho}|\varphi_{n'}\rangle$ form the elements of the density matrix. Taking the trace of this matrix we have

$$\begin{aligned} \mathrm{Tr}\hat{\rho} &= \sum_n \langle\varphi_n\, |\hat{\rho}|\, \varphi_n\rangle = \sum_i \sum_n \langle\varphi_n|\psi_i\rangle\, p_i\, \langle\psi_i|\varphi_n\rangle \\ &= \sum_i \sum_n p_i\, \langle\psi_i|\varphi_n\rangle\, \langle\varphi_n|\psi_i\rangle = \sum_i p_i = 1. \end{aligned} \tag{A9}$$

Since $\hat{\rho}$ is Hermitian (as is evident from its construction in Eq. (A4)), the diagonal elements $\langle\varphi_n|\hat{\rho}|\varphi_n\rangle$ must be real, and it follows from Eq. (A9) that

$$0 \le \langle\varphi_n|\, \hat{\rho}\, |\varphi_n\rangle \le 1. \tag{A10}$$

Now let us consider the square of the density operator: $\hat{\rho}^2 = \hat{\rho} \bullet \hat{\rho}$. For a pure state where $\hat{\rho} = |\psi\rangle\langle\psi|$ it follows that

$$\hat{\rho}^2 = |\psi\rangle\, \langle\psi \mid \psi\rangle\, \langle\psi| = |\psi\rangle\, \langle\psi| = \hat{\rho} \tag{A11}$$

and thus

$$\mathrm{Tr}\hat{\rho}^2 = \mathrm{Tr}\hat{\rho} = 1. \tag{A12}$$

For a statistical mixture

$$\hat{\rho}^2 = \sum_i \sum_j p_i p_j |\psi_i\rangle\langle\psi_i \mid \psi_j\rangle\langle\psi_j|. \tag{A13}$$

Taking the trace we have

$$\begin{aligned}
\mathrm{Tr}\hat{\rho}^2 &= \sum_n \langle\varphi_n| \hat{\rho}^2 |\varphi_n\rangle \\
&= \sum_n \sum_i \sum_j p_i p_j \langle\varphi_n \mid \psi_i\rangle\langle\psi_i \mid \psi_j\rangle\langle\psi_j \mid \varphi_n\rangle \\
&= \sum_i \sum_j p_i p_j |\langle\psi_i \mid \psi_j\rangle|^2 \\
&\leq \left[\sum_i p_i\right]^2 = 1.
\end{aligned} \tag{A14}$$

The equality holds only if $|\langle\psi_i \mid \psi_j\rangle|^2 = 1$ for *every* pair of states $|\psi_i\rangle$ and $|\psi_j\rangle$. This is possible only if all the $|\psi_i\rangle$ are collinear in Hilbert space, i.e. equivalent up to an overall phase factor. Thus we have the following criteria for pure and mixed states:

$$\begin{aligned}
&\mathrm{Tr}\hat{\rho}^2 = 1, \text{for a pure state,} \\
&\mathrm{Tr}\hat{\rho}^2 < 1, \text{ for a mixed state.}
\end{aligned} \tag{A15}$$

Perhaps a simple example is warranted at this point. Consider a superposition of, say, the vacuum and one-photon number states

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\phi}|1\rangle), \tag{A16}$$

where ϕ is just some phase. The density operator associated with this state is given by

$$\hat{\rho}_\psi = |\psi\rangle\langle\psi| = \frac{1}{2}[|0\rangle\langle 0| + |1\rangle\langle 1| + e^{i\phi}|1\rangle\langle 0| + e^{-i\phi}|0\rangle\langle 1|]. \tag{A17}$$

On the other hand, the density operator for an equally populated mixture of vacuum and one-photon states is

$$\hat{\rho}_\mathrm{M} = \frac{1}{2}[|0\rangle\langle 0| + |1\rangle\langle 1|]. \tag{A18}$$

The two density operators differ by the presence of the "off-diagonal", or "coherence", terms in the former, such terms being absent in the case of the mixture. The absence of the coherence terms is, of course, what makes the distinction between a state exhibiting full quantum-mechanical behavior and one that does not. It is easy to check that $\mathrm{Tr}\hat{\rho}_\mathrm{M}^2 = 1/2$.

For one of the states of the ensemble $|\psi_i\rangle$, by itself pure, the expectation value of some operator $\hat{O}$ is given by

$$\langle\hat{O}\rangle_i = \langle\psi_i|\hat{O}|\psi_i\rangle. \tag{A19}$$

For the statistical mixture, the ensemble average is given by

$$\langle \hat{O} \rangle = \sum_i p_i \, \langle \psi_i | \, \hat{O} \, | \psi_i \rangle, \tag{A20}$$

which is just the average of the quantum-mechanical expectation values weighted with the probabilities p_i. Formally we may write

$$\langle \hat{O} \rangle = \mathrm{Tr}(\hat{\rho} \, \hat{O}) \tag{A21}$$

since

$$\begin{aligned} \mathrm{Tr}(\hat{\rho} \, \hat{O}) &= \sum_n \langle \varphi_n | \, \hat{\rho} \, \hat{O} \, | \varphi_n \rangle \\ &= \sum_n \sum_i p_i \, \langle \varphi_n \mid \psi_i \rangle \, \langle \psi_i | \, \hat{O} \, | \varphi_n \rangle \\ &= \sum_i \sum_n p_i \, \langle \psi_i | \, \hat{O} \, | \varphi_n \rangle \, \langle \varphi_n \mid \psi_i \rangle \\ &= \sum_i p_i \, \langle \psi_i | \, \hat{O} \, | \psi_i \rangle \, . \end{aligned} \tag{A22}$$

A.3 Entangled states

Let us now consider a two-particle (or two-mode) system (known also as a bipartite system) and, for simplicity, let us assume that each particle can be in either of two one-particle states $|\psi_1\rangle$ or $|\psi_2\rangle$. Using the notation

$$\left|\psi_1^{(1)}\right\rangle, \text{ particle 1 in state 1,}$$
$$\left|\psi_2^{(1)}\right\rangle, \text{ particle 1 in state 2,}$$
$$\left|\psi_1^{(2)}\right\rangle, \text{ particle 2 in state 1,}$$
$$\left|\psi_2^{(2)}\right\rangle, \text{ particle 2 in state 2,}$$

we consider a pure two-particle superposition state (in general an entangled state)

$$|\Psi\rangle = C_1\left|\psi_1^{(1)}\right\rangle \otimes \left|\psi_2^{(2)}\right\rangle + C_2\left|\psi_2^{(1)}\right\rangle \otimes \left|\psi_1^{(2)}\right\rangle \tag{A28}$$

an example of which is given in Eq. (A2). (We have inserted the direct product symbol here for emphasis, although we generally assume it to be understood

and clear from context throughout this book.) Clearly, this can be extended for multiparticle (multipartite) systems. For such multipartite systems, we can define reduced density operators for each of the subsystems by tracing the density operator over the states of all the other systems. In the present case, with the density operator of the two-particle system given by $\hat{\rho} = |\Psi\rangle\langle\Psi|$, the reduced density operator for particle 1 is

$$\begin{aligned}\hat{\rho}^{(1)} &= \mathrm{Tr}_2\hat{\rho} = \left\langle \psi_1^{(2)}\middle|\hat{\rho}\middle|\psi_1^{(2)}\right\rangle + \left\langle \psi_2^{(2)}\middle|\hat{\rho}\middle|\psi_2^{(2)}\right\rangle \\ &= |C_1|^2\left|\psi_1^{(1)}\right\rangle\left\langle\psi_1^{(1)}\right| + |C_2|^2\left|\psi_2^{(1)}\right\rangle\left\langle\psi_2^{(1)}\right|. \end{aligned} \tag{A29}$$

This has the form for a mixed state for particle 1 as long as $C_i \neq 0,\ i = 1, 2$. Similarly, for particle 2,

$$\hat{\rho}^{(2)} = \mathrm{Tr}_1\hat{\rho} = |C_1|^2\left|\psi_1^{(2)}\right\rangle\left\langle\psi_1^{(2)}\right| + |C_2|^2\left|\psi_2^{(2)}\right\rangle\left\langle\psi_2^{(2)}\right|. \tag{A30}$$

Evidently, when one of the particles is considered without regard to the other, it is generally in a mixed state. Thus one may characterize the degree of entanglement according to the degree of purity of either of the subsystems. If $\mathrm{Tr}[\hat{\rho}^{(2)}]^2 = 1$, the state $|\Psi\rangle$ is not an entangled state; but if $\mathrm{Tr}[\hat{\rho}^{(2)}]^2 < 1$ we may conclude that $|\Psi\rangle$ describes an entanglement between subsystems 1 and 2.

Appendix B
Quantum measurement theory in a (very small) nutshell

The following is in no way meant to be an extensive review of the subject of quantum measurement theory. For more details, the reader should consult the bibliography. Here we present only those aspects of the theory necessary to understand the results of state reductive measurements, particularly for situations involving entangled states.

Suppose the operator $\hat{Q}$ is a Hermitian operator representing some observable Q. Further, we let the states $\{|q_n\rangle; n \text{ integer}\}$ be the eigenstates of $\hat{Q}$ with the eigenvalues being $q_n : \hat{Q}|q_n\rangle = q_n|q_n\rangle$. The eigenstates are complete and resolve unity according to

$$\sum_n |q_n\rangle\langle q_n| = \hat{I}. \tag{B1}$$

Each of the terms $|q_n\rangle\langle q_n|$ forms a projection operator onto the state with eigenvalue q_n, $\hat{P}(q_n) := |q_n\rangle\langle q_n|$. A state $|\psi\rangle$ (assumed normalized: $\langle\psi \mid \psi\rangle = 1$) can be expanded in terms of the eigenstates of $\hat{Q}$ according to

$$|\psi\rangle = \hat{I}\,|\psi\rangle = \sum_n |q_n\rangle\,\langle q_n \mid \psi\rangle = \sum_n c_n\,|q_n\rangle \tag{B2}$$

where the coefficient $c_n = \langle q_n \mid \psi\rangle$ is, in general, a complex number, and is known as a probability amplitude. If $|\psi\rangle$ is the state vector just before a measurement of the observable Q, the probability of obtaining the outcome q_n, $P(q_n)$, is

$$P(q_n) = \langle\psi|\hat{P}(q_n)|\psi\rangle = |c_n|^2. \tag{B3}$$

According to the orthodox (or Copenhagen) interpretation of quantum mechanics, the system described by the state vector $|\psi\rangle$ is *not* in one of the eigenstates of $\hat{Q}$ prior to the measurement, but rather, the state vector *collapses*, or *reduces*, to one of the eigenstates upon measurement, this processes sometimes represented by the symbols

$$|\psi\rangle \xrightarrow[\text{of } \hat{Q}]{\text{measurement}} |q_n\rangle. \tag{B4}$$

Prior to the measurement, the value of the observable associated with the operator $\hat{Q}$ in the system described by the state vector $|\psi\rangle$ is objectively indefinite. It is not merely a matter of only knowing what the value of the observable is

with a probability given by Eq. (B3), as would be the case if we were talking about statistical mechanics, but rather that one cannot assign definite values to observables for systems in superposition states without leading to conflict with experimental observations. The state reduction processes throws the system into a state with a definite value of the observable and subsequent measurements of that observable will return the same state. The dynamics of state reduction is not described by the Schrödinger equation. The interaction of the detector with the particle whose quantum state is being measured may destroy that particle. For example, when a photo-detector "clicks" it does so because the photon itself has been destroyed (absorbed).

It is sometimes possible to arrange experiments to perform selective, or filtered, measurements, sometimes called von Neumann projections. Such measurements are invoked in the text in Chapter 9 where polarization filters are employed and in Chapter 10 where field ionization is used in the context of cavity quantum electrodynamics. A projective measurement on one of the eigenstates $|q_n\rangle$ requires a filtration of all the other eigenstates of the operator. Mathematically we may represent this projection as $\hat{P}(q_n)|\psi\rangle$, which in normalized form reads

$$|\psi\rangle \xrightarrow[\text{measurement}]{\text{projective}} |q_n\rangle = \frac{\hat{P}(q_n)\,|\psi\rangle}{\langle\psi|\,\hat{P}(q_n)\,|\psi\rangle^{1/2}}. \tag{B5}$$

The filtering need not project onto an eigenstate of $\hat{Q}$. Suppose we could somehow filter the system so that the superposition state $|\psi_s\rangle = \frac{1}{\sqrt{2}}(|q_1\rangle + |q_2\rangle)$ is measured. The projection operator associated with this state is $\hat{P}_s = |\psi_s\rangle\langle\psi_s|$ and we have the projection

$$|\psi\rangle \xrightarrow[\text{measurement onto } |\psi_s\rangle]{\text{projective}} e^{i\varphi}\,|\psi_s\rangle = \frac{\hat{P}_s\,|\psi\rangle}{\langle\psi|\,\hat{P}_s\,|\psi\rangle^{1/2}}, \tag{B6}$$

where $\hat{P}_s|\psi\rangle = |\psi_s\rangle\langle\psi_s \mid \psi\rangle = |\psi_s\rangle(|c_1 + c_2|e^{i\varphi})/\sqrt{2}$, and $\langle\psi|\hat{P}_s|\psi\rangle = |c_1 + c_2|^2/2$. The phase φ is an irrelevant overall phase factor.

As a specific example, consider a single photon in the polarization state

$$|\theta\rangle = |H\rangle\cos\theta + |V\rangle\sin\theta \tag{B7}$$

where we have used the convention given in Chapter 9. If a Polaroid filter is placed in the beam and oriented along the horizontal (vertical) direction, the state vector of Eq. (B7) reduces to $|H\rangle(|V\rangle)$. On then other hand, if we try to project onto the state

$$|\theta^{\perp}\rangle = -|H\rangle\sin\theta + |V\rangle\cos\theta, \tag{B8}$$

where $\theta^{\perp} = \theta + \pi/2$, by orienting the filter along that direction, we find that we get no photons at all passing through: $\langle\theta^{\perp} \mid \theta\rangle = 0$. But if we orient the filter along some direction ϑ to the horizontal axis, the photon will be projected onto the polarization state

$$|\vartheta\rangle = |H\rangle\cos\vartheta + |V\rangle\sin\vartheta, \tag{B9}$$

as the reader can easily check.

Let us now turn to the case of two-mode states. If we have two modes labeled 1 and 2, we may write a general two-mode state as

$$|\psi\rangle = \sum_{n,m} c_{n\,m}\,|q_n\rangle_1\,|s_m\rangle_2\,, \tag{B10}$$

where the states $|q_n\rangle_1$ are eigenstates of $\hat{Q}_1$ defined in the Hilbert space of mode 1 and the states $|s_m\rangle$ are eigenstates of some operator $\hat{S}_2$ defined in the Hilbert space of mode 2. Certain choices of the coefficients $c_{n\,m}$ will render the state an entangled state. In either case, an unfiltered measurement of the observables $\hat{Q}_1$ and $\hat{S}_2$ reduces the state vector to

$$|\psi\rangle \xrightarrow[\text{of}\hat{Q}_1\ \text{and}\hat{S}_2]{\text{measurement}} |q_n\rangle_1\,|s_m\rangle_2\,, \tag{B11}$$

whereas, say, a projective measurement onto $|q_n\rangle_1$ reduces the state vector of Eq. (B10) to

$$\begin{aligned} |\psi\rangle \xrightarrow[\text{onto}\,|q_n\rangle_1]{\text{projection}} |\psi, q_n\rangle &= \frac{\hat{P}_1(q_n)|\psi\rangle}{\langle\psi|\hat{P}_1(q_n)|\psi\rangle^{1/2}} \\ &= |q_n\rangle_1 \frac{\sum\limits_m c_{n\,m}|s_m\rangle}{\sum\limits_m |c_{n\,m}|^2}. \end{aligned} \tag{B12}$$

The projection holds whether or not we have an entangled state, as do the more general projections of the type discussed above. Notice that if the original state *is* entangled, the projective measurement creates a factorized state. As an example, let us consider a state of the form

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle_1\,|V\rangle_2 - |V\rangle_1\,|H\rangle_2\right), \tag{B13}$$

one of the Bell states. Suppose that in mode 1 a Polaroid filter is placed at an angle of $\pi/4$ to the horizontal. This causes a projection onto the state

$$|\pi/4\rangle_1 = \frac{1}{\sqrt{2}}(|H\rangle_1 + |V\rangle_1) \tag{B14}$$

in that mode and reduces the original state vector to

$$|\psi\rangle \xrightarrow[\text{onto}|\pi/4\rangle_1]{\text{projection}} \frac{\hat{P}_1(\frac{\pi}{4})|\psi\rangle}{\langle\psi|\hat{P}_1(\frac{\pi}{4})|\psi\rangle^{1/2}} = |\pi/4\rangle_1 \frac{1}{\sqrt{2}}(|V\rangle_2 - |H\rangle_2). \tag{B15}$$

Note that a projective measurement of one part of an entangled system projects the other part into a particular state in a manner that is entirely predictable. The correlations exhibited by these projections are of a highly nonclassical nature and are ultimately responsible for the violations of Bell's inequalities as described in Chapter 9.

