Quantum Probabilities

Material used

- Widdows' *Geometry of Meaning*, chapters 5-8
- *Quantum Logic and Probability Theory* by Alexander Wilce. *The Stanford Encyclopedia of Philosophy (Spring 2006 Edition)*
- *Quantum Mechanics – An Empiricists View* by Bas C. van Fraassen

1 Vector spaces with inner product
2 Boolean algebras vs. orthoalgebras
3 Application: Word-vectors and search engines
4 Quantum probabilities
5 Aspects of quantum computation
6 Applications in the macroworld

General Motivation

- Classical Probability theories and all modifications considered so far are based on a *Boolean algebra over W*, i.e. a set $\mathcal{F}$ of subsets of W that contains W and is closed under intersection, union and complementation. The elements of $\mathcal{F}$ are called events or propositions.
- The mathematics of Quantum Theory (QT) is based on vector spaces (Hilbert spaces) and 'events' or 'propositions' are considered as vector sub-spaces. The algebra of these subspaces is non-Boolean with regard to the corresponding operations (intersection, addition, ortho-complement)
- This requires a non-classical definition of probabilities!

Mixture vs. superposition

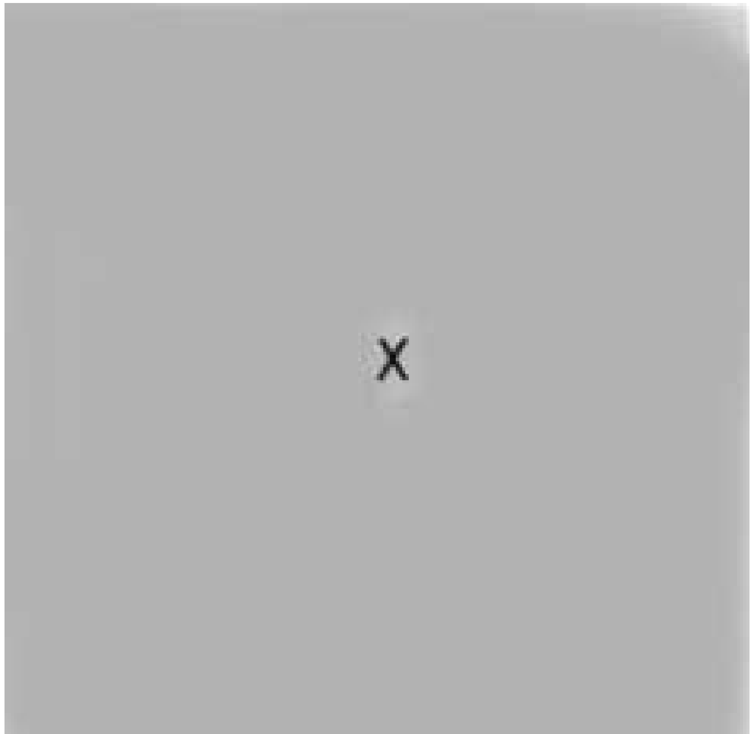
- Quantum mechanics makes a distinction between mixture and superposition of states. E.g., a light wave normally is a mixture of many photons. Sometimes, however, it comes to a superposition (vector-addition) of photons (double split experiment); see http://video.google.com/videoplay?docid=-4237751840526284618&q=quantum
- Mixture and superposition of states can also be found in the macroworld including the mental realm.
  - percepts
  - concepts
  - meanings

Afterimages and superposition

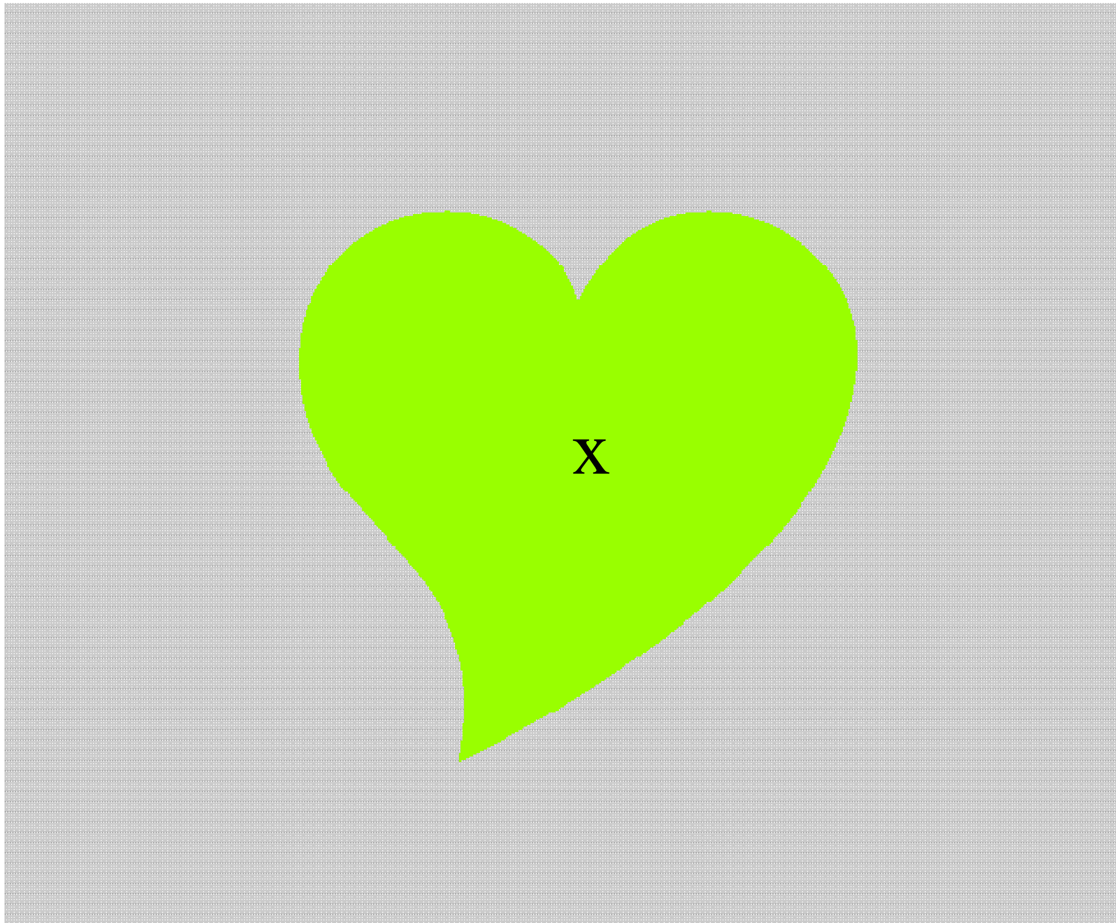
Als ich gegen Abend in ein Wirtshaus eintrat und ein wohlgewachsenes Mädchen mit blendendweißem Gesicht, schwarzen Haaren und einem scharlachroten Mieder zu mir ins Zimmer trat, blickte ich sie, die in einiger Entfernung vor mir stand, in der Halbdämmerung scharf an. Indem sie sich nun darauf hinweg-bewegte, sah ich auf der mir entgegen stehenden Wand ein schwarzes Gesicht, mit einem hellen Schein umgeben, und die übrige Bekleidung der völlig deutlichen Figur erschien in einem schönen Meergrün.

Afterimages and superposition

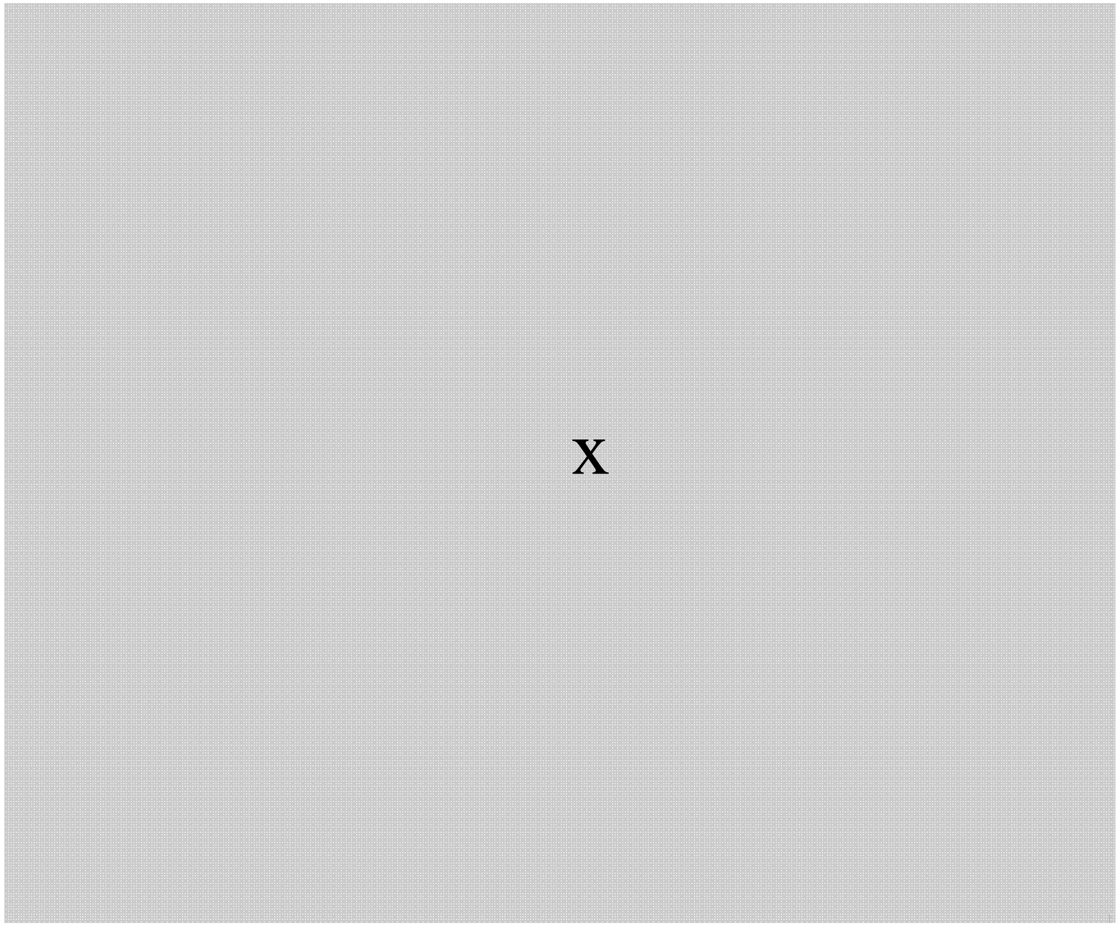
Als ich gegen Abend in ein Wirtshaus eintrat und ein wohlgewachsenes Mädchen mit blendendweißem Gesicht, schwarzen Haaren und einem scharlachroten Mieder zu mir ins Zimmer trat, blickte ich sie, die in einiger Entfernung vor mir stand, in der Halbdämmerung scharf an. Indem sie sich nun darauf hinweg-bewegte, sah ich auf der mir entgegen stehenden Wand ein schwarzes Gesicht, mit einem hellen Schein umgeben, und die übrige Bekleidung der völlig deutlichen Figur erschien in einem schönen Meergrün.



