Introduction to Dynamical Systems

Herbert Wiklicky
herbert@imperial.ac.uk

Imperial College London

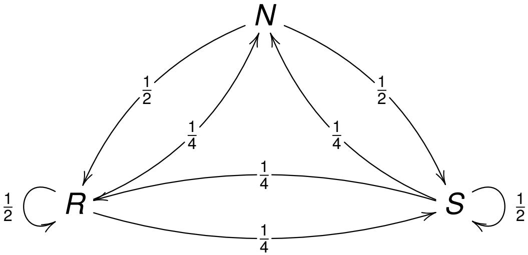
Bertinoro, June 2013

The Land of Oz

The Land of Oz is blessed with many things, but not by good weather. They never have two nice days in a row. If they have a nice day, the chance of rain or snow the next day are the same. If there is rain or snow the chances are even that the weather stays the same for the next day. If there is a change from snow or rain, only half of the time is this a change to a nice day.

The Land of Oz

From this we obtain the transition probabilities between nice (N), rainy (R) and snowy (S) days:



The Land of Oz

We can then define the following transition matrix:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

From Grinstead & Snell: *Introduction to Probability*, p406; available as GNU book on http://www.dartmouth.edu/∼chance

The Land of Oz

We can then define the following transition matrix:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

From Grinstead & Snell: *Introduction to Probability*, p406;
available as GNU book on http://www.dartmouth.edu/∼chance

Discrete Time Markov Chain

Given a finite set of states $S = \{s_1, \ldots, s_r\}$.

A discrete time Markov chain (DTMC) on S is defined via a stochastic matrix $\mathbf{P}$ as a above, i.e. an $r \times r$ (square) matrix with entries $0 \leq p_{ij} \leq 1$ and such that all row sums are equal to one, i.e.

$$\sum_j p_{ij} = 1.$$

Discrete Time Markov Processes

Let $\mathbf{P}$ be the transition matrix of a DTMC. The entry in $p_{ij}^{(n)}$ in the n-th matrix power $\mathbf{P}^n$ gives the probability that the Markov chain, starting in state s_i, will be in state s_j after exactly n steps.

At any time step we can describe the probabilities of being in a certain state s_i by a probability u_i. These probabilities define a probability distribution, i.e. a row vector

$$\mathbf{u} = (u_1, u_2, \cdots, u_r)$$

such that $0 \leq u_i \leq 1$ and $\sum_i u_i = 1$.

For any stochastic matrix $\mathbf{P}$ and probability distribution $\mathbf{u}$ the multiplication $\mathbf{uP}$ is again a probability distribution.

Discrete Time Markov Processes

Let **P** be the transition matrix of a DTMC. The entry in $p_{ij}^{(n)}$ in the n-th matrix power $\mathbf{P}^n$ gives the probability that the Markov chain, starting in state s_i, will be in state s_j after exactly n steps.

At any time step we can describe the probabilities of being in a certain state s_i by a probability u_i. These probabilities define a probability distribution, i.e. a row vector

$$\mathbf{u} = (u_1, u_2, \cdots, u_r)$$

such that $0 \leq u_i \leq 1$ and $\sum_i u_i = 1$.

For any stochastic matrix **P** and probability distribution **u** the multiplication **uP** is again a probability distribution.

Discrete Time Markov Processes

Let **P** be the transition matrix of a DTMC. The entry in $p_{ij}^{(n)}$ in the n-th matrix power $\mathbf{P}^n$ gives the probability that the Markov chain, starting in state s_i, will be in state s_j after exactly n steps.

At any time step we can describe the probabilities of being in a certain state s_i by a probability u_i. These probabilities define a probability distribution, i.e. a row vector

$$\mathbf{u} = (u_1, u_2, \cdots, u_r)$$

such that $0 \leq u_i \leq 1$ and $\sum_i u_i = 1$.

For any stochastic matrix **P** and probability distribution **u** the multiplication **uP** is again a probability distribution.

The Land of Oz

Consider the initial probability distributions $\mathbf{u} = (0, 1, 0)$ and $\mathbf{v} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ in the Oz Example. The vector $\mathbf{u}$ describes a situation where we are certain that we start with a nice day (N), while $\mathbf{v}$ corresponds to one where we assume the same chances of having a rainy (R), nice (N) or snowy (S) day.