Appendix D
Nonlinear optics and spontaneous parametric down-conversion

Down-conversion results from a nonlinear interaction of pump radiation with media where the induced polarization is so strongly effected by the radiation that it deforms beyond the linear response that generates the usual dispersion and absorption. For our purposes, we can expand the nonlinear polarization in a power series in the applied radiation field. Crystals commonly used in nonlinear optics are highly anisotropic and their response is described in tensorial form according to

$$\hat{P}_i = \chi^{(1)}_{i,j}\hat{E}_j + \chi^{(2)}_{i,j,k}\hat{E}_j\hat{E}_k + \chi^{(3)}_{i,j,k,l}\hat{E}_j\hat{E}_k\hat{E}_l + \cdots \tag{D1}$$

where $\chi^{(m)}$ is the mth-order electric susceptibility tensor [1] and where repeated indices imply a sum. The energy density is then $\varepsilon_0 E_i P_i$ and thus the *second-order* contribution to the Hamiltonian, the interaction Hamiltonian, is

$$\hat{H}^{(2)} = \varepsilon_0 \int_V d^3\mathbf{r}\,\chi^{(2)}_{i,j,k}\hat{E}_i\hat{E}_j\hat{E}_k, \tag{D2}$$

where the integral is over the interaction volume. We now represent the components of the fields as Fourier integrals of the form

$$\hat{E}(\mathbf{r},t) = \int d^3\mathbf{k}\Big[\hat{E}^{(-)}(\mathbf{k})\,e^{-i[\omega(\mathbf{k})\,t-\mathbf{k}\cdot\mathbf{r}]} + \hat{E}^{(+)}(\mathbf{k})\,e^{i[\omega(\mathbf{k})\,t-\mathbf{k}\cdot\mathbf{r}]}\Big], \tag{D3}$$

where

$$\hat{E}^{(-)}(\mathbf{k}) = i\sqrt{\frac{2\pi\,\hbar\omega(\mathbf{k})}{V}}\;\hat{a}^{\dagger}(\mathbf{k}), \text{ and } \hat{E}^{(-)}(\mathbf{k}) = i\sqrt{\frac{2\pi\,\hbar\omega(\mathbf{k})}{V}}\;\hat{a}(\mathbf{k}). \tag{D4}$$

The operators $\hat{a}(\mathbf{k})$ and $\hat{a}^{\dagger}(\mathbf{k})$ are the annihilation and creation operators respectively of photons with momentum $\hbar\mathbf{k}$.* If we substitute field expressions of the above form into Eq. (D2) and retain only the terms important for the case when the signal and idler modes are initially in vacuum states, we obtain the interaction Hamiltonian

$$\hat{H}_{\mathrm{I}}(t) = \varepsilon_0 \int_V d^3\mathbf{r} \int d^3\mathbf{k}_s d^3\mathbf{k}_i\,\chi^{(2)}_{lmn} \times \hat{E}^{(+)}_{\mathrm{pl}} e^{i\left[\omega_p(\mathbf{k}_p)\,t-\mathbf{k}_p\cdot\mathbf{r}\right]}\hat{E}^{(-)}_{\mathrm{sm}} e^{-i[\omega_s(\mathbf{k}_s)\,t-\mathbf{k}_s\cdot\mathbf{r}]}\hat{E}^{(-)}_{\mathrm{in}} e^{-i[\omega_i(\mathbf{k}_i)\,t-\mathbf{k}_i\cdot\mathbf{r}]} + H.c. \tag{D5}$$

* The commutation relations for these operators where the wave vector is continuous take the form $[\hat{a}(\mathbf{k}), \hat{a}^{\dagger}(\mathbf{k}')] = \delta^{(3)}(\mathbf{k}-\mathbf{k}')$.

The conversion rates for the process depend on the second or electric susceptibility $\chi^{(2)}$ but typically have efficiencies in the range 10^{-7} to 10^{-11}, extremely low rates. For this reason, in order to obtain significant output in the signal and idler beams it is necessary to pump the medium with a very strong coherent field which we can model as a classical field obtained from a laser as long as we are interested in interactions over a short enough time such that depletion of pump photons can be ignored – the parametric approximation. The pump laser is usually in the ultraviolet while the photons arising from the down-conversion are usually in the visible spectral range.

From time dependent perturbation theory, assuming the initial states of the signal and idler modes are in vacuum states, which we denote for the moment as $|\Psi_0\rangle$, we obtain, to first order, $|\Psi\rangle \approx |\Psi_0\rangle + |\Psi_1\rangle$ where [2]

$$\begin{aligned} |\Psi_1\rangle &= -\frac{i}{\hbar}\int d\,t\hat{H}(t)|\Psi_0\rangle \\ &= \mathcal{N}\int d^3\mathbf{k}_s d^3\mathbf{k}_i \delta(\omega_p - \omega_s(\mathbf{k}_s) - \omega_i(\mathbf{k}_i)) \\ &\quad \times \delta^{(3)}(\mathbf{k}_p - \mathbf{k}_s - \mathbf{k}_i)\hat{a}_s^\dagger(\mathbf{k}_s)\hat{a}_i^\dagger(\mathbf{k}_i)|\Psi_0\rangle, \end{aligned} \tag{D6}$$

where $\mathcal{N}$ is a normalization factor into which all constants have been absorbed. One sees that the delta functions contain the phase matching conditions

$$\begin{aligned} \omega_p &= \omega_s + \omega_i, \\ \mathbf{k}_p &= \mathbf{k}_s + \mathbf{k}_i. \end{aligned} \tag{D7}$$

In the case of type I phase matching, we end up with the state given by Eq. (9.9), which we arrived at by assuming specific momenta which can be post-selected by the placement of a screen with properly located holes over the output of the down-converter.

References

[1] See R. W. Boyd, *Nonlinear Optics*, 2nd edition (New York: Academic Press, 2003).

[2] See Y. Shih, *Rep. Prog. Phys.*, **66** (2003), 1009.

1.2 Completely Symmetric and Antisymmetric States

We begin with the single-particle states $|i\rangle$: $|1\rangle$, $|2\rangle$, ... The single-particle states of the particles 1, 2, ..., α, ..., N are denoted by $|i\rangle_1$, $|i\rangle_2$, ..., $|i\rangle_\alpha$, ..., $|i\rangle_N$. These enable us to write the basis states of the N-particle system

$$|i_1, \ldots, i_\alpha, \ldots, i_N\rangle = |i_1\rangle_1 \ldots |i_\alpha\rangle_\alpha \ldots |i_N\rangle_N \,, \tag{1.2.1}$$

where particle 1 is in state $|i_1\rangle_1$ and particle α in state $|i_\alpha\rangle_\alpha$, etc. (The subscript outside the ket is the number labeling the particle, and the index within the ket identifies the state of this particle.)

Provided that the $\{|i\rangle\}$ form a complete orthonormal set, the product states defined above likewise represent a *complete orthonormal system* in the

space of N-particle states. The symmetrized and antisymmetrized basis states are then defined by

$$S_{\pm}\,|i_1, i_2, \ldots, i_N\rangle \equiv \frac{1}{\sqrt{N!}} \sum_P (\pm 1)^P P\,|i_1, i_2, \ldots, i_N\rangle \ . \tag{1.2.2}$$

In other words, we apply all $N!$ elements of the permutation group S_N of N objects and, for fermions, we multiply by (-1) when P is an odd permutation. The states defined in (1.2.2) are of two types: completely symmetric and completely antisymmetric.

Remarks regarding the properties of $S_{\pm} \equiv \frac{1}{\sqrt{N!}} \sum_P (\pm 1)^P P$:

(i) Let S_N be the permutation group (or symmetric group) of N quantities.

Assertion: For every element $P \in S_N$, one has $PS_N = S_N$.
Proof. The set PS_N contains exactly the same number of elements as S_N and these, due to the group property, are all contained in S_N. Furthermore, the elements of PS_N are all different since, if one had $PP_1 = PP_2$, then, after multiplication by P^{-1}, it would follow that $P_1 = P_2$.
Thus

$$PS_N = S_N P = S_N \ . \tag{1.2.3}$$

(ii) It follows from this that

$$PS_+ = S_+ P = S_+ \tag{1.2.4a}$$

and

$$PS_- = S_- P = (-1)^P S_- . \tag{1.2.4b}$$

If P is even, then even elements remain even and odd ones remain odd. If P is odd, then multiplication by P changes even into odd elements and vice versa.

$$PS_+\,|i_1, \ldots, i_N\rangle = S_+\,|i_1, \ldots, i_N\rangle$$
$$PS_-\,|i_1, \ldots, i_N\rangle = (-1)^P S_-\,|i_1, \ldots, i_N\rangle$$
Special case $P_{ij} S_-\,|i_1, \ldots, i_N\rangle = -S_-\,|i_1, \ldots, i_N\rangle$.

(iii) If $|i_1, \ldots, i_N\rangle$ contains single-particle states occurring more than once, then $S_+\,|i_1, \ldots, i_N\rangle$ is no longer normalized to unity. Let us assume that the first state occurs n_1 times, the second n_2 times, etc. Since $S_+\,|i_1, \ldots, i_N\rangle$ contains a total of $N!$ terms, of which $\frac{N!}{n_1! n_2! \ldots}$ are different, each of these terms occurs with a multiplicity of $n_1! n_2! \ldots$.

$$\langle i_1, \ldots, i_N|\, S_+^{\dagger} S_+\,|i_1, \ldots, i_N\rangle = \frac{1}{N!}(n_1! n_2! \ldots)^2 \frac{N!}{n_1! n_2! \ldots} = n_1! n_2! \ldots$$

Thus, the *normalized Bose basis functions* are

$$S_+ |i_1, \dots, i_N\rangle \frac{1}{\sqrt{n_1! n_2! \dots}} = \frac{1}{\sqrt{N! n_1! n_2! \dots}} \sum_P P |i_1, \dots, i_N\rangle . \quad (1.2.5)$$

(iv) A further property of $S_\pm$ is

$$S_\pm^2 = \sqrt{N!} S_\pm , \quad (1.2.6a)$$

since $S_\pm^2 = \frac{1}{\sqrt{N!}} \sum_P (\pm 1)^P P S_\pm = \frac{1}{\sqrt{N!}} \sum_P S_\pm = \sqrt{N!} S_\pm$. We now consider an arbitrary N-particle state, which we expand in the basis $|i_1\rangle \dots |i_N\rangle$

$$|z\rangle = \sum_{i_1, \dots, i_N} |i_1\rangle \dots |i_N\rangle \underbrace{\langle i_1, \dots, i_N | z\rangle}_{c_{i_1, \dots, i_N}} .$$

Application of $S_\pm$ yields

$$S_\pm |z\rangle = \sum_{i_1, \dots, i_N} S_\pm |i_1\rangle \dots |i_N\rangle \, c_{i_1, \dots, i_N} = \sum_{i_1, \dots, i_N} |i_1\rangle \dots |i_N\rangle \, S_\pm c_{i_1, \dots, i_N}$$

and further application of $\frac{1}{\sqrt{N!}} S_\pm$, with the identity (1.2.6a), results in

$$S_\pm |z\rangle = \frac{1}{\sqrt{N!}} \sum_{i_1, \dots, i_N} S_\pm |i_1\rangle \dots |i_N\rangle \, (S_\pm c_{i_1, \dots, i_N}). \quad (1.2.6b)$$

Equation (1.2.6b) implies that every symmetrized state can be expanded in terms of the symmetrized basis states (1.2.2).

1.3 Bosons

1.3.1 States, Fock Space, Creation and Annihilation Operators

The state (1.2.5) is fully characterized by specifying the occupation numbers

$$|n_1, n_2, \dots\rangle = S_+ |i_1, i_2, \dots, i_N\rangle \frac{1}{\sqrt{n_1! n_2! \dots}} . \quad (1.3.1)$$

Here, n_1 is the number of times that the state 1 occurs, n_2 the number of times that state 2 occurs, ... Alternatively: n_1 is the number of particles in state 1, n_2 is the number of particles in state 2, ... The sum of all occupation numbers n_i must be equal to the total number of particles:

$$\sum_{i=1}^{\infty} n_i = N. \quad (1.3.2)$$

Apart from this constraint, the n_i can take any of the values $0, 1, 2, \ldots$. The factor $(n_1! n_2! \ldots)^{-1/2}$, together with the factor $1/\sqrt{N!}$ contained in S_+, has the effect of normalizing $|n_1, n_2, \ldots\rangle$ (see point (iii)). These states form a complete set of completely symmetric N-particle states. By linear superposition, one can construct from these any desired symmetric N-particle state.

We now combine the states for $N = 0, 1, 2, \ldots$ and obtain a complete orthonormal system of states for arbitrary particle number, which satisfy the orthogonality relation[6]

$$\langle n_1, n_2, \ldots | n_1', n_2', \ldots\rangle = \delta_{n_1, n_1'} \delta_{n_2, n_2'} \ldots \tag{1.3.3a}$$

and the completeness relation

$$\sum_{n_1, n_2, \ldots} |n_1, n_2, \ldots\rangle \langle n_1, n_2, \ldots| = \mathbb{1} \, . \tag{1.3.3b}$$

This extended space is the *direct sum* of the space with no particles (vacuum state $|0\rangle$), the space with one particle, the space with two particles, etc.; it is known as *Fock space*.

The operators we have considered so far act only within a subspace of fixed particle number. On applying $\mathbf{p}, \mathbf{x}$ etc. to an N-particle state, we obtain again an N-particle state. We now define *creation* and *annihilation operators*, which lead from the space of N-particle states to the spaces of $N \pm 1$-particle states:

$$a_i^\dagger \, |\ldots, n_i, \ldots\rangle = \sqrt{n_i + 1} \, |\ldots, n_i + 1, \ldots\rangle \, . \tag{1.3.4}$$

Taking the adjoint of this equation and relabeling $n_i \to n_i'$, we have

$$\langle \ldots, n_i', \ldots| \, a_i = \sqrt{n_i' + 1} \, \langle \ldots, n_i' + 1, \ldots| \, . \tag{1.3.5}$$

Multiplying this equation by $|\ldots, n_i, \ldots\rangle$ yields

$$\langle \ldots, n_i', \ldots| \, a_i \, |\ldots, n_i, \ldots\rangle = \sqrt{n_i} \, \delta_{n_i', n_i} \, .$$

Expressed in words, the operator a_i reduces the occupation number by 1.
Assertion:

$$a_i \, |\ldots, n_i, \ldots\rangle = \sqrt{n_i} \, |\ldots, n_i - 1, \ldots\rangle \quad \text{for } n_i \geq 1 \tag{1.3.6}$$

and

$$a_i \, |\ldots, n_i = 0, \ldots\rangle = 0 \, .$$

[6] In the states $|n_1, n_2, \ldots\rangle$, the n_1, n_2 etc. are arbitrary natural numbers whose sum is not constrained. The (vanishing) scalar product between states of differing particle number is defined by (1.3.3a).