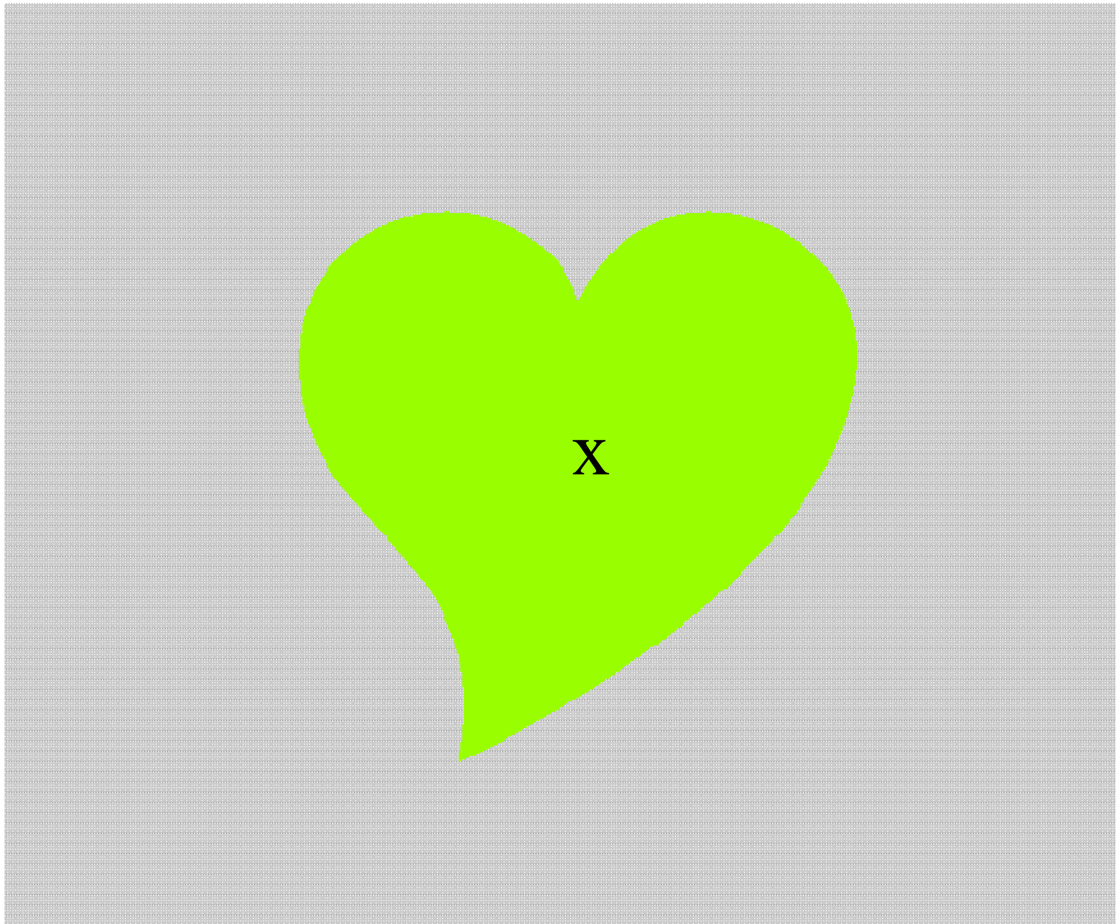
More afterimages



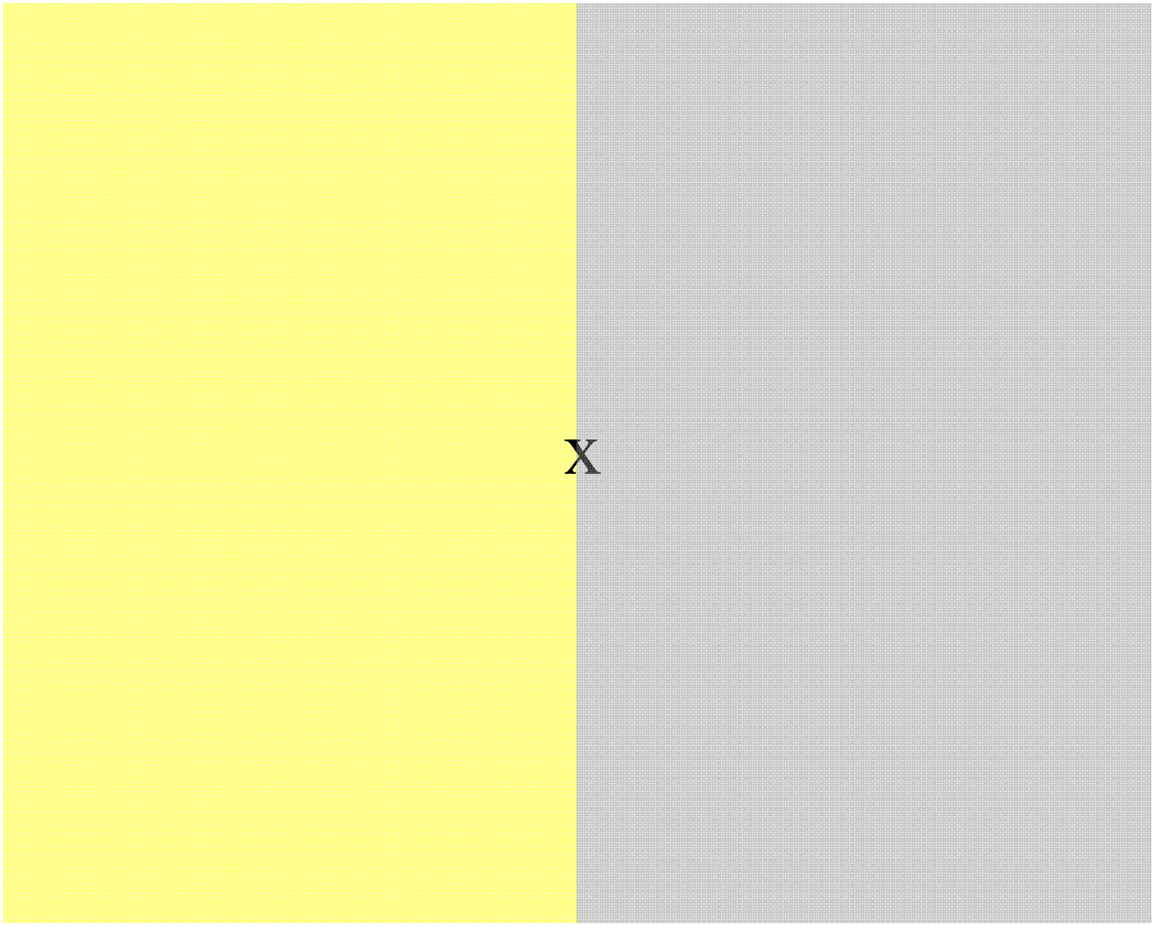
More afterimages



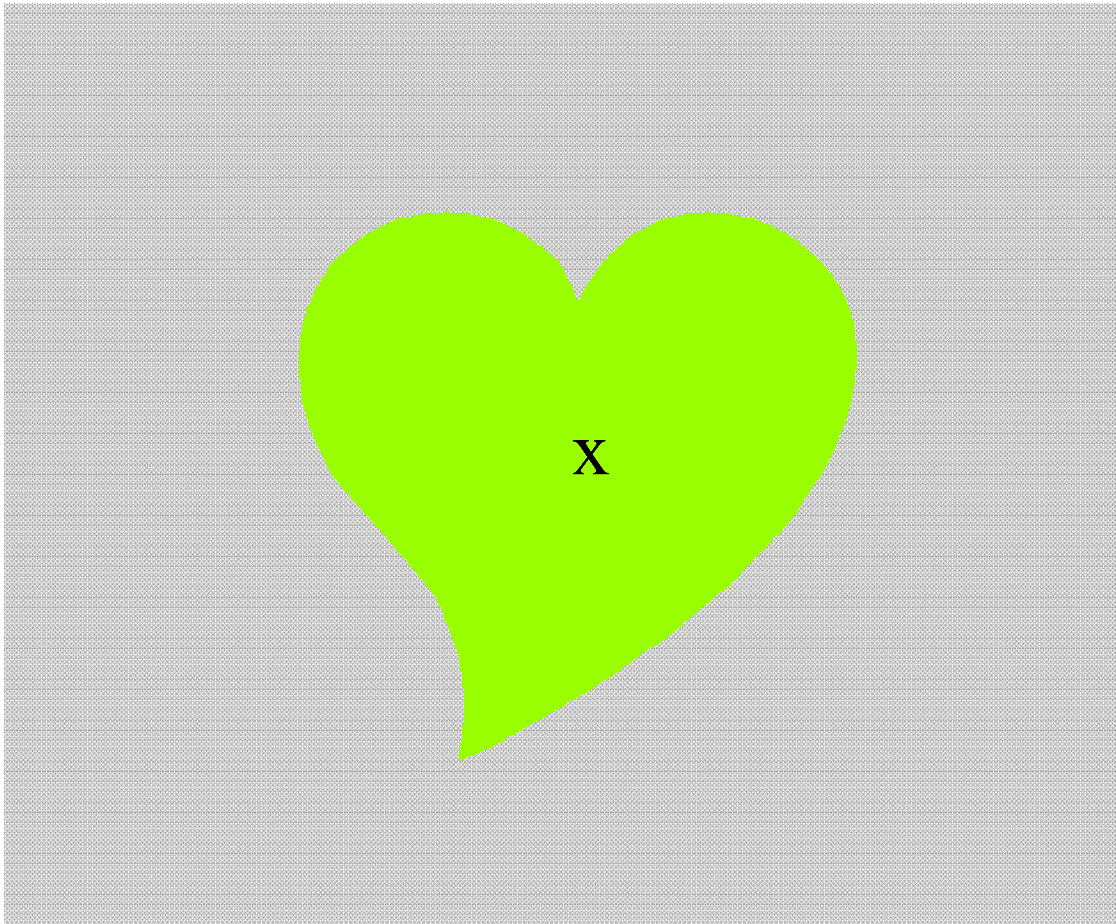
Superposition



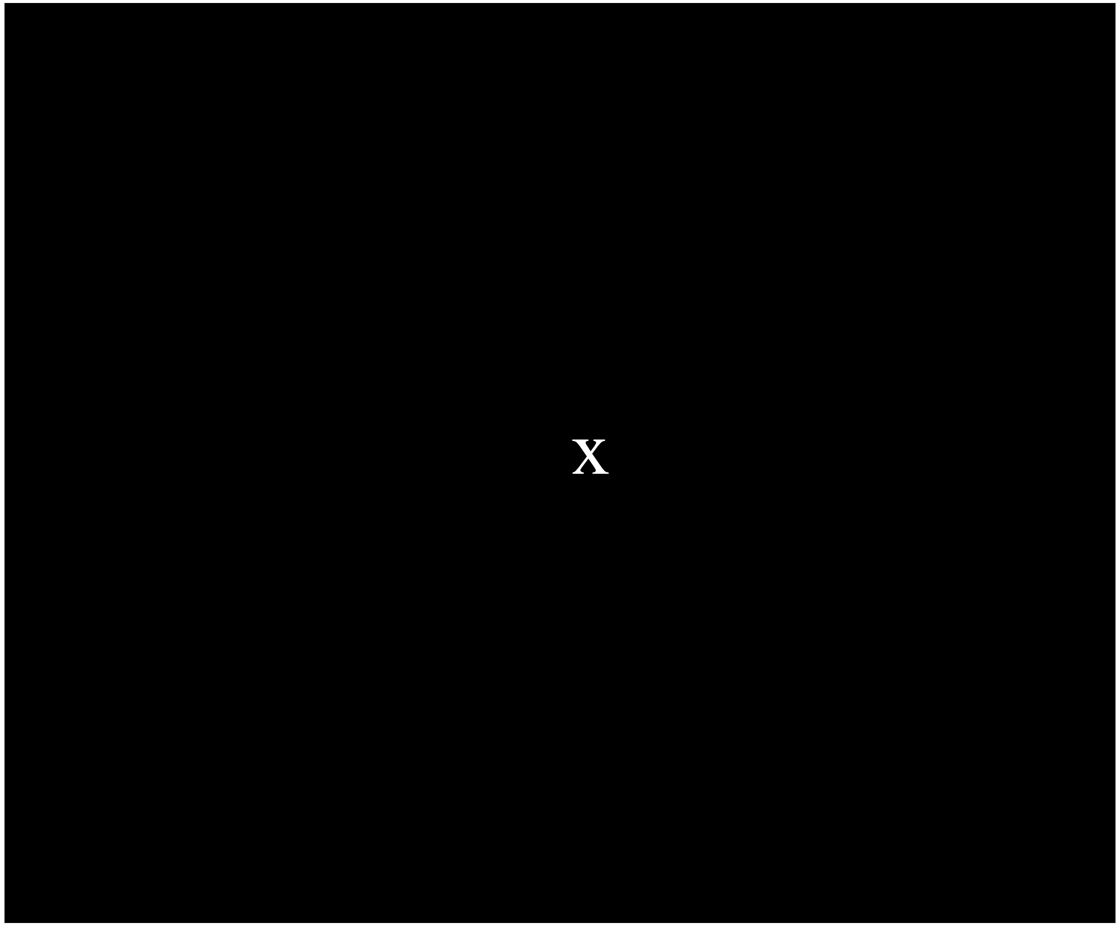
Superposition