The Land of Oz

Consider the initial probability distributions $\mathbf{u} = (0, 1, 0)$ and $\mathbf{v} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ in the Oz Example.

$$\mathbf{uP} = \left(\frac{1}{2}, 0, \frac{1}{2}\right) \quad \mathbf{uP}^2 = \left(\frac{3}{8}, \frac{1}{4}, \frac{3}{8}\right)$$

Consider the initial probability distributions $\mathbf{u} = (0, 1, 0)$ and $\mathbf{v} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ in the Oz Example.

$$\mathbf{vP} = \begin{pmatrix} 0.41667 \\ 0.16667 \\ 0.41667 \end{pmatrix}^T \quad \mathbf{vP}^2 = \begin{pmatrix} 0.39583 \\ 0.20833 \\ 0.39583 \end{pmatrix}^T \quad \mathbf{vP}^3 = \begin{pmatrix} 0.40104 \\ 0.19792 \\ 0.40104 \end{pmatrix}^T$$

$$\mathbf{vP}^4 = \begin{pmatrix} 0.39974 \\ 0.20052 \\ 0.39974 \end{pmatrix}^T \quad \cdots \quad \mathbf{vP}^{100} = \begin{pmatrix} 0.40000 \\ 0.20000 \\ 0.40000 \end{pmatrix}^T$$

Convention

Note that in the theory of Markov chains one usually is concerned with probability distributions as row vectors. Therefore, probability vectors are post-multiplied by the stochastic matrix **P** defining a Markov chain.

The usual pre-multiplication could be realised via:

$$\mathbf{P}\mathbf{u} = (\mathbf{u}^T\mathbf{P}^T)^T$$

Dynamical Systems

Introductory remarks. In dynamics we deal with physical systems whoes state at time t is completely specified by the values of n real variables

$$x_1, x_2, \ldots, x_n.$$

Accordingly the system is such that the rates of change of these variables, namely

$$dx_1/dt, dx_2/dt, \ldots, dx_n/dt,$$

merely depend upon the values of the variables themselves, so that the laws of motion can be expresses by means of n differential equations of the first order

$$dx_i/dt = X_i(x_1, x_2, \ldots, x_n) \quad (i = 1, \ldots, n).$$

George D. Birkhoff. *Dynamical Systems*, volume 9 of *Colloquium Publications*. AMS, 1927.

Dynamical Systems (Bhatia/Szegö 1970)

...the symbol X denotes a metric space [...] and R stands for the set of real numbers.

1.1 Definition. A dynamical system on X is a triplet (X, R, π), where π is a map from the product space $X \times R$ into the space X satisfying the following axioms:

1.1.1 $\pi(x, 0) = x$ for every $x \in X$ (identity axiom),

1.1.2 $\pi(\pi(x, t_1), t_2) = \pi(x, t_1 + t_2)$ for every $x \in X$ and $t_1, t_2 \in R$ (group axiom),

1.1.3 π is continuous (continuity axiom).

Nam Parshad Bhatia and Giorgio P. Szegö. *Stability Theory of Dynamical Systems*, volume 161 of *Grundlehren der mathematischen Wissenschaften*.

Definition

A general dynamical system is a triple (G, π, X) with $(G, \cdot)$ a group, X any set and $\pi : G \times X \to X$ with:

Identity Axiom

$$\pi(e, x) = x$$

for all $x \in X$ and $e \in G$ unit.

Homomorphism Axiom

$$\pi(g, \pi(h, x)) = \pi(gh, x)$$

for all $x \in X$ and $g, h \in G$.

Elements of a General Dynamical System

A general dynamical systems is made up of three ingredients:

Phase Space: a set X where “things happen”. This can have additional structure (topology, norm, etc.)

Phase Group: the group G which allows us to “combine” the partial dynamics to obtain a global picture.

Group Action: the way in which the dynamics of the group G is implement on the phase space X.

Elements of a General Dynamical System

A general dynamical systems is made up of three ingredients:

Phase Space: a set X where "things happen". This can have additional structure (topology, norm, etc.)

Phase Group: the group G which allows us to "combine" the partial dynamics to obtain a global picture.

Group Action: the way in which the dynamics of the group G is implement on the phase space X.

Elements of a General Dynamical System

A general dynamical systems is made up of three ingredients:

Phase Space: a set X where "things happen". This can have additional structure (topology, norm, etc.)

Phase Group: the group G which allows us to "combine" the partial dynamics to obtain a global picture.

Group Action: the way in which the dynamics of the group G is implement on the phase space X.

Variations of Dynamical System

Typical phase groups are $\mathbb{Z}$ (integers) or $\mathbb{R}$ (reals) for so called discrete time or continuous time models.

To investigate, for example, symmetries of the phase space it is also often the case that one considers so-called Lie Groups as transformation groups.

Typically we will request that the group action preserves the structure of the phase space, i.e. $\pi(g, .)$ is a structure preserving morphism on X for all $g \in G$.

An option is to drop invertability to get one-sided dynamical systems by taking G to be a semi-group (e.g. the naturals $\mathbb{N}$).

Variations of Dynamical System

Typical phase groups are $\mathbb{Z}$ (integers) or $\mathbb{R}$ (reals) for so called discrete time or continuous time models.

To investigate, for example, symmetries of the phase space it is also often the case that one considers so-called Lie Groups as transformation groups.

Typically we will request that the group action preserves the structure of the phase space, i.e. $\pi(g, .)$ is a structure preserving morphism on X for all $g \in G$.