Proof:

$$
\begin{aligned}
a_i \,|\ldots, n_i, \ldots\rangle &= \sum_{n_i'=0}^{\infty} |\ldots, n_i', \ldots\rangle \,\langle \ldots, n_i', \ldots| \, a_i \,|\ldots, n_i, \ldots\rangle \\
&= \sum_{n_i'=0}^{\infty} |\ldots, n_i', \ldots\rangle \,\sqrt{n_i}\, \delta_{n_i'+1, n_i} \\
&= \begin{cases} \sqrt{n_i}\,|\ldots, n_i - 1, \ldots\rangle & \text{for } n_i \geq 1 \\ 0 & \text{for } n_i = 0 \end{cases} .
\end{aligned}
$$

The operator $a_i^\dagger$ increases the occupation number of the state $|i\rangle$ by 1, and the operator a_i reduces it by 1. The operators $a_i^\dagger$ and a_i are thus called *creation* and *annihilation operators.* The above relations and the completeness of the states yield the *Bose commutation relations*

$$
[a_i, a_j] = 0, \quad [a_i^\dagger, a_j^\dagger] = 0, \quad [a_i, a_j^\dagger] = \delta_{ij} \,. \tag{1.3.7a,b,c}
$$

Proof. It is clear that (1.3.7a) holds for $i = j$, since a_i commutes with itself. For $i \neq j$, it follows from (1.3.6) that

$$
\begin{aligned}
a_i a_j \,|\ldots, n_i, \ldots, n_j, \ldots\rangle &= \sqrt{n_i}\sqrt{n_j}\,|\ldots, n_i - 1, \ldots, n_j - 1, \ldots\rangle \\
&= a_j a_i \,|\ldots, n_i, \ldots, n_j, \ldots\rangle
\end{aligned}
$$

which proves (1.3.7a) and, by taking the hermitian conjugate, also (1.3.7b).
For $j \neq i$ we have

$$
\begin{aligned}
a_i a_j^\dagger \,|\ldots, n_i, \ldots, n_j, \ldots\rangle &= \sqrt{n_i}\sqrt{n_j + 1}\,|\ldots, n_i - 1, \ldots, n_j + 1, \ldots\rangle \\
&= a_j^\dagger a_i \,|\ldots, n_i, \ldots, n_j, \ldots\rangle
\end{aligned}
$$

and

$$
\begin{aligned}
&\left(a_i a_i^\dagger - a_i^\dagger a_i\right) |\ldots, n_i, \ldots, n_j, \ldots\rangle = \\
&\qquad \left(\sqrt{n_i + 1}\sqrt{n_i + 1} - \sqrt{n_i}\sqrt{n_i}\right) |\ldots, n_i, \ldots, n_j, \ldots\rangle
\end{aligned}
$$

hence also proving (1.3.7c).

Starting from the *ground state* $\equiv$ *vacuum state*

$$
|0\rangle \equiv |0, 0, \ldots\rangle \,, \tag{1.3.8}
$$

which contains no particles at all, we can construct all states:
single-particle states

$$
a_i^\dagger \,|0\rangle \,, \ldots \,,
$$

two-particle states

$$
\frac{1}{\sqrt{2!}} \left(a_i^\dagger\right)^2 |0\rangle \,, a_i^\dagger a_j^\dagger \,|0\rangle \,, \ldots
$$

and the general many-particle state

$$|n_1, n_2, \ldots\rangle = \frac{1}{\sqrt{n_1! n_2! \ldots}} \left(a_1^\dagger\right)^{n_1} \left(a_2^\dagger\right)^{n_2} \ldots |0\rangle \,. \tag{1.3.9}$$

Normalization:

$$a^\dagger |n-1\rangle = \sqrt{n}\,|n\rangle \tag{1.3.10}$$

$$\left\| a^\dagger |n-1\rangle \right\| = \sqrt{n}$$

$$|n\rangle = \frac{1}{\sqrt{n}} a^\dagger |n-1\rangle \ .$$

1.4 Fermions

1.4.1 States, Fock Space, Creation and Annihilation Operators

For fermions, one needs to consider the states $S_- |i_1, i_2, \dots, i_N\rangle$ defined in (1.2.2), which can also be represented in the form of a determinant:

$$S_- |i_1, i_2, \dots, i_N\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} |i_1\rangle_1 & |i_1\rangle_2 & \cdots & |i_1\rangle_N \\ \vdots & \vdots & \ddots & \vdots \\ |i_N\rangle_1 & |i_N\rangle_2 & \cdots & |i_N\rangle_N \end{vmatrix} \quad . \tag{1.4.1}$$

The determinants of one-particle states are called *Slater determinants*. If any of the single-particle states in (1.4.1) are the same, the result is zero. This is a statement of the Pauli principle: two identical fermions must not occupy the same state. On the other hand, when all the i_α are different, then this antisymmetrized state is normalized to 1. In addition, we have

$$S_- \, |i_2, i_1, \ldots\rangle = -S_- \, |i_1, i_2, \ldots\rangle \, . \tag{1.4.2}$$

This dependence on the order is a general property of determinants.

Here, too, we shall characterize the states by specifying their occupation numbers, which can now take the values 0 and 1. The state with n_1 particles in state 1 and n_2 particles in state 2, etc., is

$$|n_1, n_2, \ldots\rangle \, .$$

The state in which there are no particles is the vacuum state, represented by

$$|0\rangle = |0, 0, \ldots\rangle \ .$$

This state must not be confused with the null vector!

We combine these states (vacuum state, single-particle states, two-particle states, ...) to give a state space. In other words, we form the direct sum of the state spaces for the various fixed particle numbers. For fermions, this space is once again known as *Fock space*. In this state space a scalar product is defined as follows:

$$\langle n_1, n_2, \ldots | n_1{}', n_2{}', \ldots\rangle = \delta_{n_1, n_1{}'} \delta_{n_2, n_2{}'} \ldots \ ; \tag{1.4.3a}$$

i.e., for states with equal particle number (from a single subspace), it is identical to the previous scalar product, and for states from different subspaces it always vanishes. Furthermore, we have the completeness relation

$$\sum_{n_1=0}^{1} \sum_{n_2=0}^{1} \ldots |n_1, n_2, \ldots\rangle \, \langle n_1, n_2, \ldots| = \mathbb{1} \, . \tag{1.4.3b}$$

Here, we wish to introduce creation operators $a_i^\dagger$ once again. These must be defined such that the result of applying them twice is zero. Furthermore, the order in which they are applied must play a role. We thus define the creation operators $a_i^\dagger$ by

$$\begin{aligned} S_- \, |i_1, i_2, \ldots, i_N\rangle &= a_{i_1}^\dagger a_{i_2}^\dagger \ldots a_{i_N}^\dagger \, |0\rangle \\ S_- \, |i_2, i_1, \ldots, i_N\rangle &= a_{i_2}^\dagger a_{i_1}^\dagger \ldots a_{i_N}^\dagger \, |0\rangle \, . \end{aligned} \tag{1.4.4}$$

Since these states are equal except in sign, the anticommutator is

$$\{a_i^\dagger, a_j^\dagger\} = 0, \tag{1.4.5a}$$

which also implies the impossibility of double occupation

$$\left(a_i^\dagger\right)^2 = 0. \tag{1.4.5b}$$

The anticommutator encountered in (1.4.5a) and the commutator of two operators A and B are defined by

$$\begin{aligned} \{A,B\} &= [A,B]_+ = AB + BA \\ [A,B] &= [A,B]_- = AB - BA\,, \end{aligned} \tag{1.4.6}$$

Given these preliminaries, we can now address the precise formulation. If one wants to characterize the states by means of occupation numbers, one has to choose a particular ordering of the states. This is arbitrary but, once chosen, must be adhered to. The states are then represented as

$$|n_1, n_2, \ldots\rangle = \left(a_1^\dagger\right)^{n_1} \left(a_2^\dagger\right)^{n_2} \ldots |0\rangle \;, \quad n_i = 0, 1. \tag{1.4.7}$$

The effect of the operator $a_i^\dagger$ must be

$$a_i^\dagger \,|\ldots, n_i, \ldots\rangle = (1 - n_i)(-1)^{\sum_{j<i} n_j} \,|\ldots, n_i + 1, \ldots\rangle\,. \tag{1.4.8}$$

The number of particles is increased by 1, but for a state that is already occupied, the factor $(1 - n_i)$ yields zero. The phase factor corresponds to the number of anticommutations necessary to bring the $a_i^\dagger$ to the position i.
The adjoint relation reads:

$$\langle\ldots, n_i, \ldots|\, a_i = (1 - n_i)(-1)^{\sum_{j<i} n_j} \,\langle\ldots, n_i + 1, \ldots|\,. \tag{1.4.9}$$

This yields the matrix element

$$\langle\ldots, n_i, \ldots|\, a_i \,|\ldots, n_i', \ldots\rangle = (1 - n_i)(-1)^{\sum_{j<i} n_j} \delta_{n_i+1, n_i'}. \tag{1.4.10}$$

We now calculate

$$\begin{aligned} a_i \,|\ldots, n_i', \ldots\rangle &= \sum_{n_i} |n_i\rangle \,\langle n_i|\, a_i \,|n_i'\rangle \\ &= \sum_{n_i} |n_i\rangle \,(1 - n_i)(-1)^{\sum_{j<i} n_j} \delta_{n_i+1, n_i'} \\ &= (2 - n_i')(-1)^{\sum_{j<i} n_j} \,|\ldots, n_i' - 1, \ldots\rangle \, n_i'. \end{aligned} \tag{1.4.11}$$

Here, we have introduced the factor n_i', since, for $n_i' = 0$, the Kronecker delta $\delta_{n_i+1, n_i'} = 0$ always gives zero. The factor n_i' also ensures that the right-hand side cannot become equal to the state $|\ldots, n_i' - 1, \ldots\rangle = |\ldots, -1, \ldots\rangle$.

To summarize, the effects of the creation and annihilation operators are

$$\begin{aligned} a_i^\dagger \,|\ldots, n_i, \ldots\rangle &= (1 - n_i)(-1)^{\sum_{j<i} n_j} \,|\ldots, n_i + 1, \ldots\rangle \\ a_i \,|\ldots, n_i, \ldots\rangle &= n_i (-1)^{\sum_{j<i} n_j} \,|\ldots, n_i - 1, \ldots\rangle\,. \end{aligned} \tag{1.4.12}$$

It follows from this that

$$a_i a_i^\dagger \,|\ldots, n_i, \ldots\rangle = (1-n_i)(-1)^{2\sum_{j<i} n_j}(n_i+1)\,|\ldots, n_i, \ldots\rangle$$
$$= (1-n_i)\,|\ldots, n_i, \ldots\rangle \tag{1.4.13a}$$
$$a_i^\dagger a_i \,|\ldots, n_i, \ldots\rangle = n_i(-1)^{2\sum_{j<i} n_j}(1-n_i+1)\,|\ldots, n_i, \ldots\rangle$$
$$= n_i\,|\ldots, n_i, \ldots\rangle\ , \tag{1.4.13b}$$

since for $n_i \in \{0,1\}$ we have $n_i^2 = n_i$ and $(-1)^{2\sum_{j<i} n_j} = 1$. On account of the property (1.4.13b) one can regard $a_i^\dagger a_i$ as the occupation-number operator for the state $|i\rangle$. By taking the sum of (1.4.13a,b), one obtains the anticommutator

$$[a_i, a_i^\dagger]_+ = 1.$$

In the anticommutator $[a_i, a_j^\dagger]_+$ with $i \neq j$, the phase factor of the two terms is different:

$$[a_i, a_j^\dagger]_+ \propto (1-n_j)n_i(1-1) = 0\ .$$

Likewise, $[a_i, a_j]_+$ for $i \neq j$, also has different phase factors in the two summands and, since $a_i a_i\,|\ldots, n_i, \ldots\rangle \propto n_i(n_i-1) = 0$, one obtains the following *anticommutation rules* for *fermions:*

$$[a_i, a_j]_+ = 0,\quad [a_i^\dagger, a_j^\dagger]_+ = 0,\quad [a_i, a_j^\dagger]_+ = \delta_{ij}. \tag{1.4.14}$$

III. THE OPTICAL FIBER

Figure 7-7a shows an optical fiber, which consists of a (cylindrical) central dielectric core clad by a material of slightly lower refractive index. The corresponding refractive index distribution (in the transverse direction) is given by:

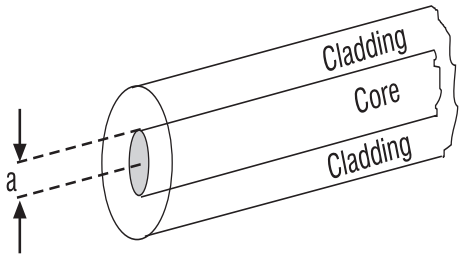
$$\begin{aligned} n &= n_1 \quad \text{for} \quad r < a \\ n &= n_2 \quad \text{for} \quad r > a \end{aligned} \tag{7-4}$$

where n_1 and $n_2\,(<n_1)$ represent respectively the refractive indices of core and cladding and a represents the radius of the core. We define a parameter Δ through the following equations.

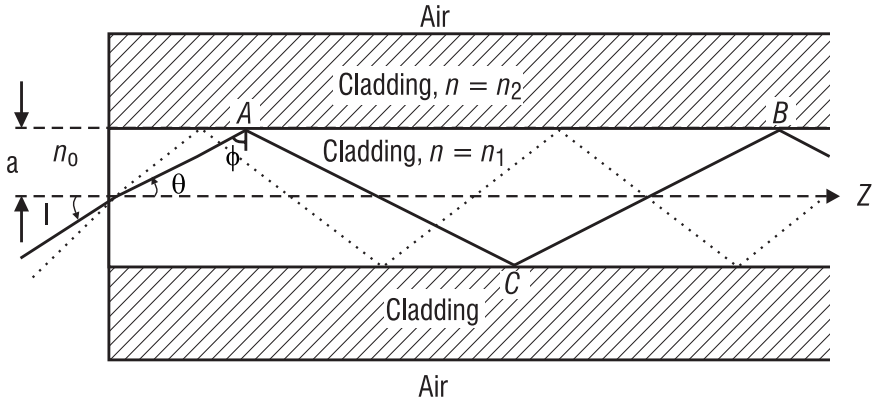
$$\Delta \equiv \frac{n_1^2 - n_2^2}{2n_2^2} \tag{7-5}$$

When $\Delta << 1$ (as is indeed true for silica fibers where n_1 is very nearly equal to n_2) we may write

$$\Delta = \frac{(n_1 + n_2)(n_1 - n_2)}{2n_1^2} \approx \frac{(n_1 - n_2)}{n_1} \approx \frac{(n_1 - n_2)}{n_2} \tag{7-6}$$



(a)



(b)

Figure 7-7 *(a) A glass fiber consists of a cylindrical central core clad by a material of slightly lower refractive index. (b) Light rays impinging on the core-cladding interface at an angle greater than the critical angle are trapped inside the core of the fiber.*

From the discussion in the above example it follows that, for a very-high-information-carrying system, it is necessary to reduce the pulse dispersion. Two alternative solutions exist—one involves the use of near-parabolic-index fibers and the other involves single-mode fibers. We look at these next.

VIII. PARABOLIC-INDEX FIBERS (PIF)

In a step-index fiber such as that pictured in Figure 7-7, the refractive index of the core has a constant value. By contrast, in a parabolic-index fiber, the refractive index in the core decreases continuously (in a quadratic fashion) from a maximum value at the center of the core to a constant value at the core-cladding interface. The refractive index variation is given by

$$\begin{aligned} n^2(r) &= n_1^2\left[1-2\Delta\left(\frac{r}{a}\right)^2\right] \qquad 0<r<a \quad \text{core} \\ n^2(r) &= n_2^2 = n_1^2(1-2\Delta) \qquad r>a \quad \text{cladding} \end{aligned} \tag{7-21}$$

with Δ as defined in Equation 7-5. For a typical (multimode) parabolic-index silica fiber, $\Delta \approx 0.01$, $n_2 \approx 1.45$, and $a \approx 25$ μm. On the other hand, for a typical plastic fiber (see Section XII), $n_1 \approx 1.49$, $n_2 \approx 1.40$, and $a \approx 500$ μm.

Since the refractive index decreases as one moves away from the center of the core, a ray entering the fiber is continuously bent toward the axis of the fiber, as depicted in Figure 7-12.