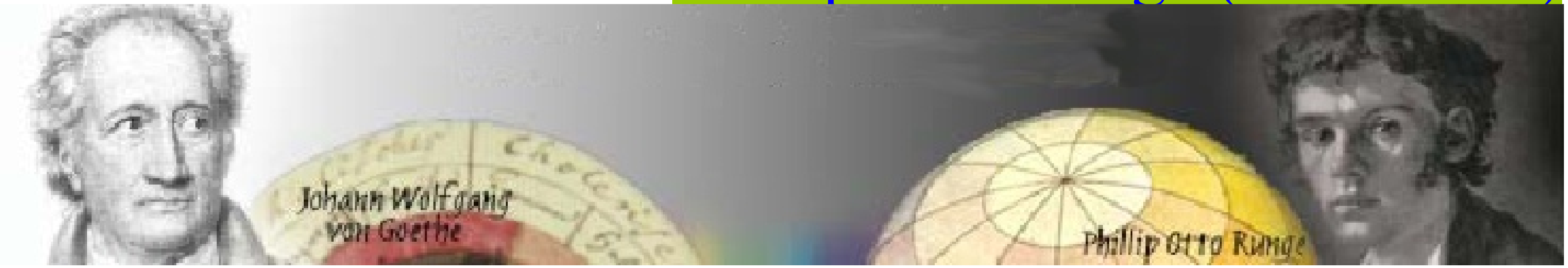
Churchland's chimeric afterimages

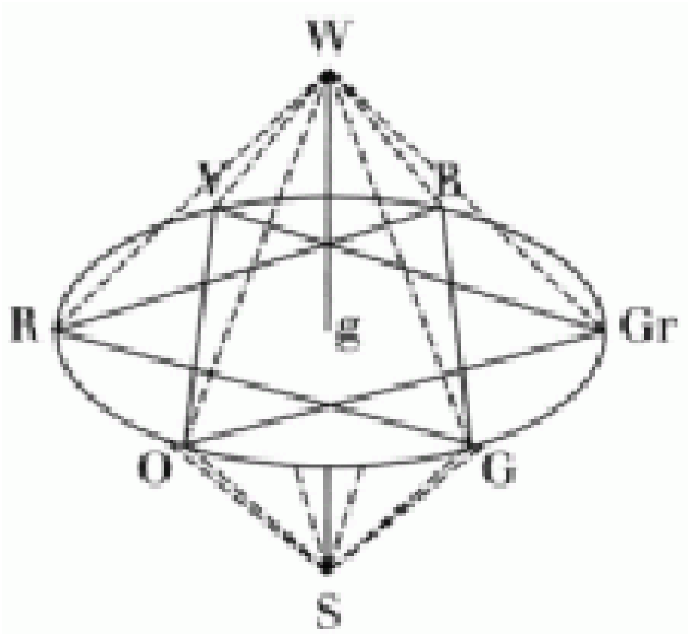


Churchland's chimeric afterimages

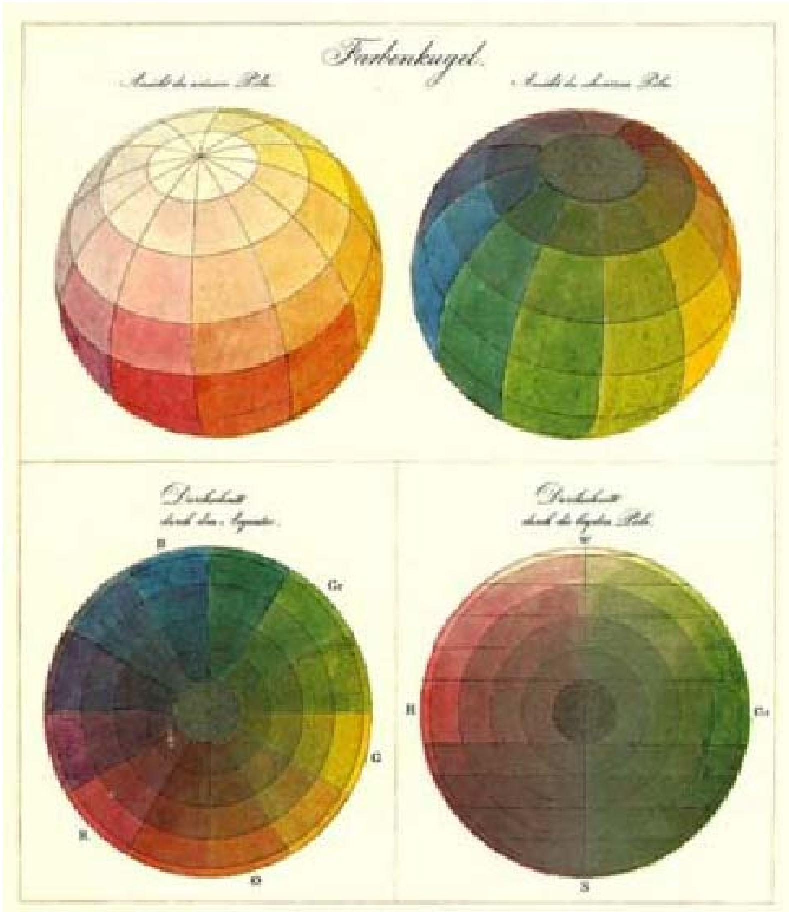


Phillip Otto Runge (1777-1810)