An option is to drop invertability to get one-sided dynamical systems by taking G to be a semi-group (e.g. the naturals $\mathbb{N}$).

Variations of Dynamical System

Typical phase groups are $\mathbb{Z}$ (integers) or $\mathbb{R}$ (reals) for so called discrete time or continuous time models.

To investigate, for example, symmetries of the phase space it is also often the case that one considers so-called Lie Groups as transformation groups.

Typically we will request that the group action preserves the structure of the phase space, i.e. $\pi(g, .)$ is a structure preserving morphism on X for all $g \in G$.

An option is to drop invertability to get one-sided dynamical systems by taking G to be a semi-group (e.g. the naturals $\mathbb{N}$).

Variations of Dynamical System

Typical phase groups are $\mathbb{Z}$ (integers) or $\mathbb{R}$ (reals) for so called discrete time or continuous time models.

To investigate, for example, symmetries of the phase space it is also often the case that one considers so-called Lie Groups as transformation groups.

Typically we will request that the group action preserves the structure of the phase space, i.e. $\pi(g, .)$ is a structure preserving morphism on X for all $g \in G$.

An option is to drop invertability to get one-sided dynamical systems by taking G to be a semi-group (e.g. the naturals $\mathbb{N}$).

(Phase) Groups

Definition

A group G is a set with two maps (product and inverse)

$$. \cdot . : G \times G \to G \text{ and } .^{-1} : G \to G$$

fulfilling:

(i) $(xy)z = x(yz)$ for all $x, y, z \in G$ associativity axiom.

(ii) $\exists e \in G$ such that $ex = xe = x$ f or all $x \in G$ identity axiom.

(iii) $x^{-1}x = xx^{-1} = e$ for all $x \in G$ inverse axiom.

Here the group is presented multiplicatively, some groups are represented additively, e.g. $(\mathbb{Z}, +)$ and $(\mathbb{R}, +)$.

(Phase) Groups

Definition

A group G is a set with two maps (product and inverse)

$$. \cdot . : G \times G \to G \text{ and } .^{-1} : G \to G$$

fulfilling:

(i) $(xy)z = x(yz)$ for all $x, y, z \in G$ associativity axiom.

(ii) $\exists e \in G$ such that $ex = xe = x$ f or all $x \in G$ identity axiom.

(iii) $x^{-1}x = xx^{-1} = e$ for all $x \in G$ inverse axiom.

Here the group is presented multiplicatively, some groups are represented additively, e.g. $(\mathbb{Z}, +)$ and $(\mathbb{R}, +)$.

G-Spaces

Definition

Let G be a group. A G-Space is a set S and a map $\tau : G \times S \to S$ so that

$$\tau(e, s) = s \quad \text{all } s \in S$$

and

$$\tau(g, \tau(h, s)) = \tau(gh, s)$$

for all $g, h \in G$ and $s \in S$. τ is also called an action of G on S.

We write $\tau_g(s) = \tau(g, s)$ so $\tau_g : S \to S$ and we have $\tau_g \tau_h = \tau_{gh}$ as well as $\tau_g \tau_{g^{-1}} = \tau_{g^{-1}} \tau_g = \tau_e = \text{id}$, see e.g. [3, I.2]

Phase Spaces

There are many choices for the phase space of a dynamical system, among them we could mention:

Topological Spaces and require that $\pi(g, .)$ a homeomorphism.

Measurable Spaces with $\pi(g, .)$ to be measure preserving.

Vector Spaces like $\mathbb{R}^n$ with π a linear map or operator.

Strings of Symbols in an alphabet Σ as in Symbolic Dynamics.

Differentiable Manifolds as, e.g., in Classical Mechanics.

Topological Vector Spaces, Toplogical Groups, Lie Groups, ...

Phase Spaces

There are many choices for the phase space of a dynamical system, among them we could mention:

Topological Spaces and require that $\pi(g, .)$ a homeomorphism.

Measurable Spaces with $\pi(g, .)$ to be measure preserving.

Vector Spaces like $\mathbb{R}^n$ with π a linear map or operator.

Strings of Symbols in an alphabet Σ as in Symbolic Dynamics.

Differentiable Manifolds as, e.g., in Classical Mechanics.

Topological Vector Spaces, Toplogical Groups, Lie Groups, . . .

Phase Spaces

There are many choices for the phase space of a dynamical system, among them we could mention:

Topological Spaces and require that $\pi(g, .)$ a homeomorphism.
Measurable Spaces with $\pi(g, .)$ to be measure preserving.
Vector Spaces like $\mathbb{R}^n$ with π a linear map or operator.
Strings of Symbols in an alphabet Σ as in Symbolic Dynamics.
Differentiable Manifolds as, e.g., in Classical Mechanics.
Topological Vector Spaces, Toplogical Groups, Lie Groups, ...