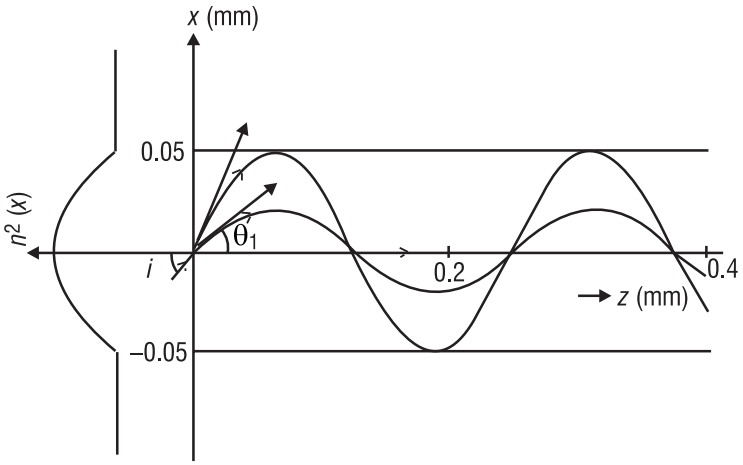


Figure 7-12 *Different ray paths in a parabolic-index fiber*

This follows from Snell's law because the ray continuously encounters a medium of lower refractive index and hence bends continuously away from the normal. Even though rays making larger angles with the fiber axis traverse a longer path, they do so now in a region of lower refractive index (and hence greater speed). The longer path length is almost compensated for by a greater average speed such that all rays take approximately the same amount of time in traversing the fiber. This leads to a much smaller pulse dispersion. The detailed calculations are

IX. MATERIAL DISPERSION

We first define the *group index*. To do this we return to Equation 7-1 where we noted that the velocity of light in a medium is given by

$$\boxed{v = c/n} \tag{7-24}$$

Here n is the refractive index of the medium, which, in general, depends on the wavelength. The dependence of the refractive index on wavelength leads to what is known as *dispersion,* discussed in Module 1.3, *Basic Geometrical Optics*. In Figure 7-13 we have shown a narrow pencil of a white light beam incident on a prism. Since the refractive index of glass depends on the wavelength, the angle of refraction will be different for different colors. For example, for crown glass the refractive indices at 656.3 nm (orange), 589.0 nm (yellow), and 486.1 nm (green) are respectively given by 1.5244, 1.5270, and 1.5330. Thus, if the angle of incidence $i = 45°$ the angle of refraction, r, will be $r = 27.64°$, 27.58°, and 27.47° for the orange, yellow, and blue colors respectively. The incident white light will therefore *disperse* into its constituent colors—the dispersion will become more evident at the second surface of the prism as seen in Figure 7-13.

XI. SINGLE-MODE FIBERS

While discussing step-index fibers, we considered light propagation inside the fiber as a set of many rays bouncing back and forth at the core-cladding interface (see Figure 7-7). There the angle θ could take a continuum of values lying between 0 and $\cos^{-1}(n_2/n_1)$, i.e.,

$$0 < \theta < \cos^{-1}(n_2/n_1)$$

For $n_2 = 1.5$ and $\Delta \approx \frac{n_1 - n_2}{n_1} = 0.01$, we would get $n_2/n_1 \simeq$ and $\cos^{-1}(\frac{n_2}{n_1}) = 8.1°$, so

$$0 < \theta < 8.1°$$

Now, when the core radius (or the quantity Δ) becomes very small, ray optics does not remain valid and one has to use the more accurate wave theory based on Maxwell's equations. This wave theory is fairly involved [see, e.g., Chapters 7 and 8 of Ghatak and Thyagarajan]. Here we just give some of the important results for a step-index fiber.

In wave theory, one introduces the parameter

$$V = \frac{2\pi}{\lambda_0} a\sqrt{n_1^2 - n_2^2} = \frac{2\pi}{\lambda_0} a n_1 \sqrt{2\Delta} \approx \frac{2\pi}{\lambda_0} a n_2 \sqrt{2\Delta} \qquad (7\text{-}28)$$

where Δ has been defined earlier (see Equation 7-5) and $n^1 \simeq n^2$. The quantity V is often referred to as the "*V-number*" or the "*waveguide parameter*" of the fiber. It can be shown that, if

$$V < 2.4045$$

only one guided mode (as if there is only one discrete value of θ) is possible and the fiber is known as a *single-mode fiber*. Further, for a step-index single-mode fiber, the corresponding (discrete) value of θ is approximately given by the following empirical formula

$$\cos\theta \approx 1 - \Delta\left[1 - \left(1.1428 - \frac{0.996}{V}\right)^2\right] \tag{7-29}$$

We may mention here that because of practical considerations the value of Δ ranges from about 0.002 to about 0.008.

Fiber Optics: How Fused Fiber Optic Couplers Work

Introduction

This technical note will describe how a fused optical fiber coupler works and how it is made. Once an understanding of this has been obtained, some of the principals of operation will be made clear at once (bi-directionality, for example).

The Fused Biconical Taper Process

A fused coupler basically consists of two, parallel optical fibers that have been twisted, stretched and fused together so that their cores are very close to each other. This forms a Coupling Region as shown in Figure 1 below. The length of this Coupling Region, L, determines the coupling ratio from one fiber to the other. During the manufacturing process, light is launched into an input port, P, and the output power from each output port is carefully monitored. When the desired coupling ratio is achieved, the fully automated manufacturing process is stopped. The resulting coupler is essentially one fiber with two cores that are very near to one another. This process is known as the Fused Biconical Taper (FBT) process.

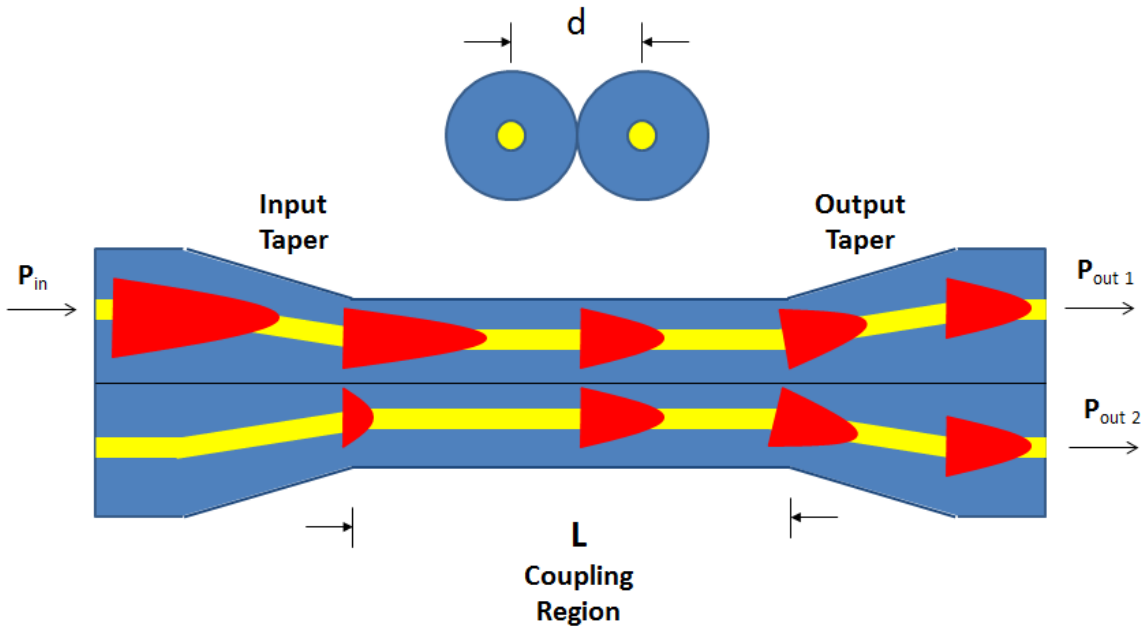


Fig. 1 The Fused Biconical Taper process.

How a Fused Coupler Works

The intensity profile of light traveling down a singlemode fiber is essentially Gaussian; that is, the intensity is greatest in the center and tapers off as the core/cladding interface is approached. The tail ends of the Gaussian profile extend slightly through the core and into the cladding. This tail is called the evanescent wave. Figure 2 represents the cross section of a light wave in an optical fiber. The vertical dashed lines represent the fiber core/cladding boundary. The red tails are the evanescent wave.

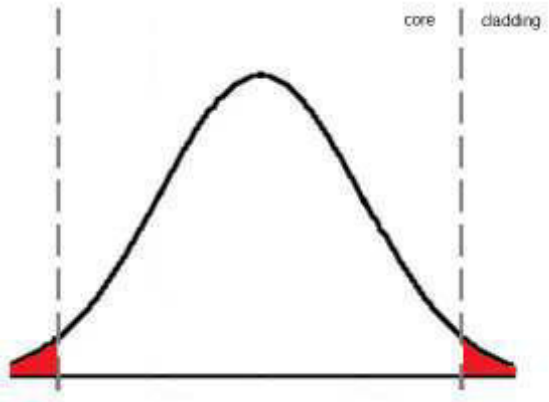


Figure 2. Light propagating down an optical fiber.
The red region represents the evanescent wave.

In the FBT process the cores of two identical parallel fibers are so close to one another that the evanescent wave can “leak” from one fiber core into the other core. This allows an exchange of energy, analogous to the energy exchange that takes place with two coupled pendulums. In fact, this is an excellent analogy. The amount of energy exchange is dependent upon the proximity of the two cores, d, and the length over which this exchange takes place, L (see Figure 1). It is easy to see that if the coupling length is long enough, a complete transfer of energy can take place from one core into the other. If the length is longer still, the process will continue, shifting the energy back into the original core. By selecting the proper length, any given power transfer ratio can be realized. This is how a 50/50 or a 10/90 coupler is made.

Let’s consider a few examples using Figure 3. Assume that this device is a 50/50 coupler.

Figure 3. A 1x2 or a 2x2 coupler.

1.2 Basic Principles of Operation

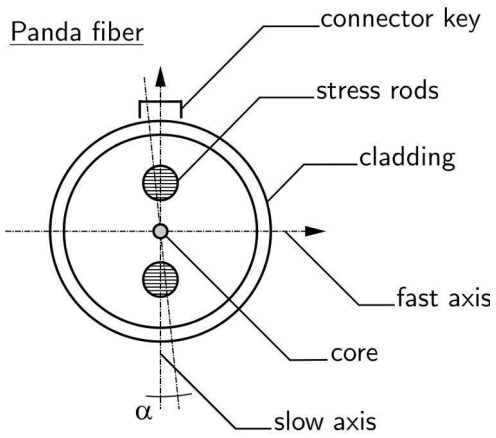
As pointed out in the previous section, the qu2PI is not an interferometer in the traditional sense, i.e. the light is not first split and later recombined like in standard Michelson or Mach-Zehnder interferometers. Rather the time-correlated photons provided by a photon-pair source enter the two input ports of the balanced beam-splitter. In the qu2PI, the fused 50:50 polarization maintaining fiber coupler is used as the beam-splitter. This all-fiber solution ensures a good spatial-mode overlap of the interfering photons, and provided that the linear polarization with correct orientation is launched into the input fibers of the coupler, also a good polarization-mode overlap of the photons. The latter is ensured due to the use of polarization-maintaining fibers, which preserve the linear-polarization of the light that is launched into the fibers; (for a short introduction to polarization maintaining fiber technology please see the boxed text on the next page). Moreover, the coupler is designed such that for linear polarization, which is maintained in the fibers, the splitting ratio is very close to the ideal 50:50 at the operating wavelength.

Either arm of the two-photon interferometer contains a short free-space optical line, i.e. the fiber-coupled photons coming from the source are collimated using a lens and after a short free-space transmission collected again into the input fibers of the coupler. The free-space lines have a double role. First, the two rotatable wave plates inserted in either arm allow to adjust the polarization of the incoming photon pairs as required. Second, one of the lines has a variable length, which makes the tuning of the path difference between the two arms possible. The tuning is controlled by the manual translation stage equipped with the differential micrometer screw. This enables the required micrometer-resolved steps for achieving the precise time overlap of the photons and scanning through the interference dip. For maximum spectral overlap the qu2PI contains two removable interference filters with transmission bandwidth significantly smaller compared to the natural spectral width of the photons. The insertion of filters thus leads not only to higher interference visibility but also to a wider interference dip. This is particularly useful when searching the dip, because larger scanning steps are possible.

Polarization maintaining fiber technology

All standard optical fibers are to some degree birefringent due to manufacturing imperfections, which break the perfect circular symmetry of the fiber core. Moreover, any mechanical or thermal stresses exerted on the fibres, change the birefringence within the fiber core. Birefringence means that the two degenerate, orthogonally polarized modes supported in the fiber propagate with generally different phase velocities. Therefore, the light launched into the single-mode fiber with linear polarization quickly evolves into a state of arbitrary polarization. To solve this problem, the so-called polarization maintaining fibers were developed. They do not aim at suppressing the birefringence to a maximum degree possible, but on the contrary, they induce a large amount of birefringence within the fiber core, so that the two polarization eigenmodes effectively decouple. The two linear polarization directions in question are often referred to as fast and slow axes of the fiber. These are the only modes to be maintained in the fiber regardless of any external mechanical or thermal perturabations; all the other polarization states gradually change in an uncontrolled way during their transmission through the fiber.

The additional birefringence in the fiber is induced by forming a non-circular fiber core (shaped induced birefringence) or by applying constant stresses within the fiber-core structure (stressed induced birefringence). The latter type is used in today's most popular polarization-maintaining fibers – Panda fibers – which are formed by two stress rods of a modified glass composition on opposite sides of the core. An advantage of Panda fibers is that their fiber core size and numerical aperture is compatible with regular single mode fiber, thereby ensuring minimum losses in devices using both types of fibers. When Panda fibers are connectorized, it is important that the stress rods are parallel or orthogonal with the connector key. The connector key then gives a reference direction which coincides with slow or fast axis of the fiber. The performance of polarization maintaining fibers or other devices using these fibers are usually measured in terms of polarization extinction ratio. This is the ratio (expressed in decibels) of the optical power in the desired transmitted polarization to that in the blocked polarization. Naturally, this ratio is degraded by the angular misalignment between the light polarization axis and the slow/fast axis of the fiber. If the misalignment is given by angle α, then the output polarization extinction ratio is limited to $10\log\left(\tan^2\alpha\right)$.



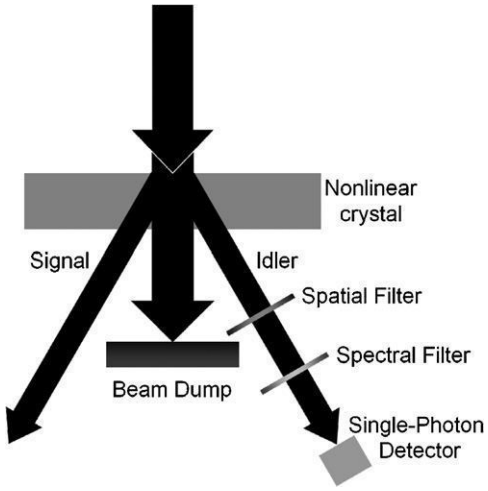


Figure 9. An example of a heralded photon source. After [88].

2.5. Heralded photons

Perhaps one of the largest disadvantages of the single-photon sources described so far is that the emission is random. It is not possible to tell if a particular excitation pulse has generated a single-photon emission until that single photon is detected. It is possible to use parametric down-conversion to generate a 'heralded' source of single photons—'heralded' meaning that there is an indication of when a single photon was generated via detection of its pair photon [87, 88]. As shown in figure 9 a crystal with an optical nonlinearity (often of $\chi^{(2)}$ type) has a pump laser of a short wavelength (and therefore, high frequency) passed through it [88]. A beam dump is used at the far side of the crystal to block any pump laser photons which are transmitted through the crystal. The photon may occasionally annihilate to produce a photon pair at well-defined wavelengths with a total energy (measured as the sum of the frequencies) which is the same as the pump laser photons. These photons are emitted in a cone centred on the pump laser with one photon at a particular wavelength known as the 'signal' and the other the 'idler' [87, 88]. The relationship between the parameters of the signal, idler and pump photons is given by energy and momentum conservation and can be expressed as

$$\begin{aligned} & E_S + E_I = E_P \\ \text{or} \quad & \nu_S + \nu_I = \nu_P \\ \text{or} \quad & \frac{1}{\lambda_S} + \frac{1}{\lambda_I} = \frac{1}{\lambda_P} \end{aligned} \tag{12}$$

where E_S is the energy of the signal photon, E_I the energy of the idler photon and E_P the energy of the pump photon. Similarly, ν denotes frequency and λ wavelength. Since the signal and idler photons are generated as pairs, if spatial and spectral filtering is used on each of the two generated photon streams it is possible to use detection of one (say the 'idler') to confirm, or herald, the existence of the other (the 'signal') [87, 88].