Hue, Saturation, Luminance

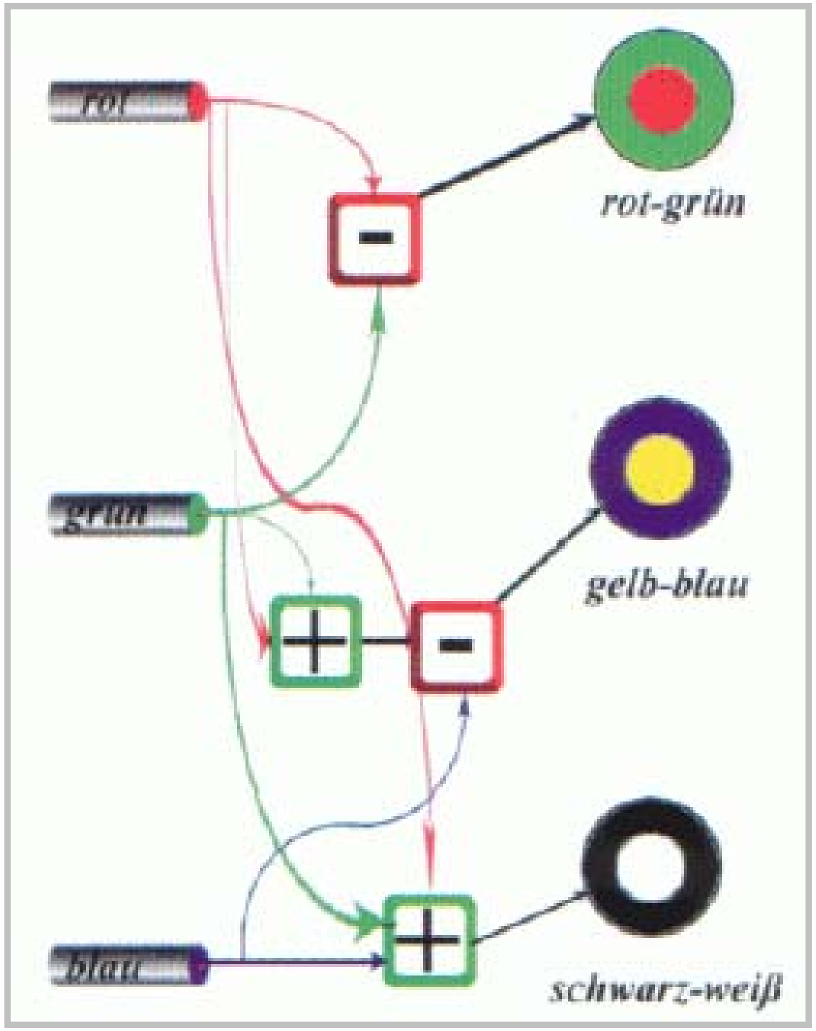


Simple neural network

The simple network connects the receptors on the retina (3 types of cones: RED, GREEN, BLUE) with the visual cortex (two opponent systems red-green, yellow-blue; one non-opponent system black-white).

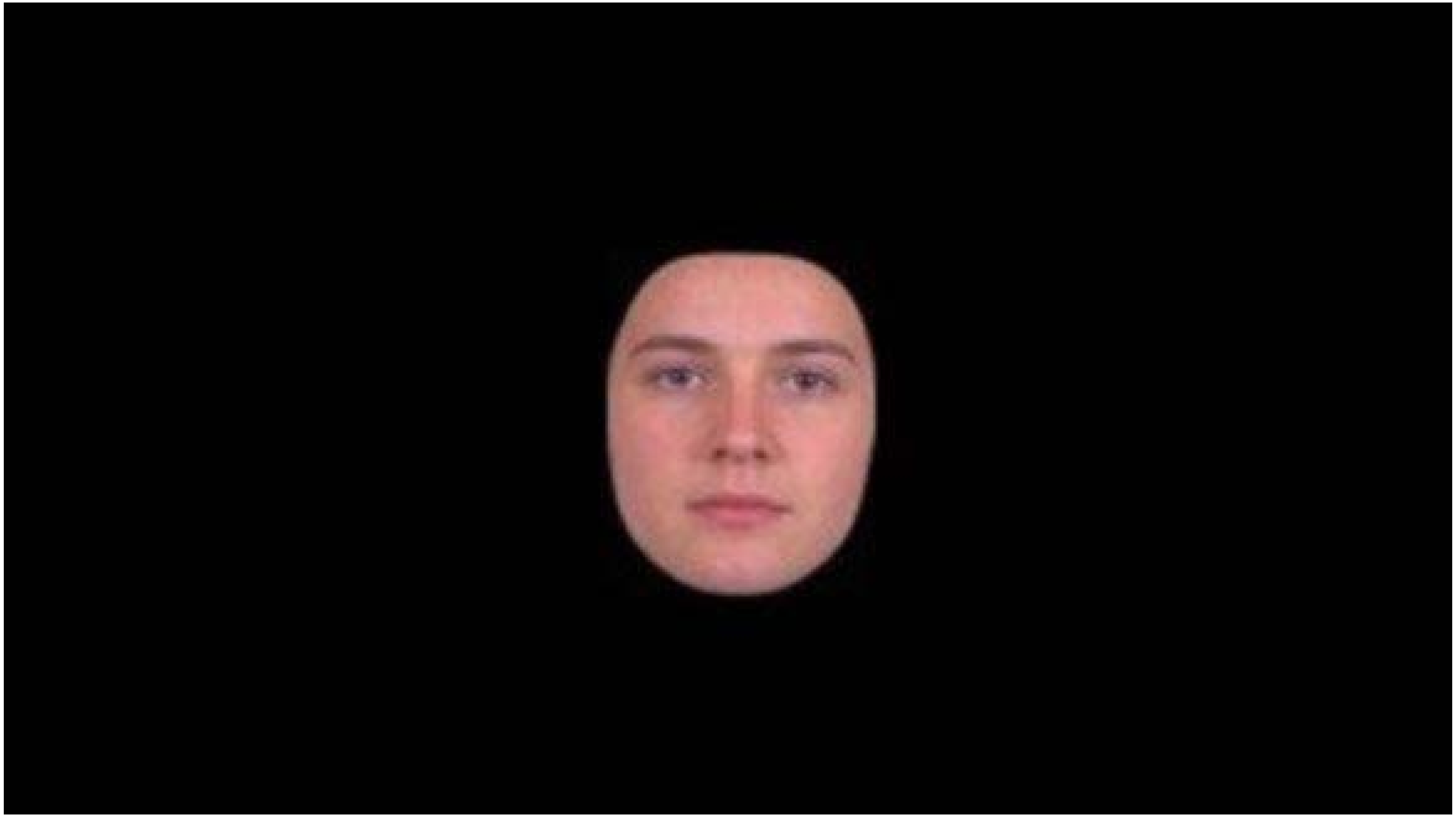
Hering, Ewald 1964 [1920]. *Outlines of a Theory of the Light Sense*. Cambridge, Mass.: Harvard University Press.

Jameson, D. and L.M. Hurvich (1955) Some quantitative aspects of an opponent-colors theory. Journal of the Optical Society of America, 45:546-52.