Phase Spaces

There are many choices for the phase space of a dynamical system, among them we could mention:

Topological Spaces and require that $\pi(g, .)$ a homeomorphism.

Measurable Spaces with $\pi(g, .)$ to be measure preserving.

Vector Spaces like $\mathbb{R}^n$ with π a linear map or operator.

Strings of Symbols in an alphabet Σ as in Symbolic Dynamics.

Differentiable Manifolds as, e.g., in Classical Mechanics.

Topological Vector Spaces, Toplogical Groups, Lie Groups, ...

Phase Spaces

There are many choices for the phase space of a dynamical system, among them we could mention:

Topological Spaces and require that $\pi(g, .)$ a homeomorphism.

Measurable Spaces with $\pi(g, .)$ to be measure preserving.

Vector Spaces like $\mathbb{R}^n$ with π a linear map or operator.

Strings of Symbols in an alphabet Σ as in Symbolic Dynamics.

Differentiable Manifolds as, e.g., in Classical Mechanics.

Topological Vector Spaces, Toplogical Groups, Lie Groups, ...

Phase Spaces

There are many choices for the phase space of a dynamical system, among them we could mention:

Topological Spaces and require that $\pi(g,.)$ a homeomorphism.

Measurable Spaces with $\pi(g,.)$ to be measure preserving.

Vector Spaces like $\mathbb{R}^n$ with π a linear map or operator.

Strings of Symbols in an alphabet Σ as in Symbolic Dynamics.

Differentiable Manifolds as, e.g., in Classical Mechanics.

Topological Vector Spaces, Toplogical Groups, Lie Groups, ...

Phase Spaces

There are many choices for the phase space of a dynamical system, among them we could mention:

Topological Spaces and require that $\pi(g,.)$ a homeomorphism.

Measurable Spaces with $\pi(g,.)$ to be measure preserving.

Vector Spaces like $\mathbb{R}^n$ with π a linear map or operator.

Strings of Symbols in an alphabet Σ as in Symbolic Dynamics.

Differentiable Manifolds as, e.g., in Classical Mechanics.

Topological Vector Spaces, Toplogical Groups, Lie Groups, ...

Group Action

Definition

Let (G, π, X) be a dynamical system. The orbit of a point $x \in X$ is given by

$$O_G(x) = \{\pi(g, x) \mid g \in G\}.$$

Definition

Let (G, π, X) be a dynamical system. The group action π is

transitive iff

$$\forall x, x' \in X : \; O_G(x) = O_G(x').$$

faithful iff

$$g \mapsto \pi(g, x) \text{ is injective.}$$

free iff

$$\forall x \in X : \; g \mapsto \pi(g, x) \text{ is injective.}$$

Group Action

Definition

Let (G, π, X) be a dynamical system. The orbit of a point $x \in X$ is given by

$$O_G(x) = \{\pi(g, x) \mid g \in G\}.$$

Definition

Let (G, π, X) be a dynamical system. The group action π is

transitive iff

$$\forall x, x' \in X : \; O_G(x) = O_G(x').$$

faithful iff

$$g \mapsto \pi(g, x) \text{ is injective.}$$

free iff

$$\forall x \in X : \; g \mapsto \pi(g, x) \text{ is injective.}$$

Group Action

Definition

Let (G, π, X) be a dynamical system. The orbit of a point $x \in X$ is given by

$$O_G(x) = \{\pi(g, x) \mid g \in G\}.$$

Definition

Let (G, π, X) be a dynamical system. The group action π is

transitive iff

$$\forall x, x' \in X : \; O_G(x) = O_G(x').$$

faithful iff

$$g \mapsto \pi(g, x) \text{ is injective.}$$

free iff

$$\forall x \in X : \; g \mapsto \pi(g, x) \text{ is injective.}$$

Group Action

Definition

Let (G, π, X) be a dynamical system. The orbit of a point $x \in X$ is given by

$$O_G(x) = \{\pi(g, x) \mid g \in G\}.$$

Definition

Let (G, π, X) be a dynamical system. The group action π is

transitive iff

$$\forall x, x' \in X : \; O_G(x) = O_G(x').$$

faithful iff

$$g \mapsto \pi(g, x) \text{ is injective.}$$

free iff

$$\forall x \in X : \; g \mapsto \pi(g, x) \text{ is injective.}$$

Group Action

Definition

Let (G, π, X) be a dynamical system. The orbit of a point $x \in X$ is given by

$$O_G(x) = \{\pi(g, x) \mid g \in G\}.$$

Definition

Let (G, π, X) be a dynamical system. The group action π is

transitive iff

$$\forall x, x' \in X : \; O_G(x) = O_G(x').$$

faithful iff

$$g \mapsto \pi(g, x) \text{ is injective.}$$

free iff

$$\forall x \in X : \; g \mapsto \pi(g, x) \text{ is injective.}$$

Elements of Ergodic Theory

Topological Dynamical System

Definition

A topological dynamical system is a dynamical system (G, π, X) with the elements:

G is a topological group, i.e. $\cdot$ is continuous,

X is a topological space,

and π fulfills the

Continuity Axiom:

$$\pi : G \times X \to X \text{ is continuous.}$$

Toplological Spaces

Definition

A topological space is a set X together with a family of sub-sets $\tau \subseteq \mathcal{P}(X)$, the topology (of open sets), iff

1. $\emptyset \in \tau$ and $X \in \tau$.
2. $\bigcap_{i=0}^{n} O_i \in \tau$ for $O_i \in \tau$ (finite).
3. $\bigcup_{i \in I} O_i \in \tau$ for $O_i \in \tau$ (arbitrary).