This approach has been used on a number of occasions to generate heralded single photons for several different applications. In 2004, Fasel *et al* reported the generation of 1550 nm wavelength photons (with a 6.9 nm spectral width) heralded by 810 nm photons from a 532 nm wavelength pump source in a type I $KNbO_3$ bulk non-linear crystal [89]. The 1550 nm wavelength photons were heralded with in excess of 60% probability and the probability of multi-photon emission was reduced by a factor of up to 500 when compared to Poissonian light. Castelletto *et al* reported the highly efficient generation of heralded single photons at a wavelength of 1550 nm and the collection of these into single spatial and spectral modes [90]. In 2007, Soujaeff *et al* published details of a heralded single-photon source producing signal photons at a wavelength of 1550 nm and herald photons at 521 nm [91]. At an input laser pulse repetition frequency of 82 MHz, the experiment produced single photons at a mean rate of 216 kHz.

3. Single-photon detection

3.1. Characterization parameters

An important parameter in the choice of detector is the detection efficiency (η_D) which is the probability that an incident photon generates a measurable current pulse, assuming that the time between photon arrivals is greater than the dead-time of the detection system. The quantum efficiency (η_q) is the probability that an incident photon generates an electron–hole pair and is related to η_D by a factor which depends on the detector. Figure 10 shows a comparison of the detection efficiency spectra for three detectors of types which will be described in more detail later.

In photon-counting detectors, events will occur in the absence of incident radiation, or 'dark events'. The origin of these dark counts will vary depending on the detector types and operating conditions, for example in a photomultiplier thermionic emission at the photocathode will give rise to a quantifiable dark count rate. In photon-counting photomultipliers and avalanche photodiodes, the phenomenon of afterpulsing also occurs to form a contribution to the dark count rate. This happens when a detector event generated at a time t generates a subsequent event at a time $t + \Delta t$, where Δt is a largely unpredictable unique time difference for each photo-generated pulse. In addition there are further 'dark' (not generated by a photon) events, the exact cause of which

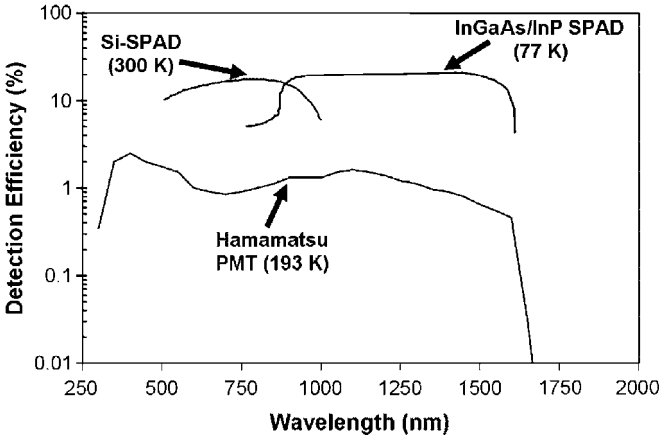


Figure 10. A comparison of the typical detection efficiencies of three different example detector technologies: a photo-multiplier tube (PMT) from Hamamatsu cooled to 193 K; an indium-gallium-arsenide/indium-phosphide (InGaAs/InP) single-photon avalanche diode (SPAD) cooled to 77 K [143] and a silicon (Si) SPAD at room temperature (300 K).

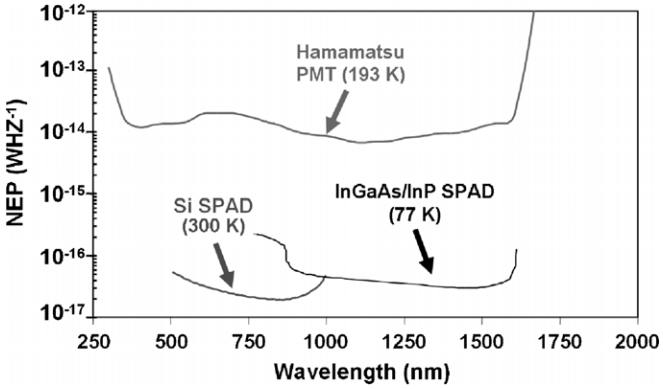


Figure 11. A comparison of the typical noise equivalent power (NEP) of three different detector technologies: a photo-multiplier tube (PMT) from, Hamamatsu cooled to 193 K; an indium gallium arsenide/indium-phosphide (InGaAs/InP) single-photon avalanche diode (SPAD) cooled to 77 K [143] and a silicon (Si) SPAD at room temperature (300 K).

depends on the configuration of the detector used but are often triggered by thermal excitation.

In many single-photon experiments, there are advantages in maintaining a low dark count rate as non-photon generated events will contribute to the overall error rate in measurements. A useful quantity in the characterization of single-photon detectors is the noise equivalent power (NEP), which is defined as

$$\mathrm{NEP} = \frac{hc}{\lambda \eta_q}\sqrt{2N_D} \tag{13}$$

where λ is the wavelength of the incident photon, N_D is the dark count rate and η_q is the quantum efficiency, as defined before. The NEP is defined as the signal power required to attain a unity signal-to-noise ratio within a one second integration time [92] and is shown in figure 11 for the same three detectors presented in figure 10.

For many applications, the timing jitter of the detector is an important characterization parameter. The timing jitter is a measure of the degree of variation in the delay between the arrival time of a photon at the detector and the output of an electrical pulse. It is common to take many different measurements of this delay and plot them in a histogram. The timing jitter is then quoted as being the full-width at half-maximum of the peak in this histogram. In time-of-flight laser-ranging, the timing jitter of the detector will affect the depth resolution of the measurement [93] while in QKD, the timing jitter can lead to intersymbol interference and increase the error rate of the key exchange process [94].

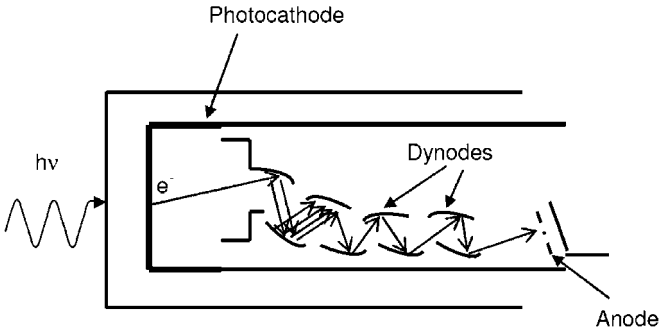


Figure 12. Schematic of a 'focused dynode chain' photomultiplier tube (PMT); the process of electron multiplication from each dynode stage causes a measurable current to be detected [72].

3.2. Photomultiplier tubes

Figure 12 shows a representational diagram of the operation of possibly the earliest form of single-photon detector to be used in a wide-range of applications; the photomultiplier tube (PMT) [95–98]. The device is formed from a vacuum tube with a photocathode at the entrance and a series of following dynodes arranged prior to an anode at the far end. The photocathode absorbs the incident photons and emits up to one electron per incident photon as it does so. Depending on the material composition of the photocathode, PMTs can be effective for detection of light at varying wavelengths. Generally, the photocathode consists of a thin evaporated layer of alkali metal compounds and one or more group V elements. Once an electron is emitted from the photocathode, it is accelerated towards, and collides with, the first positively charged dynode. This collision releases further electrons which are then accelerated towards the second dynode. Each successive dynode in the PMT is charged to a higher positive potential than the preceding one resulting in amplification as the increasing number of electrons collide with later dynodes. When compared to the semiconductor detectors which will be described later, PMTs have a high internal gain. If N dynodes are used, each with secondary electron coefficient α, then the multiplication factor, M, is given by

$$M = \alpha^N \tag{14}$$

Typical values of $\alpha = 4$ and $N = 10$ lead to a multiplication factor of 10^6 [99]. Typically, PMTs have multiplication factors which lie in the range 10^4–10^7.

When compared to the smaller active areas of semiconductor single-photon detectors, the large

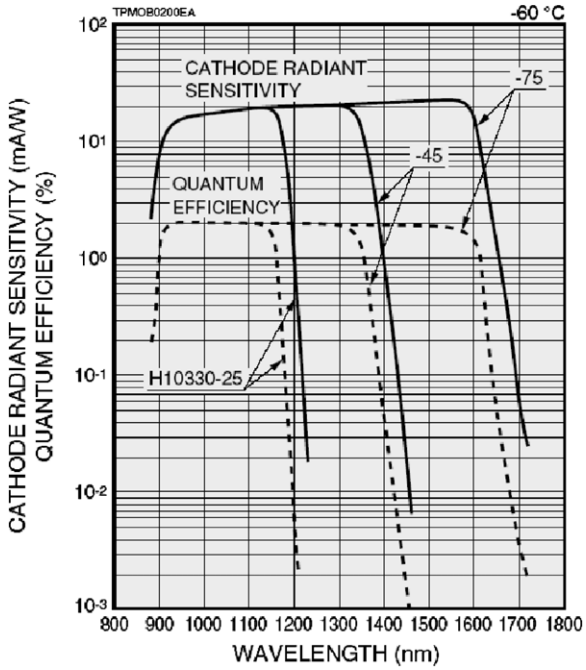


Figure 13. Quantum efficiency of three Hamamatsu PMT devices at −60 °C [75].

photocathodes used in PMTs (which can be up to 1000 times larger) are an advantage—facilitating improved optical collection, especially from extended optical sources [100]. However, the requirement to have several dynodes means that PMTs are contained in relatively large physical packages (although there have been advances in miniaturization [100]), have a poor mechanical stability and require high bias voltages, typically of the order of 1 kV. Although the single-photon detection efficiencies of PMTs are generally not greatly dependent on the wavelength of the incident photons, they typically exhibit low single-photon detection efficiencies [101, 102] and their timing jitter (which is mainly dependent on the fluctuations in transit times from the photocathode to the anode) is typically in the region of 1 ns [103]. The combination of these factors means that in applications where higher detection efficiencies and lower jitter are required, such as time-of-flight laser ranging [93, 104], PMTs are less likely to be used. The afterpulsing in PMTs is most likely caused by ion feedback or by luminescence of the dynode material and the glass of the tube while the dark counts are typically induced by thermal excitation [80].

If the photocathode is fabricated from InGaAs, and the device cooled to 200 K, it is possible to extend the spectrum of operation for a PMT into the near-infrared wavelengths up to 1700 nm [105]. The cooling is required to reduce the dark count rate which is typically around 10^5 counts per second at this temperature. Figure 13 shows the quantum efficiency of three devices manufactured by Hamamatsu; at best their efficiency is ∼2%. The cathode radiant sensitivity curve shows the ratio of the current transmitted by the photocathode to the incident radiant power of a specific wavelength. The quantum

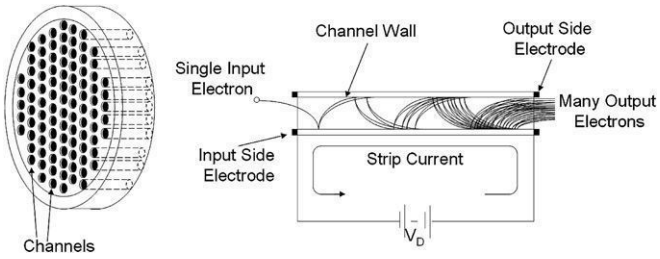


Figure 14. Micro-channel-plate (MCP) schematic and operating mechanism [106].

efficiencies of the three detectors follow different gradients as they are related by

$$\eta_q = S\frac{h\nu}{e} \tag{15}$$

where S is the cathode radiant sensitivity, ν is the frequency of the light incident upon the detector and e is charge on an electron.

3.3. Microchannel plates

It is possible to consider microchannel plates (MCP) as an array of devices which operate along principles similar to those employed in the PMT. A schematic of the structure of a typical MCP is shown in figure 14, along with the method of operation. An MCP is assembled from an array of multiple ∼10 μm diameter thin glass capillaries arranged into a disc. Unlike in the case of the PMT, where there were separate dynodes positioned along the length of the tube, in an MCP the inner wall of each capillary is coated with a photo-emissive material and biased at each end, so that it acts as a continuous dynode [106]. The single input electron is accelerated by the positive potential towards the inner wall of the channel where it collides, releasing secondary electrons which then initiate further collisions along the length of the channel resulting in an exponential multiplication of the electron flux. MCP detectors require a high voltage of the order of 1 to 3 kV to operate but do not show the same multiplication factors as PMT detectors. Experiments by Becker indicated that some MCP devices do not exhibit any afterpulsing effects at timescales of up to 150 ns [80]. Typically, MCPs will exhibit timing jitters of approximately 20 to 30 ps [107].

The large 2D arrayed input means that the MCPs are good candidates to be used as detectors in imaging systems such as image intensifiers or time-of-flight laser-ranging systems [108].

3.4. Avalanche photodiodes

During the last 30 or so years, more rugged and efficient alternatives to dynode-based detectors have become more widely used: semiconductor detectors known as avalanche photodiodes. A diode is a device formed by a junction between a semiconductor with an excess of holes (p-type) and a semiconductor with an excess of carriers (n-type). These n-type and p-type regions are formed by the introduction of immobile dopant centres of opposite polarity within

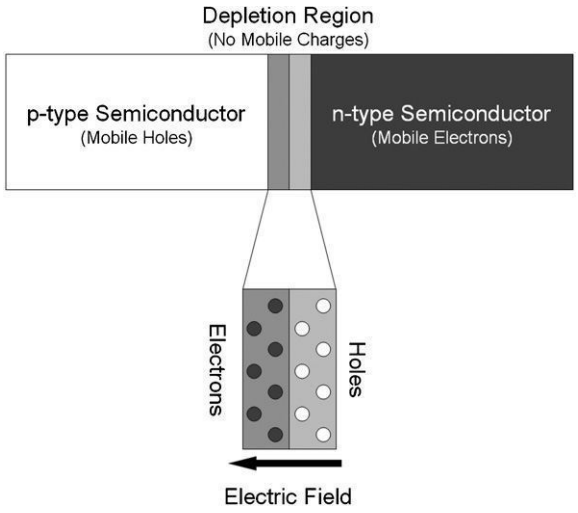


Figure 15. A schematic diagram of the structure of a diode. At the junction between the n-type and p-type semiconductor the mobile charges diffuse and create an electric field which prevents further charge diffusion. After [27].

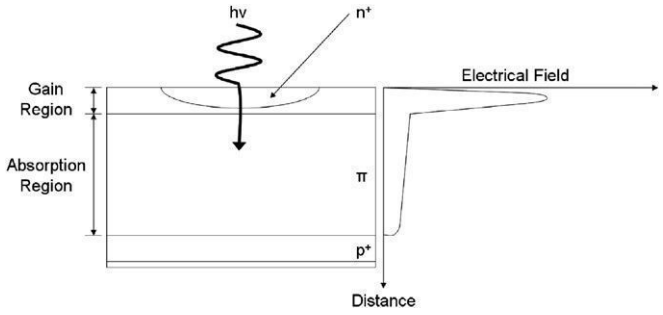


Figure 16. The typical structure of a reach through avalanche photodiode (APD) is shown on the left while the right shows the electrical field within the device. After [100].