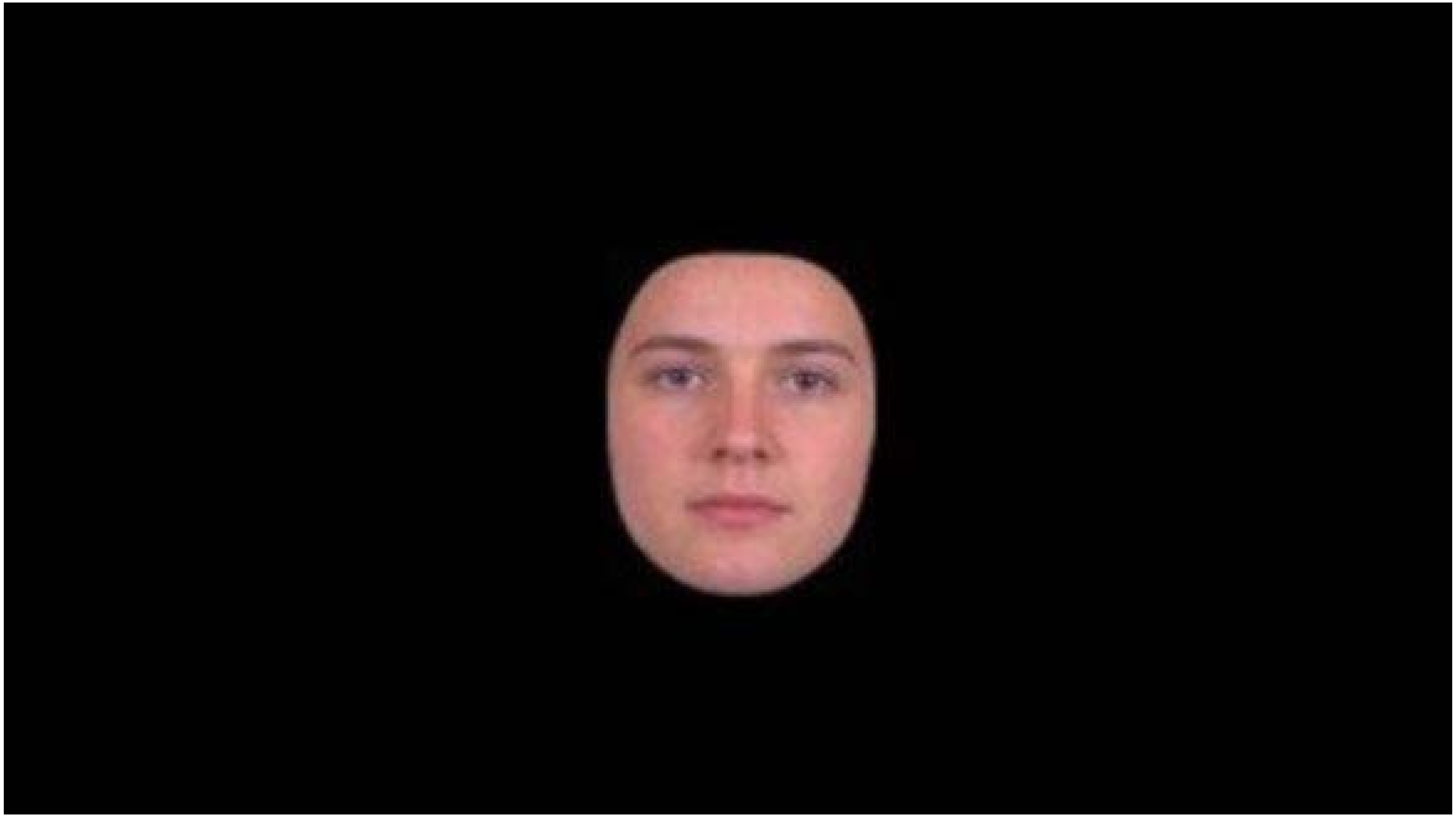
Superposition of faces

Superposition of faces

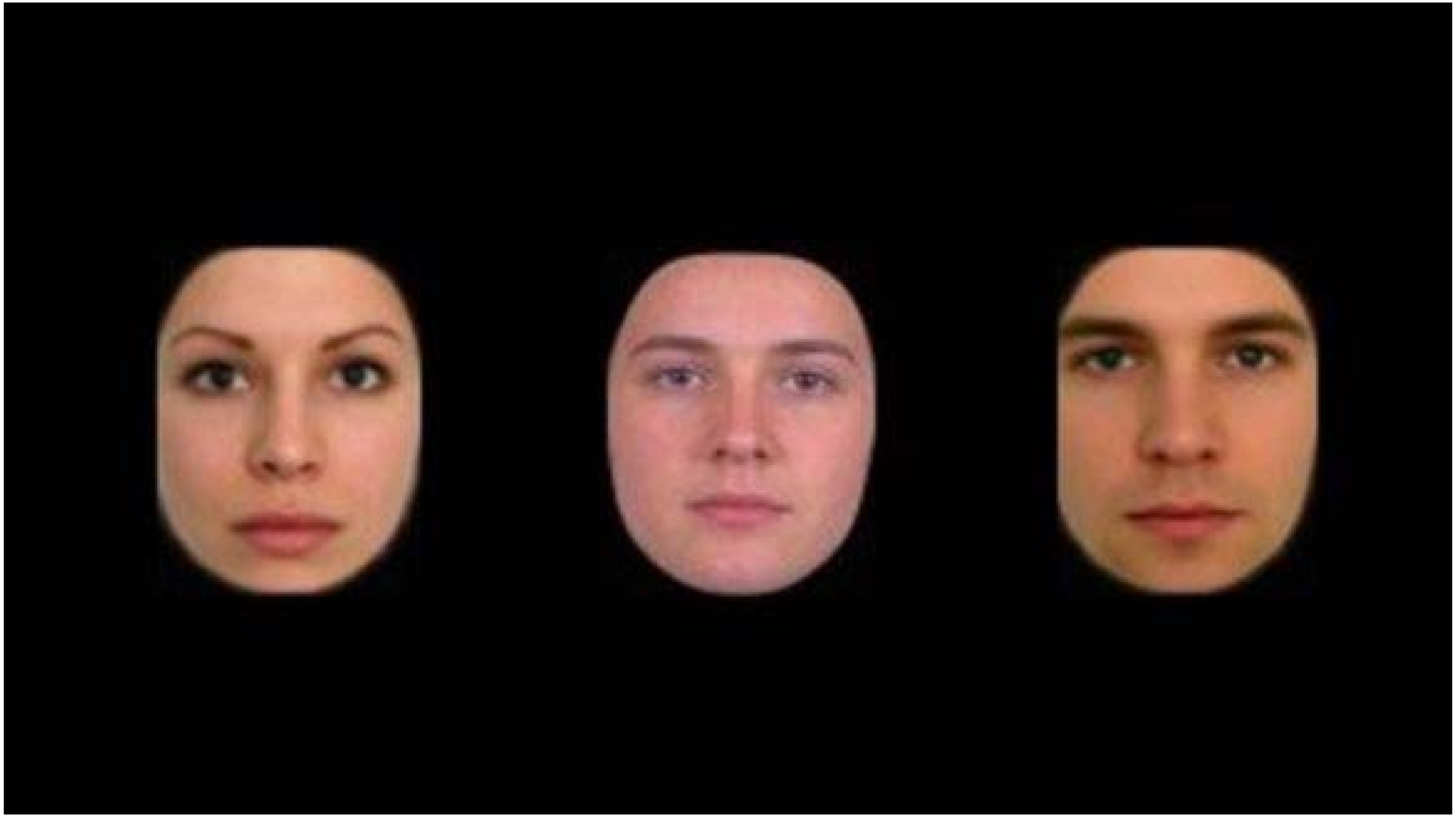


Superposition of faces

Superposition of faces



Superposition of faces



The postulates of quantum mechanics (informal)

1. **States**. States of a physical system are represented by vectors in a Hilbert space (→ superposition)

2. **Observables**. Observables are represented by Hermitian operators (i.e. operators having real eigen values)

3. **Evolution**. The evolution of a (closed) system is described by a unitary transformation (leaving the scalar product of two vectors invariant)

The postulates of quantum mechanics (informal)

4. **Measurement**. Measurement is an external observation of a system and so disturbs its unitary evolution. A quantum state u can be measured by use of a set of orthogonal projections $u_1, \ldots, u_n$; u collapses into the state u_i with probability $|\langle u_i , u\rangle|^2$ (→ uncertainty principle)

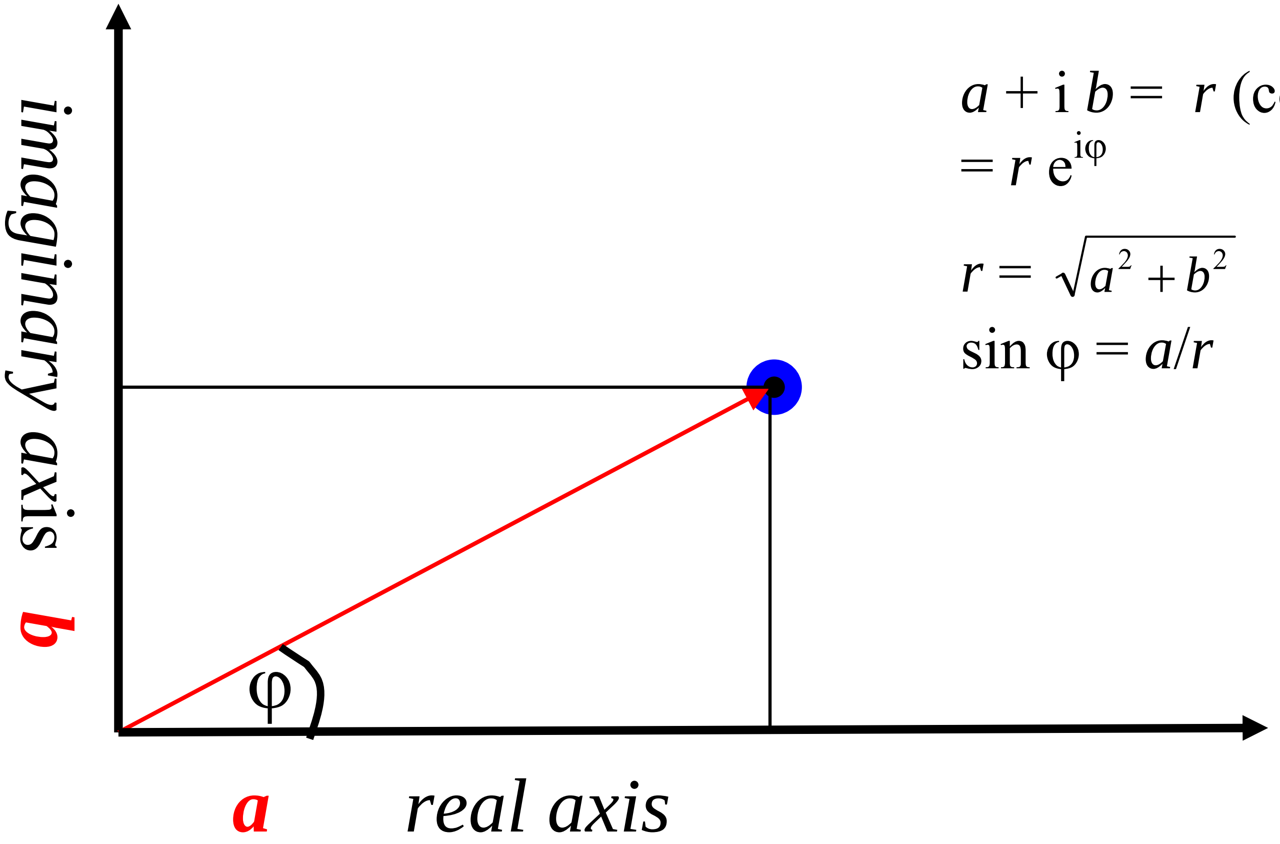
5. **Composite systems**. The state space of a composite system is the tensor product of the state spaces of its components. The unitary evolution of composed states $u \otimes v$ can lead to states that cannot longer be expressed as a tensor product of two vectors in the original two subsystems.
(→ entanglement)

Potential application

- Applications in the micro world
  - Quantum Computation
- Applications in the macro world
  - face recognition
  - manipulation of word vectors and geometrical meanings; modelling concepts
  - opinion forming
  - interference effects in perception and cognition
  - understanding the logic of activation patterns in connectionism
  - modelling types of personality (C.G. Jung)

Appendix: a note on complex numbers

Complex numbers can be understood as points/vectors in a 2-dimensional space.



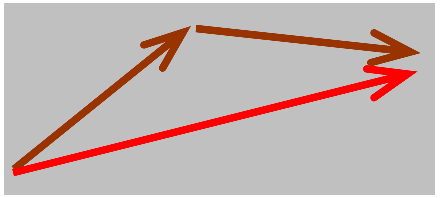
$a + \mathrm{i}\, b = r\,(\cos \varphi + \mathrm{i} \sin \varphi)$
$= r\, \mathrm{e}^{\mathrm{i}\varphi}$

$r = \sqrt{a^2 + b^2}$

$\sin \varphi = a/r$

Some rules

1. $u+v = v+u$
2. $(u+v)+w = u+(v+w)$
3. $u\,v = v\,u$
4. $u\,(v\,w) = (u\,v)\,w$
5. $\lambda(u+v) = \lambda u+\lambda v$
6. $(u+v)^* = u^* + v^*$
7. $(u\,v)^* = u^*\,v^*$; where u^* is the complex conjugate: $(a + \mathrm{i}\,b)^* = a - \mathrm{i}b$.