The sets $O \in \tau$ are called open sets. The complements $A = X \setminus O$ of open sets are closed sets.

Definition

A metric space is a set X and a real-valued function $d(.,.)$, a metric, on $X \times X$ which satisfies:

1. $d(x,y) \geq 0$
2. $d(x,y) = 0 \iff x = y$
3. $d(x,y) = d(y,x)$
4. $d(x,z) \leq d(x,y) + d(y,z)$

Complete Metric Spaces

In a metric space we can define a basis for the topology open sets via open balls, i.e. sets $B(x, \varepsilon) = \{x' \mid d(x, x') < \varepsilon\}$, i.e. open sets are those which are unions of open balls.

Given a sequence $(x_i)_{i \in \mathbb{N}}$ of points in a topological space. We say that it converges if there exists $x = \lim x_i$ such that for all neighbourhoods $U(x)$ of x there $\exists N$ s.t. for $n > N : x_n \in U(x)$.

A sequence of elements $(x_i)_{i \in \mathbb{N}}$ in a metric space (X, d) is called a Cauchy sequence if

$$\forall \varepsilon > 0 \; \exists N : n, m \geq N \Rightarrow d(x_n, x_m) < \varepsilon.$$

A metric space (X, d) in which all Cauchy sequences converge is called complete (metric) space.

Complete Metric Spaces

In a metric space we can define a basis for the topology open sets via open balls, i.e. sets $B(x, \varepsilon) = \{x' \mid d(x, x') < \varepsilon\}$, i.e. open sets are those which are unions of open balls.

Given a sequence $(x_i)_{i \in \mathbb{N}}$ of points in a topological space. We say that it converges if there exists $x = \lim x_i$ such that for all neighbourhoods $U(x)$ of x there $\exists N$ s.t. for $n > N : x_n \in U(x)$.

A sequence of elements $(x_i)_{i \in \mathbb{N}}$ in a metric space (X, d) is called a Cauchy sequence if

$$\forall \varepsilon > 0 \; \exists N : n, m \geq N \Rightarrow d(x_n, x_m) < \varepsilon.$$

A metric space (X, d) in which all Cauchy sequences converge is called complete (metric) space.

Complete Metric Spaces

In a metric space we can define a basis for the topology open sets via open balls, i.e. sets $B(x, \varepsilon) = \{x' \mid d(x, x') < \varepsilon\}$, i.e. open sets are those which are unions of open balls.

Given a sequence $(x_i)_{i\in\mathbb{N}}$ of points in a topological space. We say that it converges if there exists $x = \lim x_i$ such that for all neighbourhoods $U(x)$ of x there $\exists N$ s.t. for $n > N : x_n \in U(x)$.

A sequence of elements $(x_i)_{i\in\mathbb{N}}$ in a metric space (X, d) is called a Cauchy sequence if

$$\forall\varepsilon > 0 \;\exists N : n, m \geq N \Rightarrow d(x_n, x_m) < \varepsilon.$$

A metric space (X, d) in which all Cauchy sequences converge is called complete (metric) space.

Complete Metric Spaces

In a metric space we can define a basis for the topology open sets via open balls, i.e. sets $B(x, \varepsilon) = \{x' \mid d(x, x') < \varepsilon\}$, i.e. open sets are those which are unions of open balls.

Given a sequence $(x_i)_{i \in \mathbb{N}}$ of points in a topological space. We say that it converges if there exists $x = \lim x_i$ such that for all neighbourhoods $U(x)$ of x there $\exists N$ s.t. for $n > N : x_n \in U(x)$.

A sequence of elements $(x_i)_{i \in \mathbb{N}}$ in a metric space (X, d) is called a Cauchy sequence if

$$\forall \varepsilon > 0 \; \exists N : n, m \geq N \Rightarrow d(x_n, x_m) < \varepsilon.$$

A metric space (X, d) in which all Cauchy sequences converge is called complete (metric) space.