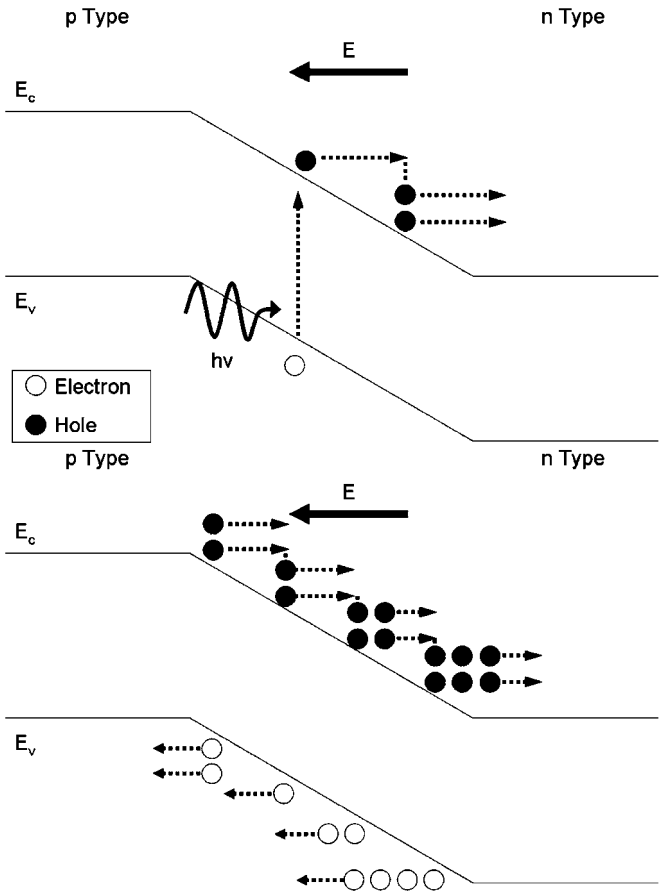


Figure 17. The process of impact ionization in an avalanche photodiode (APD). In the top schematic an electron–hole pair is generated by the absorption of a photon in the depletion region of a reverse biased APD—a process known as impact ionization. In the lower schematic the electron causes further impact ionizations, multiplying the number of electron–hole pairs and producing a self-sustaining avalanche of carriers.

the material. Because of the diffusion of carriers set up by the discontinuity between the two differently doped semiconductors, a region of electric field is formed which is depleted of all free carriers (as shown in figure 15) [100]. When a voltage is applied so that the n-type semiconductor is at a higher potential than the p-type, the junction is said to be reverse biased. When the electric field is sufficiently high, then a drifting electron can gain enough kinetic energy that, on collision with an atom in the lattice, it can knock an electron out of its bound state and promote it into the conduction band, releasing a hole in the valence band. This phenomenon is known as impact ionization, and a similar effect is seen for holes drifting in an electric field. These resulting carriers will then be accelerated in the high electric field and may undergo further impact ionization events, resulting in an avalanche of carriers—hence the term avalanche photodiode or APD.

Figure 16 shows a schematic of an example of a silicon reach-through avalanche photodiode and the corresponding electric field. The absorption region (in which the energy of the photon is absorbed) is produced from a thick layer of an undoped or lightly doped (intrinsic) semiconductor which is introduced between the p and n layers. This thick layer is introduced to ensure a high level of absorptance of incident photons within the depleted region, as well as ensuring predominantly electron injection into the high-field multiplication region, the electron impact ionization coefficient in silicon being higher than the hole impact ionization coefficient. The use of this thick layer also helps reduce the capacitance of the device. In the gain region (also known as the multiplication region) the injected electron drifts from the point of absorption and undergoes impact ionization, initiating the avalanche process. This secondary electron–hole pair can lead to further impact ionization which will generate another electron–hole pair in a continuing process, as shown in figure 17 [100].

If the electric field applied across the device is sufficiently high to be above the avalanche breakdown threshold, it is possible for a single photo-generated carrier to induce a self-sustaining avalanche where the positive feedback from hole and electron impact ionization means that the avalanche process cannot stop. The point at which this occurs can be easily seen from the current–voltage characteristics of the device, as shown in figure 18. As the reverse bias voltage is increased, the current increases steadily until the voltage reaches the breakdown voltage where the current increases

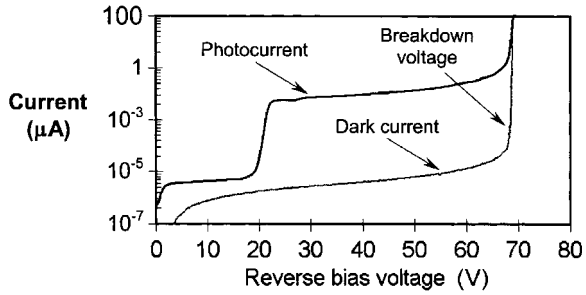


Figure 18. An example graph showing the change in photocurrent and dark current of an avalanche photodiode (APD) as the reverse bias voltage is increased towards breakdown voltage. This graph is for an InGaAs device manufactured by Fujitsu for use at a wavelength of 1300 nm.

rapidly, indicating that the carriers are beginning to multiply; however, at such a bias level the photocurrent will be linearly proportional to the incident light level. Above avalanche breakdown, a single electron (or hole) can initiate a self-sustaining current, the onset of which will be readily detectable by external thresholding circuitry. Such a mode of operation can be regarded as purely digital: the output is the same whether one or more photons are incident at the same time, quite different to the linear multiplication region of operation. When operated in this photon-counting, or Geiger, mode, the device is termed a single-photon avalanche diode or SPAD [109]. In order to stop the self-sustaining avalanche, it is necessary to reset the detector so that it is ready to receive more photons, i.e. the process of 'quenching' the avalanche.

The single-photon detection efficiency increases with the increasing excess bias voltage above avalanche breakdown primarily due the increase in the avalanche triggering probability, and possibly also due to the increase in the thickness of the depletion region [110]. The dark count rate also increases with increasing excess bias voltage due to the increase of both the avalanche triggering probability and the field-enhanced dark count generation [111, 112]. The rate of increase of dark count rate is usually greater than the rate of increase in photon detection efficiency. In addition, as the excess bias voltage is increased, the FWHM timing jitter decreases [109]. Consequently, the choice of optimal excess bias depends on the application to which the detector is applied. If the contribution of counts from other light sources is likely to greatly exceed the dark count rate (such as the solar background in the time-of-flight ranging application) then it may be preferable to increase the photon detection efficiency at the cost of increasing the dark count rate. In the case of quantum key distribution, measurements over long distances (i.e. high transmission channel losses) will be dominated by the dark count rate and increased timing jitter can lead to intersymbol interference. It is therefore necessary to balance the choice of excess bias to ensure sufficient detection efficiency, low timing jitter in comparison to the clock period and a low dark count rate.

3.4.1. Quenching. The process of quenching involves detecting the leading edge of the avalanche current, then generating a closely time-correlated electrical pulse, reducing the bias voltage below breakdown level and finally restoring the voltage to the original operating level. There are three main forms of quenching applied to detectors: passive, active and gated.

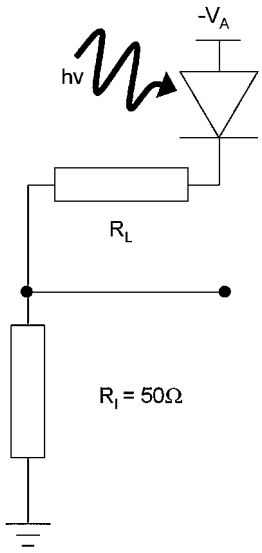


Figure 19. An example of a circuit used in passive quenching. When the avalanche photodiode (APD) is biased above the breakdown voltage by V_A an electron–hole pair can generate a self-sustaining avalanche. The avalanche discharges through the high resistance of R_L and the voltage V_A decreases. This is sensed by a comparator which produces a signal for photon counting and timing [113].

A passive quenching circuit is simply a high impedance load connected in series to the SPAD [109, 113]. As shown in figure 19, the SPAD is reverse biased through a high ballast resistor R_L, of the order of kΩ. During a self-sustaining avalanche, the resistance of the SPAD drops to a few kΩ and the most of the external bias is dropped across R_1. Passive quenching circuits can have a slow recovery time in which no further photon detection events can be registered (a deadtime), reducing the maximum count rate possible. The recovery time depends on the product of the resistor R_L and the internal capacitance of the diode. Although it is possible to reduce both values to minimize the recovery time, even at small values of R_L (∼500 kΩ) and internal capacitance (∼1 pF), the recovery time can be up to ∼1 μs [109].

Active quenching was first developed by a group at Politecnico di Milano [114] and the basic circuit is shown figure 20 [115]. The photodiode is biased from the low impedance (100 Ω) source. The avalanche current induced by a photon event triggers a fast comparator which switches the current in an emitter-coupled transistor pair. A negative pulse is then superimposed on the bias and rapidly quenches the avalanche. Due to the additional propagation delays caused by the path through the comparator, the emitter-coupled transistor pair and the leads of the feedback loop, the leading edge of this pulse is delayed with respect to the onset of the avalanche and the duration of the avalanche current pulse is set by this delay. The hold-off time after quenching is determined by the duration of the comparator output pulse and is equal to that of the avalanche pulse. The deadtime associated with this

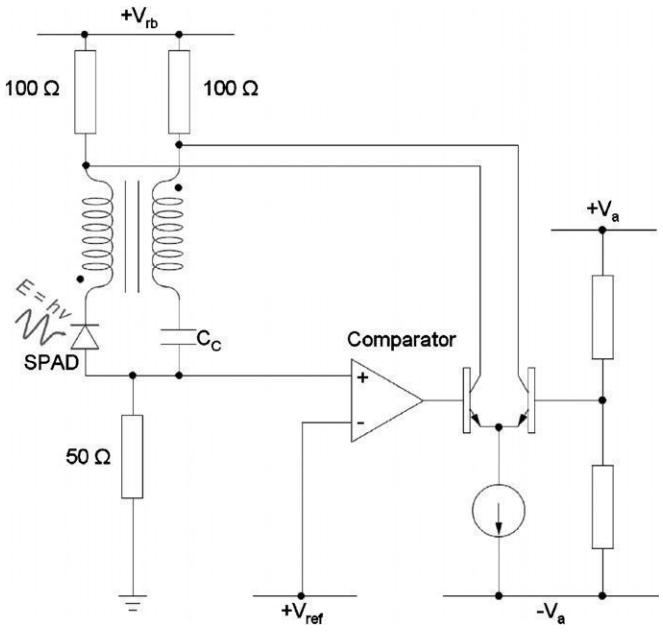


Figure 20. An example of a circuit used to perform active quenching [115].

technique is the sum of the avalanche and hold-off durations (approximately twice the duration of the propagation delay in the feedback loop) and can be of the order of a few nanoseconds—permitting photon detection rates of the order of Mbit s^{-1}.

It is important to note that the quenching voltage pulse induces a capacitive current flow in the diode which can be comparable to the avalanche current. This means that the comparator would be retriggered on the trailing edge of the pulse and the circuit would begin to oscillate. To alleviate this, the capacitive current peak is cancelled by a current pulse which has the same shape but opposite polarity, obtained by coupling the complementary output to the comparator unit via an adjustable compensating capacitor (C_C in figure 20). A transformer is used to equalize the shapes of the complementary pulses. This results in a slightly increased deadtime due to the degradation of the trailing edge of the pulse at the comparator input and the increased delay of the feedback loop.

In gated quenching, the bias voltage is only increased above the breakdown level for the duration of the period in which a photon is expected. A form of gated quenching is used for most InGaAs/InP SPADs. This approach is applicable to QKD as the arrival times of the photons at the detectors, corresponding to the individual binary digits (bits) sent by Alice, can be predicted with reasonably high accuracy. However, in other applications such as time-of-flight-ranging, this technique is less appropriate since the arrival time of the photons at the detector is dependent on the unknown distance to the target.

3.4.2. Silicon single-photon avalanche diodes. Silicon single-photon avalanche diode detectors have been used for several decades [114, 115], and have become widely used in a number of photon-counting application areas in the spectral region 400–1000 nm.

There are two main types of design of Si-SPAD architecture: thick and thin junction. Both thick [116] and thin [117] junction Si-SPADs are now commercially available technologies. As may be guessed from the different terms, the main difference between the two designs is the thickness of the depleted region in which photon absorption takes place (see figure 21) [118]. In the case of a thick junction SPAD [119] this can be a few tens of μm whereas in the case of the thin junction it is typically only a few μm [118].

Generally, the thin-junction devices have lower single-photon detection efficiency (SPDE) than the thick junction, and the long interaction length of the latter leads to improved efficiency in the near-infrared. Figure 23 shows a comparison between the detection efficiencies of typical thick and thin junction Si-SPADs [109]. It can be seen from figure 22 that a typical thin junction Si-SPAD has a peak detection efficiency of ∼52% at a wavelength of ∼525 nm, falling to ∼3% at a wavelength of 1000 nm. A typical thick junction Si-SPAD has a peak detection efficiency of ∼70% at a wavelength of ∼800 nm, falling to ∼5% at 1000 nm.

In terms of timing jitter, there have been reports of thin junction devices having exhibited jitter responses of 20 ps FWHM, although these were in small area diameter devices (∼10 μm) [120]. Thick junction devices generally exhibit timing jitter of the order of 520 ps FWHM in a 200 μm diameter active area [121]. It should be noted that some authors have reported much improved timing jitter through adaptations of read-out electronics [121].

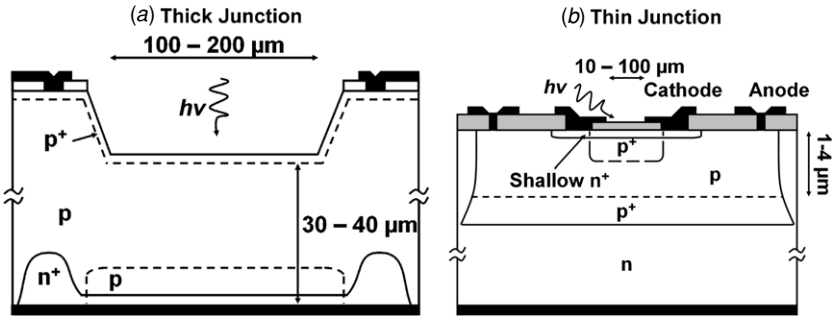


Figure 21. Left: a schematic of a thick junction silicon single-photon avalanche diode (Si-SPAD). Right: a schematic of a thin junction Si-SPAD with [118].

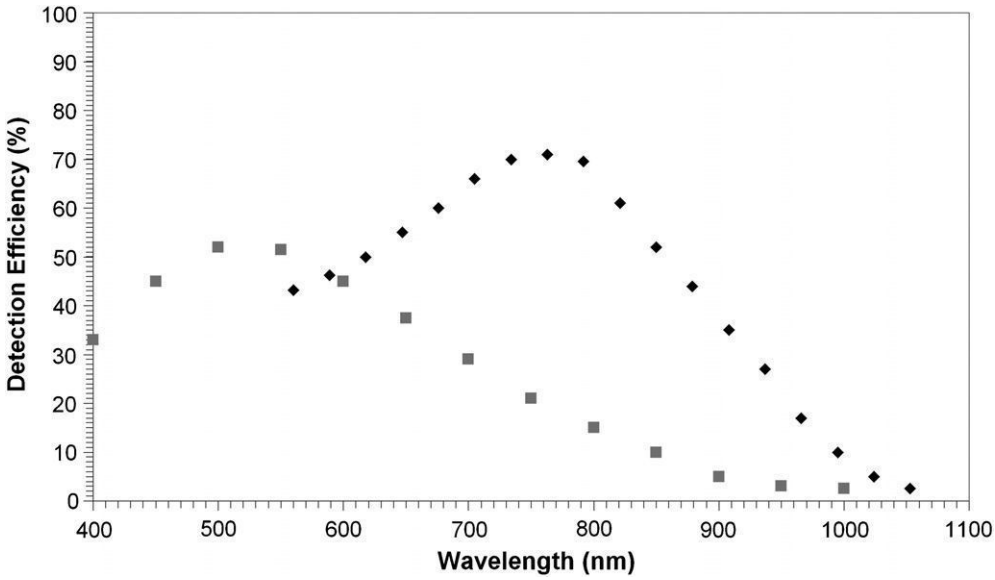


Figure 22. A comparison of the detection efficiencies of thick (diamonds, ♦) and thin junction (squares, ■) silicon single-photon avalanche diodes (Si-SPADs) [109].

The reduction in the thickness of the junction leads to a reduction in the timing jitter but also a reduction in the detection efficiency. Early designs of thin junction SPAD exhibited long diffusion tails [122–124] whereby the instrumental response has a peak followed by a long, exponential tail caused by photogenerated carriers which are generated by absorption in the substrate of the device which then reach the depletion layer by diffusion and generate further avalanches. At $\frac{1}{100}$th of the maximum, the full-width (FW1/100M) of the detector response could be as great as 650 ps for comparatively short FWHM of ∼35 ps [125]. Later version of these devices improved the FW1/100M to 130 ps by reducing the thickness of the neutral layer beneath the active junction [125].