1 Vector spaces with inner product

Formally, an inner product space is a vector space V over the field of real/complex numbers with a so-called inner product $\langle\, .\, |\, .\, \rangle$, which satisfies the following conditions for all $x, y, z \in V$ and all complex scalars a:

1. $\langle x \mid y\rangle = \langle y \mid x\rangle^*$ (*Conjugate symmetry*)
2. $\langle a\, x \mid y\rangle = a\, \langle x \mid y\rangle$ (*Linearity in the first variable*)
   $\langle x+y \mid z\rangle = \langle x \mid z\rangle + \langle y \mid z\rangle$
3. $\langle x \mid x\rangle \geq 0$ and $\langle x \mid x\rangle = 0$ if and only if $x = 0$. (*Nonnegativity and nondegeneracy*)

Example 1

In $\mathbb{R}_N$, the usual dot product $\langle x \mid y \rangle = \sum_{k=0}^{N} x_k \cdot y_k$

Example 2

In C_N, the dot product $\langle x \mid y \rangle = \sum_{k=0}^{N} x_k \cdot y_k^*$

Example 3

For complex-valued continuous-time signals in C_N[a,b]

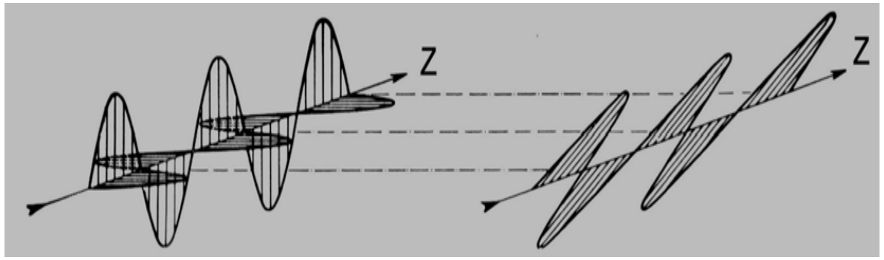
$$\langle x \mid y \rangle = \int_a^b x(t) \cdot y^*(t) dt$$

A **Hilbert space** $\mathscr{H}$ is a vector space with inner product; it is not restricted to finite dimensions. In the infinite case, it has to satisfy a certain completion requirement – the so-called Cauchy criterion with regard to the norm $\|x\|^2 = \langle x \mid x \rangle$.

The Cauchy criterion may be defined for sequences in this space (as it can in any uniform space): a sequence $\{x_n\}$ is a Cauchy sequence if for every positive real number ε there is a natural number N such that for all m, $n > N$, $\|x_n - x_m\| < \varepsilon$. $\mathscr{H}$ is a **Hilbert space** if it is complete with respect to this norm, that is if every Cauchy sequence converges to an element in the space. Thus, every Hilbert space is also a Banach space (but not vice versa).

Assume x and y to be vectors in $\mathbb{R}_N$. Show that

1. $\|x + iy\|^2 = \|x\|^2 + \|y\|^2$
2. $\|x - iy\|^2 = \|x\|^2 + \|y\|^2$
3. $\langle x + y \mid x - y\rangle = \|x\|^2 - \|y\|^2$
4. $\langle x + iy \mid x - iy\rangle = \|x\|^2 - \|y\|^2 + 2i\langle x \mid y\rangle$.



2 Boolean algebras vs. orthoalgebras

A *Boolean algebra over W* is a set $\mathcal{F}$ of subsets of W such that it contains W and is closed under intersection, union and complementation.

Instead of W we consider now a vector space $\mathcal{H}$; U, $V \subseteq \mathcal{H}$. $U \cap V$ is a vector space again. However, $\neg U$ and $U \cup V$ are normally no vector spaces (examples?). Consequently, we cannot form a *Boolean algebra over* $\mathcal{H}$ if it is required that the elements of this algebra are vector spaces!

Problem: What operations can we take in order to save the idea of propositional structures in vector spaces?

Orthoalgebras

Let $\mathcal{H}$ be a Hilbert space and U, V sub (Hilbert) spaces of $\mathcal{H}$. It's simply to show that $U \cap V$ is a Hilbert space.

Ordered by set-inclusion $\subseteq$, the closed subspaces of $\mathcal{H}$ form a complete lattice, in which the meet (greatest lower bound) of a set of subspaces is their intersection, while their join (least upper bound) is the closed span of their union (sum). There are (infinitely) many complementary closed subsets, one of them is the orthocomplement.

Intersection: $U \cap V$

Sum: $U + V = \{w\colon w = u + v \text{ for some } u \in U \text{ and some } v \in V\}$

Orthocomplement: $U^{\perp} = \{v \in \mathcal{H}\colon \forall u \in U, \langle u \mid v \rangle = 0\}$

Facts

Let $\mathcal{H}$ be a Hilbert space and U, V sub (Hilbert) spaces of $\mathcal{H}$.
Then it holds

1. $(U^{\perp})^{\perp} = \mathrm{U}$
2. $(U \cap V)^{\perp} = U^{\perp} + V^{\perp}$
3. $(U + V)^{\perp} = U^{\perp} \cap V^{\perp}$

Distributivity doesn't hold generally!
$(U+V) \cap W \neq (U \cap W) + (V \cap W)$

In $\mathbb{R}_3$ assume a vector space U spanned by the unit vector $u = \frac{1}{\sqrt{3}}(1,\ 1,\ 1)$. What is $U^{\perp}$? Give an orthonormal basis for $U^{\perp}$!

$(0, 1, -1) \rightarrow \frac{1}{\sqrt{2}}(0, 1, -1)$

$(1, 0, -1)$

$(1, -1, 0)$

$(2, -1\ -1) \rightarrow \frac{1}{\sqrt{6}}(2, -1\ -1)$

Operators in the Hilbert space

- An (linear) operator or transformation $\hat{O}$ on a Hilbert space $\mathcal{H}$ is a Hilbert space morphism of $\mathcal{H}$ into $\mathcal{H}$
- Example $\mathbb{R}^N$, $\mathcal{C}_N$: operators are represented by matrices of real/complex numbers.

$$[\hat{O}\, u]_i = \sum_j \hat{O}_{ij}\, u_j \text{ (matrix multiplication)}$$

- The *adjoint* $\hat{O}^+$ of an operator $\hat{O}$ is that operator such that

$$\langle \hat{O}^+ u \mid v \rangle = \langle u \mid \hat{O} v \rangle, \text{ for all elements } u,v \text{ of } \mathcal{H}.$$

- The *trace* of an operator is

$$\mathbf{Tr}(\hat{O}) = \sum_i \langle \mathbf{e}_i \mid \hat{O}\mathbf{e}_i \rangle, \text{ with an orthonormal basis } \{\mathbf{e}_i\} \text{ of } \mathcal{H}.$$

Observables and eigenvalues

- An eigenvalue a of an operator Ô is a complex number for which there is an element of $\mathcal{H}$ such that

$$\hat{O}\, u = \lambda\, u$$

The element u is called an eigen-state of Ô corresponding to the eigenvalue λ.

- In quantum mechanics, an observable is simply a Hermitian (also called self-adjoint) operator on a Hilbert space $\mathcal{H}$, i.e., an operator Ô such

$$\hat{O}^{+} = \hat{O}.$$

Theorem: The eigenvalues λ_i of an observable Ô are all real numbers. Moreover, the eigen-states u_i for distinct eigenvalues of an observable are orthogonal. Hence, we can write

$$\hat{O} = \Sigma_i \lambda_i \, |u_i\rangle\langle u_i|$$

- An eigenvalue is called *degenerate* if there are at least two linearly independent eigen-states for that eigenvalue. Otherwise, it is called *non-degenerate*.
- Usually, the eigenvalues of observables can be assumed to be non-degenerate. In this case, we have

$$\hat{O}\, u_i = \lambda_i\, u_i\, ,\ \langle u_i \mid u_j\rangle = \delta_{ij}$$

Example: the Q-bit

The Q-bit is a physical system each of its (nontrivial) observables has a discrete spectrum with two non-degenerate eigenvalues (say ±1): **Sp**($\hat{X}$) = **Sp**($\hat{Y}$) = **Sp**($\hat{Z}$) = {-1,+1}.

$$\hat{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \hat{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \hat{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\hat{Z}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |\updownarrow\rangle \qquad \hat{X}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad |\nearrow\rangle \qquad \hat{Y}\begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} 1 \\ i \end{pmatrix} \quad |\circlearrowleft\rangle$$

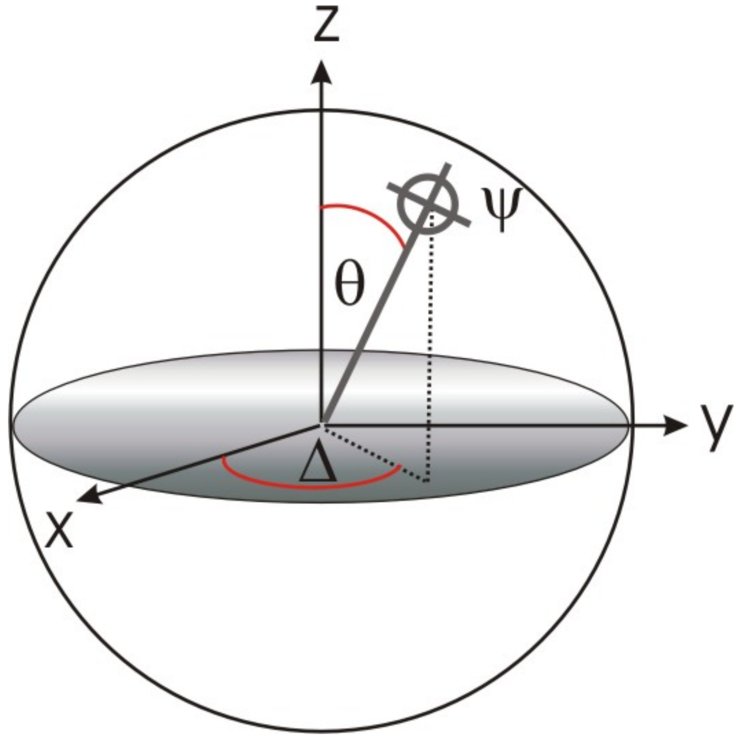
$$\hat{Z}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad |\leftrightarrow\rangle \qquad \hat{X}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = -\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad |\searrow\rangle \qquad \hat{Y}\begin{pmatrix} 1 \\ -i \end{pmatrix} = -\begin{pmatrix} 1 \\ -i \end{pmatrix} \quad |\circlearrowright\rangle$$