Continuous Functions

Definition

A function $\mathbf{T} : X \to X'$ between two topological spaces (X, τ) and (X', τ') is called

continuous iff

$$\forall O \in \tau' : \mathbf{T}^{-1}(O) \in \tau.$$

homeomorph iff

$$\mathbf{T} \text{ is a bijection, and } \mathbf{T} \text{ and } \mathbf{T}^{-1} \text{ are continuous.}$$

Continuous functions preserve limits, i.e. $\lim \mathbf{T}(x_i) = \mathbf{T}(\lim(x_i))$.

Continuous Functions

Definition

A function $\mathbf{T} : X \to X'$ between two topological spaces (X, τ) and (X', τ') is called

continuous iff

$$\forall O \in \tau' : \mathbf{T}^{-1}(O) \in \tau.$$

homeomorph iff

$$\mathbf{T} \text{ is a bijection, and } \mathbf{T} \text{ and } \mathbf{T}^{-1} \text{ are continuous.}$$

Continuous functions preserve limits, i.e. $\lim \mathbf{T}(x_i) = \mathbf{T}(\lim(x_i))$.

Continuous Functions

Definition

A function $\mathbf{T} : X \to X'$ between two topological spaces (X, τ) and (X', τ') is called

continuous iff

$$\forall O \in \tau' : \ \mathbf{T}^{-1}(O) \in \tau.$$

homeomorph iff

$\mathbf{T}$ is a bijection, and $\mathbf{T}$ and $\mathbf{T}^{-1}$ are continuous.

Continuous functions preserve limits, i.e. $\lim \mathbf{T}(x_i) = \mathbf{T}(\lim(x_i))$.

Continuous Functions

Definition

A function $\mathbf{T} : X \to X'$ between two topological spaces (X, τ) and (X', τ') is called

continuous iff

$$\forall O \in \tau' : \mathbf{T}^{-1}(O) \in \tau.$$

homeomorph iff

$$\mathbf{T} \text{ is a bijection, and } \mathbf{T} \text{ and } \mathbf{T}^{-1} \text{ are continuous.}$$

Continuous functions preserve limits, i.e. $\lim \mathbf{T}(x_i) = \mathbf{T}(\lim(x_i))$.

Measure Theoretical Dynamical System

Definition

A measure theoretic dynamical system is a dynamical system (G, π, X) with

G is a measur(abl)e space,

X is a measur(abl)e space,

and π fulfills the

Measurability Axiom:

$\pi : G \times X \to X$ is measurable.
$\pi(g, .) : X \to X$ is measure preserving $\forall g \in G$.

One can define a product measure on $G \times X$ in order to make sense of the first condition.

Measure Theoretical Dynamical System

Definition

A measure theoretic dynamical system is a dynamical system (G, π, X) with

G is a measur(abl)e space,

X is a measur(abl)e space,

and π fulfills the

Measurability Axiom:

$\pi : G \times X \to X$ is measurable.
$\pi(g, .) : X \to X$ is measure preserving $\forall g \in G$.

One can define a product measure on $G \times X$ in order to make sense of the first condition.

Measureable Spaces

Definition

Given any set X. A family σ of sub-sets $\sigma \subseteq \mathcal{P}(X)$ is called a σ-algebra iff

1. $\emptyset \in \sigma$ and $X \in \sigma$.
2. $\bigcap_{i=0}^{\infty} S_i \in \sigma$ for $S_i \in \sigma$ (countable).
3. $X \setminus S \in \sigma$ for $S \in \sigma$.

We say that (X, σ) is a measurable space, and $S \in \sigma$ are measurable sets.

By de Morgan we have also: $\bigcup_{i=0}^{\infty} S_i \in \sigma$ for $S_i \in \sigma$ (countable).

Measureable Spaces

Definition

Given any set X. A family σ of sub-sets $\sigma \subseteq \mathcal{P}(X)$ is called a σ-algebra iff

1. $\emptyset \in \sigma$ and $X \in \sigma$.
2. $\bigcap_{i=0}^{\infty} S_i \in \sigma$ for $S_i \in \sigma$ (countable).
3. $X \setminus S \in \sigma$ for $S \in \sigma$.

We say that (X, σ) is a measurable space, and $S \in \sigma$ are measurable sets.

By de Morgan we have also: $\bigcup_{i=0}^{\infty} S_i \in \sigma$ for $S_i \in \sigma$ (countable).

Measures and Measurable Functions

Definition

Given a measurable space (X, σ) then $\mu : \sigma \to \mathbb{R}^+$ is a (finite) measure if

1. $\mu(\emptyset) = 0$ (for $\mu(X) = 1$ we have a probability measure).
2. $\mu(\bigcup_{i=0}^{\infty} S_i) = \sum_{i=0}^{\infty} \mu(S_i)$ for $S_i \in \sigma$ with $S_i \cap S_j = \emptyset$ for $i \neq j$.

Definition

A function $\mathbf{T} : X \to X'$ between two measure spaces spaces (X, σ, μ) and (X', τ', μ') is called

measurable iff

$$\forall S \in \sigma' : \mathbf{T}^{-1}(S) \in \sigma.$$

measure preserving iff $\forall S \in \sigma'$ also $\mu'(S') = \mu(\mathbf{T}^{-1}(S))$.