Generally, the dark count rate of a thin junction Si-SPAD depends on the diameter of the active area [125]. Typical values at 20 °C are ∼700 count s^{-1} for 50 μm, diameter ∼3000 count s^{-1} for 100 μm diameter and ∼ 40 000 count s^{-1} for 200 μm diameter [125]. As the device is cooled, the probability of thermally generated carriers falls and the dark count rate falls. At a temperature of −25 °C the dark count rates become 5 count s^{-1}, 50 count s^{-1} and 1500 count s^{-1} respectively [125]. At a wavelength of 850 nm a typical commercial thick junction Si-SPAD at a temperature of ∼ −25 °C will exhibit an NEP of ∼8×10^{-18} W $Hz^{-1/2}$ while a 200 μm diameter active area thin junction at the same wavelength and a temperature of −25 °C will exhibit an NEP of ∼7×10^{-17} W $Hz^{-1/2}$ [125], and at room temperature the higher NEP of ∼7×10^{-16} W $Hz^{-1/2}$. However, it should be noted that reduced NEP is readily observed with smaller diameter devices, for example an NEP of 1.5×10^{-17} W $Hz^{-1/2}$ was measured using a 50 μm diameter shallow-junction Si-SPAD at −15 °C at a wavelength of 850 nm [126]. At room temperature a 50 μm diameter shallow junction Si-SPAD has demonstrated an NEP of 4.7×10^{-17} W $Hz^{-1/2}$ [127] at a wavelength of 850 nm.

Just as in the case of some single-photon sources, it is possible to use a resonant cavity to increase the efficiency of SPAD detectors [128]. It is possible to form a cavity of this type by using a reflector buried in the device and the air/semiconductor interface at the top [128, 129], leading to a higher detection efficiency for the same depletion region thickness. This can be used with thin junction SPADs to increase the detection efficiency while avoiding an increase in timing jitter [118]. The introduction of a cavity into a thin junction Si-SPAD using silicon-on-insulator (SOI) in the lower mirror was shown by Ghioni *et al* [125] to increase the detection efficiency from 10% to 34% at a wavelength of 850 nm. This prototype device has a high room temperature dark count rate of 100 000 count s^{-1} for a 50 μm diameter active area, due to dislocations induced by the SOI layers by temperature-induced strain relaxation.

As was discussed previously, the main advantage of PMT detectors is the large active area when compared to Si-SPADs. One possible solution is to produce a 'large area' detector which is a series of interconnected Si-SPADs in a grid formation. Each element in the grid (or pixel) is a SPAD. The SPADs are joined together on a common substrate and output across a common load resistor so that the electrical outputs of each pixel are summed [130]. Consequently, it is possible to use these interconnected Si-SPADs in photon number resolving experiments as, provided the photons are incident on different pixels, coincident photon events will generate an electrical output pulse which is correspondingly higher by a factor depending on the number of pixels which were illuminated. All of the electrical outputs from the individual pixels are combined, adding dark count contributions from each pixel to the total dark count rate. A typical fill factor for a 1 mm^2 SiPM is about 25% [130].

The single-photon detection efficiencies of interconnected Si-SPADs are, as would be expected, similar to those of single silicon SPADs. An example interconnected Si-SPAD matrix,

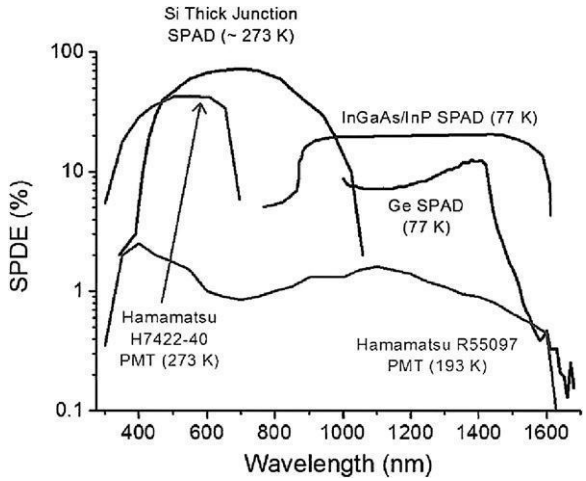


Figure 23. Plot of single-photon detection efficiencies versus wavelength for a silicon thick junction single-photon avalanche diode (SPAD), an InGaAs/InP avalanche photodiode (APD) operated in Geiger mode, a germanium SPAD, a Hamamatsu Infrared Photomultiplier Tube (PMT), and a visible wavelength high-efficiency Hamamatsu PMT.

such as that developed by SensL in 2008, has a single-photon detection efficiency of the order of 20% at a wavelength of 500 nm, decreasing to approximately 6% at 800 nm [131, 132]. The NEP of a 1 mm^2 SensL detector with 620 elements is $\sim 9 \times 10^{-15}$ W $Hz^{-1/2}$ at a temperature of -4 °C [133].

Recent advances in arrayed silicon SPADs [134–136] also provide possible options as highly sensitive focal plane arrays, and, in some cases, these arrays permit individual timing information from each pixel. Arrayed Si-SPADs consist of a 2D matrix of independently electrically addressed SPAD detectors covering a large area, typically several mm^2. The addressing and read-out electronics tend to take a high proportion of the area between pixels, leading to low fill factors. Arrayed silicon SPADs have found applications in fields such as 3D imaging [135] and astronomy [134] where they have the potential to reduce measurement durations. Niclass *et al* used a 32 × 32 array of SPADs to demonstrate a 3D imager based on time-of-flight [135], with each pixel exhibiting 115 ps timing jitter (FWHM). A separate depth measurement was performed for each pixel in the array and the 1024 independent measurements combined to produce a 3D image of the target. The NEP of a typical pixel was $\sim 6 \times 10^{-17}$ W $Hz^{-1/2}$ at a temperature of 0 °C. Zappa *et al* presented a 60 pixel array for use in astronomical applications where the photon numbers involved in measurements can be extremely low but there is a desire to undertake measurements for only a short time period due to the rapid nature of the phenomena under examination [134]. The maximum saturated count rate of this detector was 30 Mcount s^{-1} while the minimum was at the single-photon level. The NEP of a single pixel in this array was $\sim 3 \times 10^{-17}$ W $Hz^{-1/2}$ at a temperature of -10 °C.

3.4.3. Germanium single-photon avalanche diodes. In the mid-1990s, commercially available linear multiplication germanium APDs were characterized in Geiger mode [136, 137], as potential photon-counting detectors in the infrared. Their quantum efficiency performance compared to that of other SPADs and two PMTs is displayed in figure 23. The main problem with Ge SPADs was the high DCR which was reduced by cryogenic cooling. At room temperature, Ge would absorb at wavelengths beyond 1550 nm, but at the temperatures the devices were tested (77 K), the cut-off was a little under 1500 nm. The SPDE of a Ge SPAD was measured in gated mode to be a modest ∼10% and sub-100 ps timing jitter. Ge devices also suffered from a high level of afterpulsing, a phenomenon also apparent in the InGaAs/InP devices which will be discussed later.

3.4.4. Silicon germanium single-photon avalanche diodes. One step towards low-noise linear multiplication APDs was the use of separate absorption and multiplication hetero-structures. Whilst most progress was made in III–V structures, e.g. InGaAs/InP as described below, some progress was made in the use of Ge-containing absorbing layers on Si multiplication layers. The 1980s saw progress towards the development of strained layer silicon/silicon germanium (Si/SiGe) absorbing avalanche diodes grown on Si [138, 139] and used in linear multiplication mode. The devices were fabricated with the absorption in the $Si/Si_{1-x}Ge_x$ alloy layers and multiplication in the Si layers. In 2002, the first attempt at a SPAD grown using a similar approach was made by Loudon *et al* [140], where clear improvements in quantum efficiency were found in the near infrared over otherwise identical all-Si control samples. However, such devices used strained SiGe/Si layers with $Si/Si_{0.7}Ge_{0.3}$ multiple quantum well material as an absorber, where the thickness of the Ge-containing layer was kept low—a total of only 300 nm including the all-Si layers—in order to keep below the critical thickness of the layer and inhibit relaxation. Such thin layers meant that the overall absorptance in the infrared was low. In later work by Carroll *et al* [141], a 400 nm thick Ge absorber was grown on Si to demonstrate linear multiplication and dark counting above avalanche breakdown.

3.4.5. Indium-phosphide-based single-photon avalanche diodes. Currently, the most promising candidates for near-room temperature single-photon counting at a wavelength of 1550 nm are indium-phosphide (InP)-based separate absorption and multiplication avalanche diodes, particularly InGaAs/InP devices [142, 143]. These devices have been used as linear multiplication devices [144] for many years and have been more recently used above avalanche breakdown, in Geiger mode. In the last few years, specific growth and fabrication programmes aimed at designs for single-photon operation have yielded devices with improved performance [145, 146].

Figure 24 illustrates a schematic of a planar geometry InGaAs/InP SPAD detector, where incident infrared photons are absorbed in the narrow-gap InGaAs and, the photo-generated holes drift to the high-field InP, where multiplication takes place. One potentially important issue in such devices is a large valence band discontinuity in these devices which can result in the delay or recombination of devices at the interface. To combat this, a thin layer (typically 100 nm) of

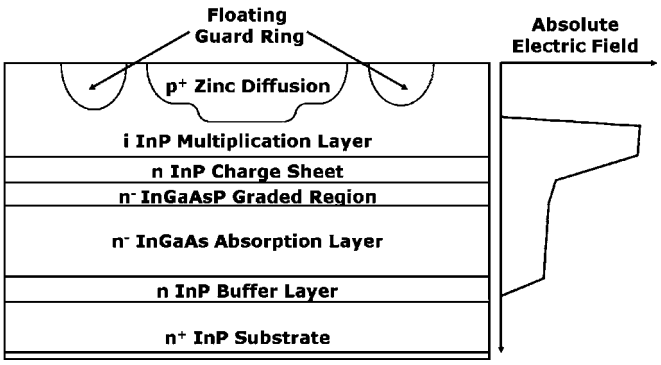


Figure 24. Layer structure of planar geometry InGaAs/InP single-photon avalanche diode. The junction is formed by the diffusion of zinc (Zn) into n-type InP and typically is fabricated with multiple diffusion steps and guard rings in order to avoid preferential edge breakdown.

intermediate gap InGaAsP grading layer is grown between the InP and InGaAs layers to provide a graded step for the drifting photogenerated holes to traverse.

Initially, most work using InGaAs/InP SPADs has used an elementary form of gated quenching operation, typically at low temperatures to counteract the effects of inherently high dark count rates. For example, Hiskett *et al* [143] reported an NEP of 1×10^{-16} W Hz$^{-1/2}$ at 77 K with a Fujitsu linear multiplication device operated in Geiger mode in 1998. By 2001, the same group reported an NEP of 4×10^{-17} W Hz$^{-1/2}$ at the same temperature with an Epitaxx linear multiplication avalanche photodiode. At all temperatures, sub-nanosecond jitter measurements were reported by a number of groups. Aside from the practicality of low temperature operation, serious performance issues also occur as a result of the longer trap lifetimes. Afterpulsing in such a SPAD occurs where an avalanche current fills mid-band gap trap states in the material which then emit carriers at a later time causing further 'dark' (i.e. not directly induced by a photon) events. The resulting higher afterpulsing rates can only be reduced by having a bias below the breakdown threshold after each event, in order that the traps can be emptied without resulting in an avalanche pulse. Consequently, only low gating rates are possible to avoid afterpulsing, where the maximum rate depends on the temperature, the gate duration as well as the constituent material properties, but will generally be prohibitively low, typically in the 1–100 kHz range. Most successful approaches for the reduction of this phenomenon have relied on improved quenching methods to reduce the charge flow per event. Improved quenching approaches, such as very rapid gating at near 1 GHz or greater [147, 148] have been used to reduce the effect of afterpulsing at 223 K temperature operation. Yuan *et al* [149] used a self-differencing circuit to reduce the charge required for the recording of each photon event resulting in reduced afterpulsing and operation at gigahertz clock rates.

These gigahertz gating approaches can work well at near room temperature operation for the application of quantum key distribution where the photon events occur in pre-defined time windows. However, some applications, such as photon-counting time-of-flight ranging [17] or time-resolved luminescence ideally require an ungated detection technique for more efficient data acquisition. Ungated operation has been realized using rapid active quenching, to permit free-running operation at 210 K [150]. Recent results with low-bias passive quenching have shown room temperature operation with no electrical gating, in a completely free-running mode, although the sensitivity was $\sim 1 \times 10^{-14}$ W Hz$^{-1/2}$ [151]. Greater understanding of the principal dark count mechanisms, eg field-assisted tunnelling in the InP, has led to further improvements in sensitivity via the introduction of longer InP multiplication region [152, 153]. Use of these devices in the low-bias regime have led to room temperature operation with no electrical gating with NEP of $\sim 1 \times 10^{-15}$ W Hz$^{-1/2}$, reducing to 5×10^{-17} W Hz$^{-1/2}$ at 210 K [154], with 10^6 count s per second operation demonstrated.

Other methods of reducing the single-photon induced avalanche pulse have been used where the feedback layers have been incorporated within the device structure. An early example of this uses InGaAs absorber layers and InAlAs multiplication grown on InP [155], and feedback provided by the avalanche pulse altering a hetero-barrier height within the structure, permitting self-quenching and self-recovery.

Despite a number of issues regarding afterpulsing and dark count rates, InP-based SPADs remain the outstanding candidate for practical single-photon detection at 1550 nm wavelength. Although significant improvements have been reported in terms of quenching approach and structure design, the major issue of the origin of afterpulsing phenomenon remains. Although several groups have reported evidence that the traps responsible for afterpulsing are found in the InP layers [156, 157], serious attempts at the removal of the defect complexes have yet to begin. A concentration of research in this area is likely to yield further improvements, leading to reduced dark count rates and higher photon counting rates.

1.1 Basic Principles of Operation

To generate entangled photon pairs, a second-order nonlinear process, usually referred to as spontaneous parametric down-conversion (SPDC), is used in the quED. In the SPDC process photons of an intense laser pump beam spontaneously convert in a nonlinear crystal with a very low probability ($\approx 10^{-11}$ for standard materials) into pairs of lower-frequency photons. Due

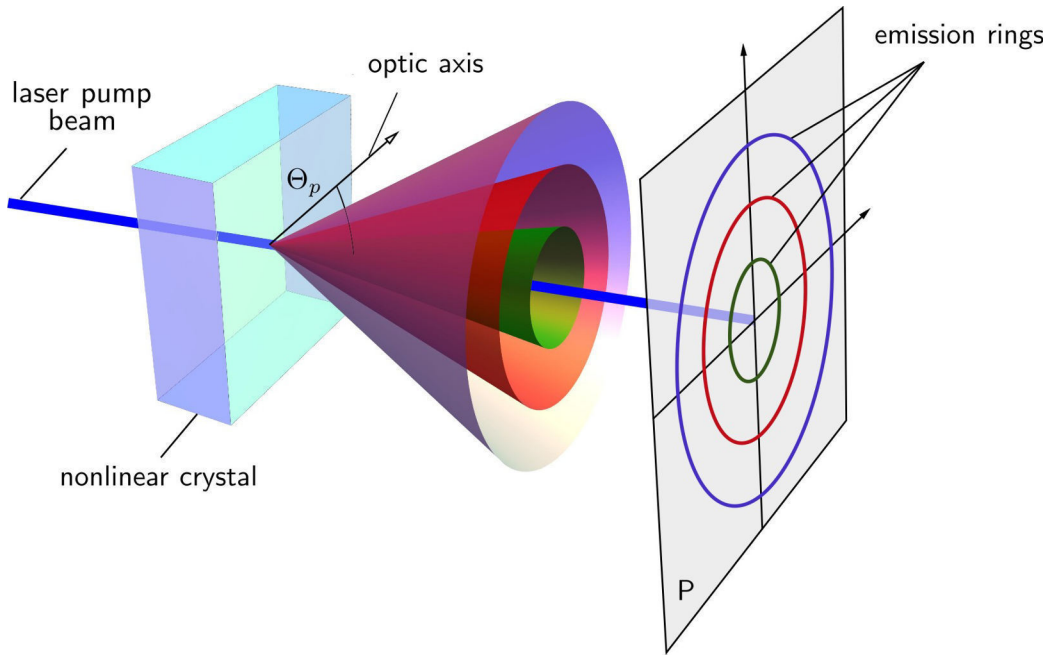


Figure 1: Spatial distribution of the down-conversion emission for type I phase matching. The transverse momentum conservation requires that down-conversion photons have to emerge from the crystal along the directions lying always on exactly opposite sides of the cone.

to energy and momentum conservation in the nonlinear interaction the possible wavelengths and emission directions of the generated photons are severely constrained. Consequently, the emission pattern is formed by cones, which imprint the characteristic rings in the plane (P) perpendicular to the pump-beam direction. In type I phase matching the cones are concentric around the pump direction, as illustrated in Fig. 1. Every cone corresponds to a distinct emitted wavelength. The opening angles of the emission cones thus depend on the wavelengths of the emitted photons, but also on the angle Θ_p between the pump direction and the optical axis; (the smaller the angle Θ_p, the smaller the opening of the cone with a given wavelength). This allows to angle tune the spatial emission of down-conversion photons as required.