The Bloch sphere

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \text{ with } \alpha^2 + \beta^2 = 1$$

$$|\psi\rangle = \cos(\theta/2)\, e^{-i\Delta/2}\, |0\rangle + \sin(\theta/2)\, e^{+i\Delta/2}\, |1\rangle$$

Figure: an arbitrary (normalized) state of the two dimensional Hilbert space can parameterized by the two spherical polar coordinates θ and Δ. Hereby, Δ corresponds to a phase shift of the two superposing states $|0\rangle$ and $|1\rangle$)



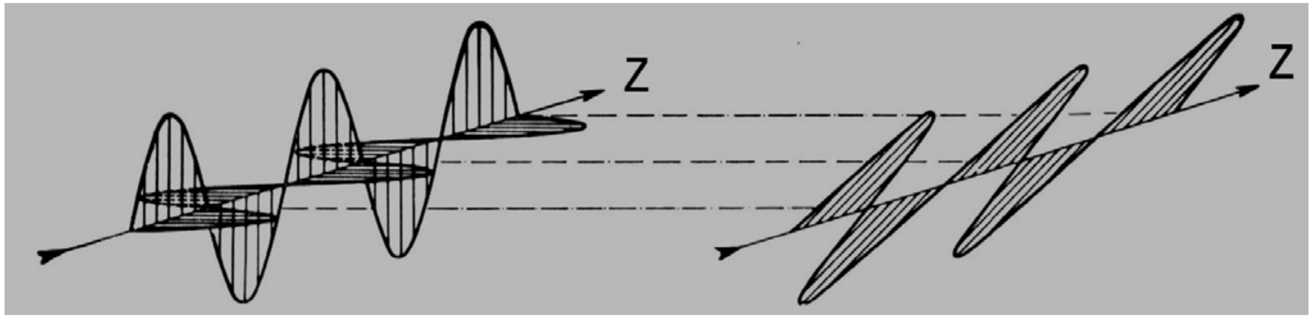
How can a Q-bit be realized?

- Two polarizations of a photon
- Alignment of a nuclear spin (proton!) in a uniform magnetic field
- Two states of an electron orbiting a single atom (ground or excited state) atom

Stationary photons as Q-bits

- **H**orizontal and **V**ertical polarization: $|\updownarrow\rangle$ and $|\leftrightarrow\rangle$
  (can be seen as basis vectors of C_2)
- Equal superposition of $|\updownarrow\rangle$ and $|\leftrightarrow\rangle$:

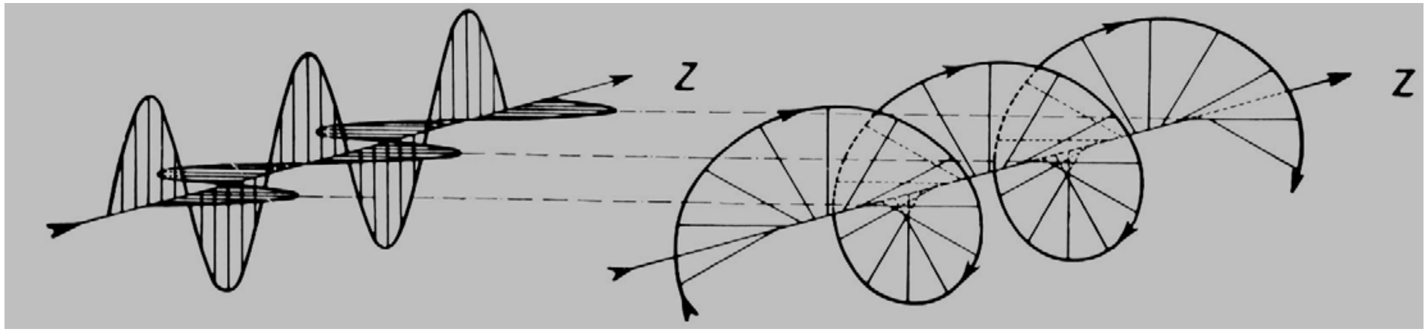
$$|\nearrow\rangle = \tfrac{1}{\sqrt{2}}(|\updownarrow\rangle + |\leftrightarrow\rangle); \quad |\searrow\rangle = \tfrac{1}{\sqrt{2}}(|\updownarrow\rangle - |\leftrightarrow\rangle)$$



Circular polarization

- 2 states ↺ and ↻ of circular polarization
- Equal superposition of of $|\updownarrow\rangle$ and $|\leftrightarrow\rangle$ (the latter with $\pm\,\pi/2$ phase shift; note $i = e^{i\,\pi/2}$):

$$|\circlearrowleft\rangle = \tfrac{1}{\sqrt{2}}(|\updownarrow\rangle + i\,|\leftrightarrow\rangle);\ |\circlearrowright\rangle = \tfrac{1}{\sqrt{2}}(|\updownarrow\rangle - i\,|\leftrightarrow\rangle)$$



Definition

A self-adjoint operator $\hat{O}$ with the property $\hat{O}^2 = \hat{O}$ is called a projection (= projection operator)

- It is not difficult to show that a self-adjoint operator $\hat{O}$ with spectrum $\mathbf{Sp}(\hat{O}) \subseteq \{0,1\}$ must be a projection
- Projection operators are in one-to-one correspondence with the closed subspaces of $\mathcal{H}$.
  - Assume a projection operator $\hat{O}$ with eigenvectors $\hat{O}(u_i) = u_i$. Then the orthonormal vectors u_i span a closed subspace U of $\mathcal{H}$.
  - Conversely, if U is such a closed subspace, then the operator $\sum_i |u_i\rangle\langle u_i|$ is a projection.

Definitions

- $\| \hat{P} \| = \{u \in \mathcal{H}:\ \hat{P} u = u\}$ (*range of an operator*)
- With the projections $\hat{P}$ and $\hat{Q}$, the projections $\hat{P}^{\perp}$, $\hat{P} \cup \hat{Q}$ and $\hat{P} \cap \hat{Q}$ are defined by the conditions:
  1. $\| \hat{P}^{\perp} \| = \| \hat{P} \|^{\perp}$
  2. $\| \hat{P} \cup \hat{Q} \| = \| \hat{P} \| + \| \hat{Q} \|$
  3. $\| \hat{P} \cap \hat{Q} \| = \| \hat{P} \| \cap \| \hat{Q} \|$

Facts

1. $\hat{P}^{\perp} u = u - \hat{P} u$ (i.e. $\hat{P}^{\perp} = \mathbf{1} - \hat{P}$)
2. $(\hat{P} \cup \hat{Q})u = \hat{P} u + \hat{Q} u$ if $\hat{P} \hat{Q} = -\hat{Q} \hat{P}$
3. $(\hat{P} \cap \hat{Q})u = \hat{P} \hat{Q} u$ if $\hat{P} \hat{Q} = \hat{Q} \hat{P}$

Axioms for an Ortho-lattice

$$a = a^{\perp\perp}$$

$$a \cup b = b \cup a$$

$$(a \cup b) \cup c = a \cup (b \cup c)$$

$$a \cup (b \cup b^{\perp}) = b \cup b^{\perp}$$

$$a \cup (a^{\perp} \cup b)^{\perp} = a$$

$$a = b \quad \Rightarrow \quad b = a$$

$$a = b \ \& \ b = c \quad \Rightarrow \quad a = c$$

$$a = b \quad \Rightarrow \quad a^{\perp} = b^{\perp}$$

$$a = b \quad \Rightarrow \quad (a \cup c) = (b \cup c)$$

Definition

$$(a \cap b) = (a^{\perp} \cup b^{\perp})^{\perp}$$

Orthomodular Law Axiom

$$(c \cup c^{\perp}) = ((a^{\perp} \cup b^{\perp})^{\perp} \cup (a \cup b)^{\perp}) \quad \Rightarrow \quad a = b$$

Foulis-Holland Theorems

$aCb \ \& \ aCc \Rightarrow a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$
$aCb \ \& \ aCc \Rightarrow b \cap (a \cup c) = (b \cap a) \cup (b \cap c)$
$aCb \ \& \ aCc \Rightarrow a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$
$aCb \ \& \ aCc \Rightarrow b \cup (a \cap c) = (b \cup a) \cap (b \cup c)$

The relation

$$a\ C b := \ a = (a \cap b) \cup (a \cap b^{\perp})$$

is read “a commutes with b”.

Conclusions

- Projection operators correspond to Boolean random variables in the classical case
- In both cases there is a one-one correspondence between the projection operator/Boolean random variable and the **range** defined by these elements (a subspace of the domain under discussion)
- Whereas Boolean random variables commute, projection operators do not necessarily.
- The familiar distributive laws of classical logic are not obeyed here. De Morgan's laws are valid (taking negation as orthocomplement)

3 Application: Word-vectors and search engines

- Representing word meanings by vectors
- Similarity and orthogonality
- Excurse about LSA
- Superposition and ambiguities

see Dominic Widdows' *Geometry of Meaning*, chapters 5-8.

Also relevant: Keith van Rijsbergen, *The Geometry of Information Retrieval*, Cambridge University Press, 2004

Word vectors and the term-document matrix

Consider the frequencies of words in certain documents. From this information we can construct a vector for each word reflecting the corresponding frequencies:

	Document 1	Document 2	Document 3
bank	0	0	4
bass	2	4	0
commercial	0	2	2
cream	2	0	0
guitar	1	0	0
fishermen	0	3	0
money	0	1	2

Document 1 is about music instruments, document 2 about fishermen, and document 3 about financial institutions.