Topological Mixing Notions

Definition

Given a topological dynamical system (G, π, X).
We say that (G, π, X) is topologically transitive if

$$\exists x \in X : O_G(x) \text{ is dense in } X.$$

We say that (G, π, X) is (topologically) minimal if

$$\forall x \in X : O_G(x) \text{ is dense in } X.$$

Definition

A discrete topological dynamical system $(\mathbf{T}, X)$ is called topologically (strong) mixing if

$$\forall U, V \subseteq X \text{ open and non-empty } \exists N : \forall n > N : \mathbf{T}^n(U) \cap V \neq \emptyset.$$

Topological Mixing Notions

Definition

Given a topological dynamical system (G, π, X).
We say that (G, π, X) is topologically transitive if

$$\exists x \in X : O_G(x) \text{ is dense in } X.$$

We say that (G, π, X) is (topologically) minimal if

$$\forall x \in X : O_G(x) \text{ is dense in } X.$$

Definition

A discrete topological dynamical system $(\mathbf{T}, X)$ is called topologically (strong) mixing if

$$\forall U, V \subseteq X \text{ open and non-empty } \exists N : \forall n > N : \mathbf{T}^n(U) \cap V \neq \emptyset.$$

Topologically Transitive

Theorem

Given a discrete topological dynamical system $(\mathbf{T}, X)$ *on a compact metric space* X *then the following conditions are equivalent:*

1. $\forall x \in X : O_{\mathbf{T}}(x)$ *is dense in* X *(topologically transitive).*
2. $\forall C \subseteq X$ *closed with* $\mathbf{T}(C) = C \Rightarrow C = X$ *or* $C = \emptyset$.
3. $\forall O \subseteq X$ *open with* $\mathbf{T}(O) = O \Rightarrow O = X$ *or* $O = \emptyset$.
4. $\forall O \subseteq X$ *open and non-empty, then* $\bigcup_{n=-\infty}^{\infty} \mathbf{T}^n(O) = X$.

Measure Theoretic Mixing Notions

Definition

Given a discrete measure theoretic dynamical system $(\mathbf{T}, X)$. We say $(\mathbf{T}, X)$ is measure theoretic transitive or ergodic if

$$\forall S \subseteq X \text{ measurable with } \mathbf{T}(S) = S \Rightarrow \mu(S) = 0 \text{ or } \mu(S) = 1.$$

Definition

A discrete measure theoretic dynamical system $(\mathbf{T}, X)$ is called strong mixing if

$$\forall S_1, S_2 \subseteq X \text{ measurable } \lim_{n\to\infty} \mu(\mathbf{T}^{-n}(S_1) \cap S_2) = \mu(S_1)\mu(S_2).$$

Measure Theoretic Mixing Notions

Definition

Given a discrete measure theoretic dynamical system $(\mathbf{T}, X)$. We say $(\mathbf{T}, X)$ is measure theoretic transitive or ergodic if

$$\forall S \subseteq X \text{ measurable with } \mathbf{T}(S) = S \Rightarrow \mu(S) = 0 \text{ or } \mu(S) = 1.$$

Definition

A discrete measure theoretic dynamical system $(\mathbf{T}, X)$ is called strong mixing if

$$\forall S_1, S_2 \subseteq X \text{ measurable } \lim_{n\to\infty} \mu(\mathbf{T}^{-n}(S_1) \cap S_2) = \mu(S_1)\mu(S_2).$$

Ergodicity

Theorem

Given a discrete measure theoretic dynamical system $(\mathbf{T}, X)$ *with* $\mathbf{T}$ *measure preserving. Then the following conditions are equivalent (with ergodic):*

1. $\forall S \subseteq X$ *measurable* $\mathbf{T}(S) = S \Rightarrow \mu(S) = 0$ *or* $\mu(S) = 1$.
2. $\forall S \subseteq X$ *measurable and* $\mu(\mathbf{T}^{-1}(S) \bigtriangleup S) = 0$ $\Rightarrow \mu(S) = 0$ *or* $\mu(S) = 1$.
3. $\forall S \subseteq X$ *measurable and* $\mu(S) > 0 \Rightarrow \mu(\bigcup_{n=-\infty}^{\infty} \mathbf{T}^{-n}(S)) = 1$.
4. $\forall S_1, S_2 \subseteq X$ *measurable and* $\mu(S_1) > 0 < \mu(S_2)$ $\Rightarrow \exists n \in \mathbb{N}$ *such that* $\mu(\mathbf{T}^{-n}(S_1) \cap S_2) = 0$.

Given a discrete measure theoretic dynamical system $(\mathbf{T}, X)$ and a function (i.e. a random variable) $f : X \to \mathbb{R}$.