To obtain polarization entanglement from SPDC, the quED utilizes a well-known method of coherent spatial overlap of the emissions from two adjacent type-I crystals. Consider two nonlinear crystals, both operated in type-I phase-matching configuration and pumped with linearly polarized light.The otherwise identical crystals are oriented such that their optical axes lie in mutually perpendicular planes. For example, let the optical axis of the first crystal be aligned in the vertical plane and the axis of the second crystal in the horizontal plane. Due to the type-I coupling, the down-conversion process occurs only in the crystal, where the pump photon is extraordinary polarized, emitting two ordinary polarized down-conversion photons into the characteristic cone. That is, with the vertically-polarized pump the down-conversion process occurs only in the first crystal emitting pairs of horizontally polarized photons, whereas with the horizontally-polarized pump it occurs only in the second crystal producing two vertically-

polarized photons. By pumping the crystals with light, linearly polarized at 45° with regard to horizontal and vertical direction, there is an equal probability that a pump photon will be down-converted in either crystal. Provided that the two emission processes are coherent with one another, which is fulfilled as long as there is no way of ascertaining whether a photon pair was produced in the first or the second crystal, the following entangled state is automatically generated:

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\big[|H\rangle_1|H\rangle_2 + e^{i\phi}|V\rangle_1|V\rangle_2\big].$$

The symbols $|H\rangle$ and $|V\rangle$ represent the horizontal and vertical polarization state of photons and the labels "1" and "2" correspond to the two spatial modes, which are in practice selected by e.g. pinholes or fibres. The relative phase ϕ is determined by the details of the phase matching and thickness of the crystals, but can be controlled by adjusting the relative phase between the horizontal and vertical components of the pump light.

The distinguishing information, which might possibly label the emission processes and thereby reduce their mutual coherence, can be either of temporal or spatial character. The latter case occurs whenever the emission modes from the two crystals are spatially distinguishable. To avoid this situation, the nonlinear crystals have to be thin enough and the down-conversion photons have to be collected into spatially single-mode channels, such as a pair of single-mode fibres. The use of thin crystals ensures that the emission cones from the two crystals overlap to a great extent. Moreover, the single-mode nature of the collection modes removes practically all the spatial information the photons may have carried before entering the fibre. Consequently, there is even in principle no way how to spatially distinguish whether the down-conversion photons are coming from the first or the second crystal and therefore pure polarization-entangled photon pairs can be detected.

In the time domain, the crystal birefringence in combination with dispersion lead to an unwanted effect as well. The arrival times of photons at the output face of the second crystal depend on their wavelengths and polarizations, which reveal the actual position of the photon-pair's origin. This leads to a partial loss of coherence between the two emission processes, and thus to the reduced entanglement quality. The detrimental temporal effect is two-fold, which is illustrated in a simplified fashion in Fig. 2. First, it is primarily the group-velocity mismatch between the pump and the down-conversion light, which causes the photon pairs born in the first crystal to be advanced with regard to those originating from the second crystal. This is usually precluded using a continuous-wave pump laser. Since a (spectrally broadband) free-running blue laser diode is used as the pump in the quED, a special birefringent crystal has to be included in the path of the pump beam. It introduces a proper temporal retardation between its horizontally and vertically polarized components and thus effectively pre-compensates the effect. Second, the dispersive delay of the down-conversion photons at non-degenerate wavelengths is different for the two emission possibilities, because the photons generated in the first crystal acquire an extra spread by propagating through the second crystal; (since the type I SPDC emission is spectrally very broadband, the detection of photons with very non-degenerate wavelengths is indeed possible in the quED). Therefore, an additional birefringent crystal has to be put behind the down-conversion crystals to counteract this second effect, too. The described double-crystal compensation technique ensures

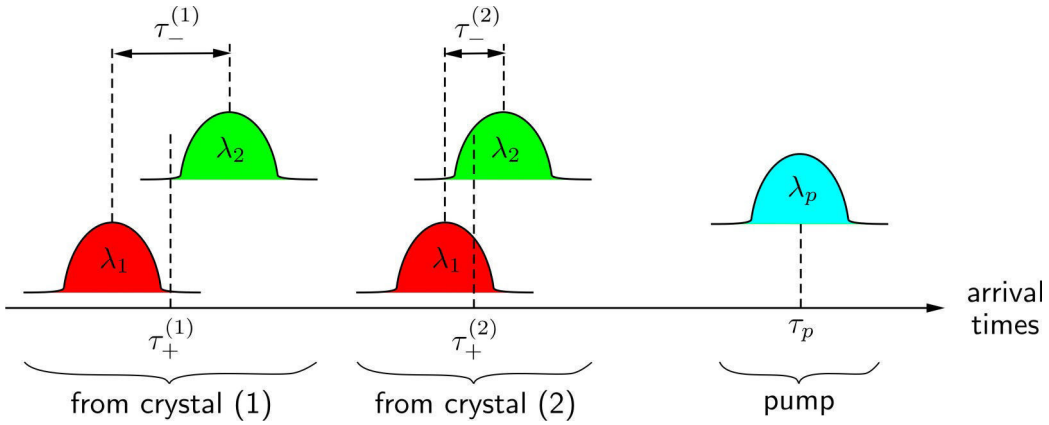


Figure 2: Explanation of the detrimental time effect inherent to SPDC emission in a two-crystal configuration. Due to crystal birefringence and dispersion, the arrival times of non-degenerate photons ($\lambda_1 \neq \lambda_2$) at the output face of the second crystal (2) differ in general for the two emission possibilities. The photon pairs from the first crystal (1) are advanced with regard to photon pairs from the second ($\tau_+^{(1)} < \tau_+^{(2)}$). Moreover, the photons originating from the first crystal experience higher dispersive delay due to their pass through the second crystal ($\tau_-^{(1)} > \tau_-^{(2)}$). Consequently, a compensation using two additional birefringent crystals erasing their temporal distinguishability has to be applied.

a complete temporal indistinguishability of the emission processes even though a free-running laser diode as a pump source is used and no spectral filtering of generated photons is applied in the quED.

3.6 Checking the Entanglement Quality (quick procedure)

1. If not already inserted, put the polarizing filters into the paths of down-conversion photons (the filters should be placed approximately perpendicular to the beam directions). Please make sure that the filters are centered on the paths of down-conversion photons.

2. Verify the existence of correlations in the horizontal-vertical polarization basis (maximum coincidence count rates for the filter combinations "horizontal/horizontal" and "vertical/vertical" and minimum coincidence count rates for the filter combinations "horizontal/vertical" and "vertical/horizontal").

3. Decide which of the Bell states should be prepared - either the Bell state Φ^-:

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}\big[|H\rangle_1|H\rangle_2 - |V\rangle_1|V\rangle_2\big],$$

showing anti-correlations in the diagonal basis or Φ^+:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\big[|H\rangle_1|H\rangle_2 + |V\rangle_1|V\rangle_2\big],$$

showing correlations in the diagonal basis.

4. Rotate the polarizing filters to positions corresponding to polarization analysis in the diagonal basis. Choose the combination of positions, which corresponds to the expected minimum in coincidence count rate for a chosen Bell state (i.e. combinations "+45°/-45°" or "-45°/+45°" for Φ^+ state and combinations "+45°/+45°" or "-45°/-45°" for Φ^- state).

6. Rotate the polarizing filters to positions corresponding to maximum in coincidence count rate in diagonal polarization basis.

7. Calculate visibility of the correlations in the diagonal polarization basis according to $(C_{\max} - C_{\min})/(C_{\max} + C_{\min})$, where $C_{\max}/C_{\min}$ is the maximum/minimum coincidence count rate. If the visibility is lower than expected we refer to section 3.7.

Understanding Correlated-Photon Based Efficiency Calibration of Photon Counting Detectors

The demand for detectors with photon counting capability grows rapidly. This is driven by advances in optical technologies, biophotonics and astronomy, or for example by novel quantum information applications. The requirement for simple, but precise, characterization of the photon-counting detectors goes hand in hand with this progress.

Conventional approach:

All traditional routes to efficiency measurements of detectors rely on comparison with externally calibrated reference sources or detectors. They always involve (i) establishing the primary source or primary detector standards, which define the radiometric scales at discrete wavelengths, (ii) extending the radiometric scales to a spectral continuum by characterizing intermediate standards, and (iii) dissemination of the scales to users.

The multistep procedure from establishing high-accuracy primary radiometric quantities to comparison of a tested device with the working standards is complex, not flexible, and resource-demanding, making the final costs for the user relatively high. Every step adds a calibration uncertainty into the entire procedure and thus a very high initial accuracy of the primary radiometric quantities is inevitably lost in the transfer chain of standards. The situation is even more complicated, because the present radiometric standards are precise at a power range well above the operating levels of photon counting detectors. Therefore, one needs calibrated attenuators, rendering the conventional calibration methods of photon counting detectors impractical and barely usable at all.

Correlated two-photon approach:

The working principle of this new attractive calibration method is based on a fundamental property of spontaneous parametric downconversion (SPDC) – the two-photon correlated emission. In SPDC, the photons from a pump laser beam spontaneously convert in a nonlinear crystal into a "train" of photon pairs. Since the photons are generated at random times in pairs like twins, the detection of one photon heralds, with certainty, the existence of the other. Not only the existence, but also the wavelength and the propagation direction of the other photon from a pair are accurately known due to restrictions of energy and momentum conservation,

$$\boldsymbol{k}_{\mathrm{p}} = \boldsymbol{k}_1 + \boldsymbol{k}_2, \quad \omega_{\mathrm{p}} = \omega_1 + \omega_2,$$

where ω and $\boldsymbol{k}$ are the photon frequencies and wave vectors and the subscripts $\mathrm{p}, 1$ and 2 refer to the pump and the two down-conversion photons, respectively.

The correlated nature of the two-photon light emitted from the SPDC offers a unique resource for absolute optical measurements of detector efficiency. The basic arrangement of the method is sketched in the figure 1. A pair of photon-counting detectors, a trigger detector (TRIG) and a device under test (DUT), are positioned behind the nonlinear crystal along the propagation directions of correlated photon pairs with given wavelengths. Whenever the TRIG registers a photon, the DUT should have ideally seen, in coincidence, one as well. But due to the finite detection efficiency of the DUT detector, only a fraction of TRIG detections is accompanied by a coincident detection at the DUT. It is just this fraction, which defines the efficiency of the DUT detector (assuming the detectors fire only due to correlated photon pairs and the DUT channel path from the point of SPDC emission to the detector is lossless).

Quantitatively, if N is the total number of correlated photon pairs emitted by SPDC in some

SPDC Emission Characteristics

The emission pattern of SPDC is formed by the cones which imprint characteristic rings in a plane perpendicular to the pump-beam direction. Each cone corresponds to a different down-conversion wavelength. If the nonlinear process is tuned to a spectral degeneracy, the photons of a pair are always located on the opposite sides of the same cone. In the case of spectral non-degeneracy the photons belong to different cones, yet they must be located on their opposite sides to obey the momentum conservation rule.

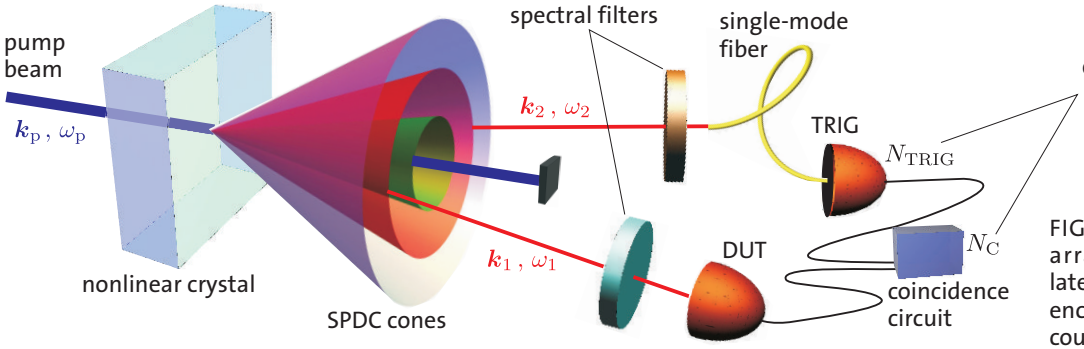


FIGURE 1. Basic optical arrangement for correlated-photon based efficiency calibration of photon-counting detectors.

arbitrary time unit, then the mean photon numbers N_{TRIG} and N_{DUT} registered by the two detectors and the number of detections N_{C} registered in coincidence are given by:

$$\begin{aligned} N_{\mathrm{TRIG}} &= \eta_{\mathrm{TRIG}} N, \\ N_{\mathrm{DUT}} &= \eta_{\mathrm{DUT}} N, \\ N_{\mathrm{C}} &= \eta_{\mathrm{TRIG}} \eta_{\mathrm{DUT}} N, \end{aligned}$$

where η_{TRIG} and η_{DUT} are the efficiencies of the detectors. As a result, the absolute value of η_{DUT} is simply detemined by:

$$\eta_{\mathrm{DUT}} = \frac{N_{\mathrm{C}}}{N_{\mathrm{TRIG}}}.$$

Remarkably, to determine the efficiency η_{DUT} of the DUT detector, the efficiency η_{TRIG} of the trigger detector need not be known! This makes the method inherently absolute. To highlight this point, one could even think of a measurement system where the DUT holds a role of its own trigger. In this measurement scenario the mutually delayed photons of a pair would be incident on the tested detector, whose time-resolved output reveals the efficiency in the same way as with a separate trigger detector.

Additional benefit comes from the spectral correlation of SPDC photons: The calibration wavelength can be defined by spectral selectivity components (with unknown transmission) put into the trigger arm. This allows the calibration regime to be transferred from some difficult spectral region to a more convenient one (e.g. from the infrared to visible by using detectors operating at widely separated spectral ranges).

Despite the advantages, there are several potentially problematic issues, which must be carefully considered, in order to turn the operation principle of the correlated-photon method into an accurate metrological measurement of detector efficiency. The issues are:

1. Any darkcount or count due to back-ground photons of the trigger detector falsely heralds the arrival of the correlated photon at the DUT detector. Therefore, the number N_{TRIG} of TRIG detections has to be subtractively corrected for the dark- and background-count rate, which can be easily measured.

2. As with false trigger events, a small fraction of coincidence detections is not due to correlated photons, but rather due to noise of detectors or background light. Thus, independent measurement has to be performed to determine the number of accidental coincidences.

3. The efficiency η_{DUT} given by the ratio of trigger-to-coincidence events includes all the losses of the DUT channel from the point of SPDC emission to the point of detection. Thus, in principle, the correlation method measures the efficiency of the entire DUT channel! The losses in the DUT path include the finite transmittance of the nonlinear crystal from the emission point, the transmittance of any pump-blocking filter and focusing optics and incomplete collection of the correlated photons. To extract the accurate efficiency of the detector DUT, all the losses have to be individually determined, and appropriate corrections performed.

System design considerations

To minimize the calibration uncertainties, ideally all the photons correlated to those recorded by the trigger detector should reach the DUT detector. To this end, one adopts an unbalanced setup geometry, in which all the spectral and spatial filtering is concentrated into the trigger channel (see fig. 1). Practically, this is realized by inserting a narrow-band spectral filter and the single-mode fiber into the trigger channel, whereas only a pump-blocking filter (with a high transmission over the entire spectral band correlated to that defined by the narrowband filter) and a large collection aperture (letting pass all the correlated photons through, while restricting the number of uncorrelated photons) is used in the DUT channel.