Normalization and similarity

- $d(u, v) = \|u - v\|$
  not very reasonable to measure similarity since larger vectors (corresponding to more frequent words) tend to be more distant from most other vectors than small vectors.
- Use unit vectors (normalization)
  $\cos(u, v) = \langle u \mid v \rangle / \|u\| \cdot \|v\|$
  $d(u, v) = \sqrt{2(1 - \cos^2(u, v))}$ (for unit vectors)
- Similarities between words and documents:
  $\cos(u, doc^i)$, where doc^i is the i^{th} unit vector (representing the document as a whole).

Example

	Document 1	Document 2	Document 3
bank	0	0	1
bass	0.447	0.894	0
commercial	0	0.707	0.707
cream	1	0	0
guitar	1	0	0
fishermen	0	1	0
money	0	0.447	0.894

	guitar *cream*	*guitar* *bass*	*guitar* *fisherman*
d(*u*,*v*)	0	0.81	1.41
cos(*u*,*v*)	1	0.477	0
cos(*v*, doc^1)	1	0.447	0
cos(*v*, doc^2)	0	0.894	1

Two Problems

- Ambiguous words: words such as *bass* are ambiguous (i. music instrument, ii. fish). If a user is only interested in one of these meanings, how are we to enable users to search for only the documents containing this meaning of the word?
  → Superposition, Vector negation

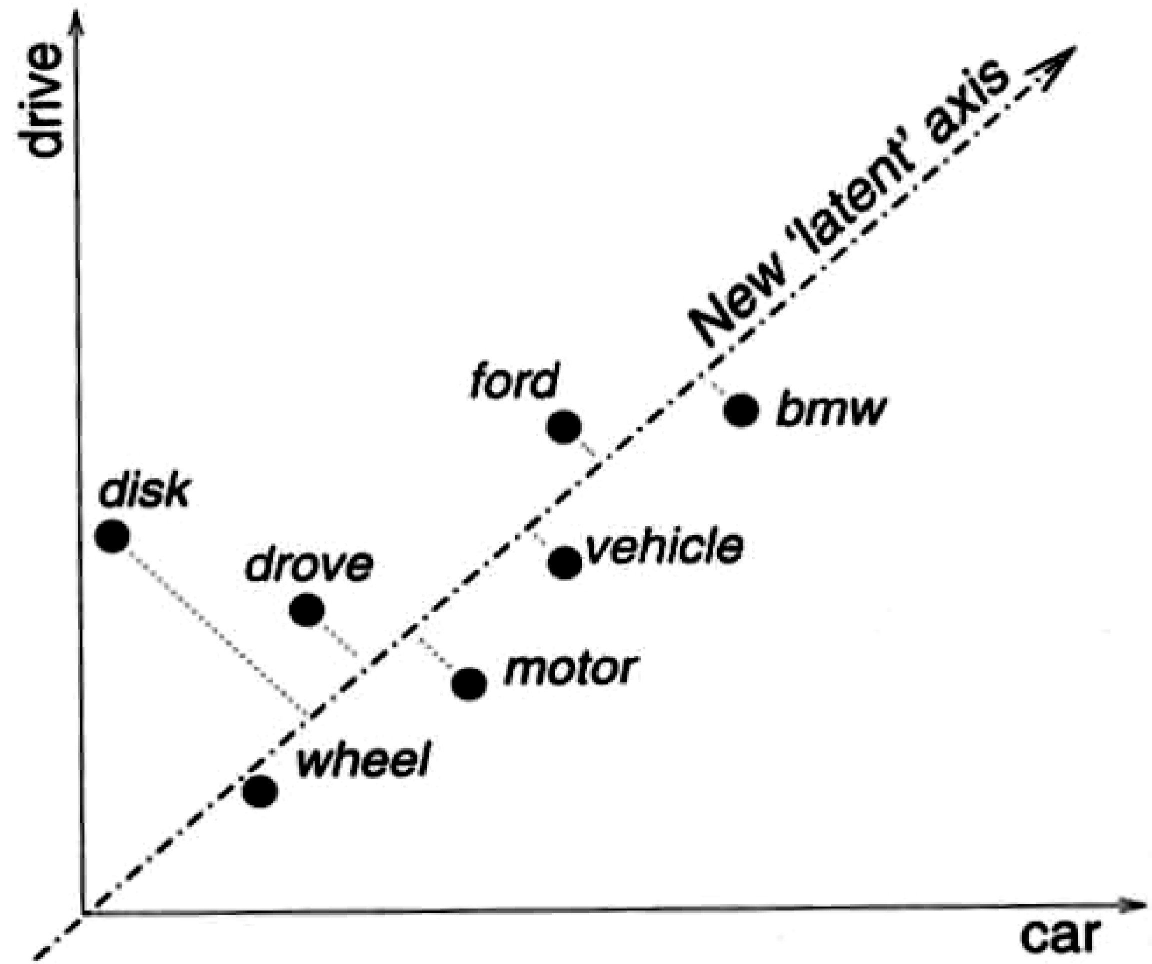
- Synonymous and other similar words: Some terms only appear singular in a document. Similar or even synonymous terms may appear much more often in a document. The user searching for one word is most likely also interested in finding documents containing the related words.
  → Vector representations after performing **L**atent **S**emantic **A**nalysis

Excurse: LSA and **S**ingular **V**alue **D**ecomposition

- The SVD can be seen as a generalization of the *spectral theorem*, which says that normal matrices can be unitarily diagonalized using a basis of *eigenvectors*, to arbitrary, not necessarily square, matrices.
- SVD: Any m×n matrix M can be factorized as the product UΣV* of three matrixes where U is an m×m unitary matrix over K (i.e., the columns of U are orthonormal), the matrix Σ is m×n with nonnegative numbers on the diagonal (called the singular values) and zeros off the diagonal, and V* denotes the conjugate transpose of V, an n×n unitary matrix over K. Such a factorization is called a singular-value decomposition of M.

Example

- Projecting words related to 2 documents (cars, driving) onto a new coordinate axis (from Widdows 2004, p. 176)



SVD: links

- http://en.wikipedia.org/wiki/Singular_value_decomposition
- http://users.pandora.be/paul.larmuseau/SVD.htm (for performing online computations)
- http://mathworld.wolfram.com/SingularValueDecomposition.html
- http://math.ut.ee/~toomas_l/linalg/
- http://www.cs.ut.ee/~toomas_l/linalg/lin2/node1.html
- http://www.cs.ut.ee/~toomas_l/linalg/lin2/node11.html#SECTION000130000000000000000

Ambiguous words and superposition

- Quantum mechanics makes a distinction between mixture and superposition of states.
- Assume that certain ambiguities can be represented by superposition:

$$w = \alpha_1\, w_1 + \alpha_2\, w_2, \text{ with } \alpha_1^{\,2} + \alpha_2^{\,2} = 1$$

- If the meanings w_1 and w_2 are orthogonal (i.e. unrelated to each other), the we can "disambiguate" w_1 and w_2 by vector negation

$$\alpha_2\, w_2 = w \text{ NOT } w_1 =_{\text{def}} w - (w \cdot w_1)\, w_1$$

Vector negation

- The vector a NOT b has the form $a - \lambda b$
- The vector a NOT b is orthogonal to b, i.e.

$$(a \text{ NOT } b) \cdot b = 0$$

→consequence in case of normalized vectors:

$$\lambda = a \cdot b$$

$$a \text{ NOT } b = a - (a \cdot b)\, b$$

Example

	Document 1	Document 2	Document 3
bank	0	0	1
bass	0.447	0.894	0
commercial	0	0.707	0.707
cream	1	0	0
guitar	1	0	0
fishermen	0	1	0
money	0	0.447	0.894

$$\textit{bass}\ \text{NOT}\ \textit{fisherman} = \begin{pmatrix} 0.447 \\ 0.894 \\ 0 \end{pmatrix} - 0.894 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.447 \\ 0 \\ 0 \end{pmatrix}$$

$$\text{After normalization: } \textit{bass}\ \text{NOT}\ \textit{fisherman} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Conclusions

- The basic for the present presentation of word vectors and searching engines is the representation of similarities and distances by the inner product representation.
- Certain ambiguities can be represented by superposition (= vector addition) of the different meaning vectors.
- The Euclidean model adopted here is not universally valid, however. For instance, the realization of colours using vector algebras has to make use of Riemannian geometry.
- Quantification is a perfectly natural idea both extending Boolean and Quantum logic, but we have not investigated it here.

Appendix: Linear algebraic proof of SVD

Let M be a rectangular matrix with complex entries. M*M is positive semidefinite, therefore Hermitian. By the spectral theorem, there exist an unitary U such that

$$U^*M^*MU = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$$

where Σ is diagonal and positive definite. Partition U appropriately so we can write

$$\begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} M^*M \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \begin{bmatrix} U_1^*M^*MU_1 & U_1^*M^*MU_2 \\ U_2^*M^*MU_1 & U_2^*M^*MU_2 \end{bmatrix} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore $U^*_1M^*MU^*_1 = \Sigma$, and $MU_2 = 0$. Define

$$W_1 = \Sigma^{-\frac{1}{2}} U_1^* M^*.$$

Then

$$W_1 M U_1 = \Sigma^{\frac{1}{2}}.$$

We see that this is almost the desired result, except that W_1 and U_1 are not unitary in general. W_1 is a partial isometry ($W_1W^*_1 = I$) while U_1 is an isometry ($U^*_1U_1 = I$). To finish the argument, one simply has to "fill out" these matrices to obtain unitaries. U_2 already does this for U_1. Similarly, one can choose W_2 such that

$$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix}$$

is unitary. Direct calculation shows

$$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} M \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \begin{bmatrix} \Sigma^{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix},$$

which is the desired result.

Notice the argument could begin with diagonalizing MM* rather than M*M (This shows directly that MM* and M*M have the same non-zero eigenvalues).