The phase average of f is defined as $\mu(f) = \int_X f(x)dx$. The time average of f is defined as $f^*(x) = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(\mathbf{T}^t(x)))dt$.

Theorem (Birkhoff)

Given a discrete measure theoretic dynamical system $(\mathbf{T}, X)$, with $\mathbf{T}$ measure preserving, and a function $f : X \to \mathbb{R}$ with $f \in L^1(X, \mu)$ then the following holds:

$$(\mathbf{T}, X) \text{ is ergodic } \Leftrightarrow \mu(f) = f^*(x) \quad \mu\text{-almost everywhere.}$$

Ergodic Theorem

Given a discrete measure theoretic dynamical system $(\mathbf{T}, X)$ and a function (i.e. a random variable) $f : X \to \mathbb{R}$.

The phase average of f is defined as $\mu(f) = \int_X f(x)dx$. The time average of f is defined as $f^*(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(\mathbf{T}^t(x))dt$.

Theorem (Birkhoff)

Given a discrete measure theoretic dynamical system $(\mathbf{T}, X)$, with $\mathbf{T}$ measure preserving, and a function $f : X \to \mathbb{R}$ with $f \in L^1(X, \mu)$ then the following holds:

$$(\mathbf{T}, X) \text{ is ergodic } \Leftrightarrow \mu(f) = f^*(x) \quad \mu\text{-almost everywhere.}$$

Elements of Linear Dynamical Systems

Linear Dynamical System

Definition

A linear dynamical system is a dynamical system (G, π, X) with

G is a group (typically $G = \mathbb{Z}$),

X is a vector space

and π fulfils the

Linearity Axiom:

$$\pi(g, .) : X \to X \text{ is linear } \forall g \in G.$$

Many versions of linear dynamical systems play an important role in control theory investigating e.g. feed back loops etc.

Linear Dynamical System

Definition

A linear dynamical system is a dynamical system (G, π, X) with

- G is a group (typically $G = \mathbb{Z}$),
- X is a vector space

and π fulfils the

Linearity Axiom:

$$\pi(g, .) : X \to X \text{ is linear } \forall g \in G.$$

Many versions of linear dynamical systems play an important role in control theory investigating e.g. feed back loops etc.

Definition

A Vector Space (over a field $\mathbb{K}$, e.g. $\mathbb{R}$ or $\mathbb{C}$) is a set $\mathcal{V}$ together with two operations:

$$\text{Scalar Multiplication} \quad .\cdot. : \mathbb{K} \times \mathcal{V} \mapsto \mathcal{V}$$

$$\text{Vector Addition} \quad .+. : \mathcal{V} \times \mathcal{V} \mapsto \mathcal{V}$$

such that ($\forall x, y, z \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{K}$):

1. $x + (y + z) = (x + y) + z$
2. $x + y = y + x$
3. $\exists o : x + o = x$
4. $\exists -x : x + (-x) = o$

1. $\alpha(x + y) = \alpha x + \alpha y$
2. $(\alpha + \beta)x = \alpha x + \beta x$
3. $(\alpha\beta)x = \alpha(\beta x)$
4. $1x = x$ $(1 \in \mathbb{K})$

Tuple Spaces

Theorem

All finite dimensional vector spaces are isomorphic to the (finite) Cartesian product of the underlying field $\mathbb{K}^n$ (i.e. $\mathbb{R}^n$ or $\mathbb{C}^m$).

Finite dimensional vectors can always be represented via their coordinates with respect to a given base, e.g.

$$\begin{aligned} x &= (x_1, x_2, x_3, \ldots, x_n) \\ y &= (y_1, y_2, y_3, \ldots, y_n) \end{aligned}$$

Algebraic Structure

$$\begin{aligned} \alpha x &= (\alpha x_1, \alpha x_2, \alpha x_3, \ldots, \alpha x_n) \\ x + y &= (x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots, x_n + y_n) \end{aligned}$$

Linear Operators

Definition

A map $\mathbf{T} : \mathcal{V} \to \mathcal{W}$ between two vector spaces $\mathcal{V}$ and $\mathcal{W}$ is called a linear map iff

1. $\mathbf{T}(x + y) = \mathbf{T}(x) + \mathbf{T}(y)$ and
2. $\mathbf{T}(\alpha x) = \alpha \mathbf{T}(x)$

for all $x, y \in \mathcal{V}$ and all $\alpha \in \mathbb{K}$ (e.g. $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$).

The set of all linear maps between $\mathcal{V}$ and $\mathcal{W}$ is denoted $\mathcal{L}(\mathcal{V}, \mathcal{W})$. For $\mathcal{V} = \mathcal{W}$ we talk about a linear operator on $\mathcal{V}$.

On normed vector spaces the continuous or equivalently bounded linear operators are of particular interest, i.e.

$$\mathcal{B}(\mathcal{V}) = \{\mathbf{T} \mid \|\mathbf{T}\| = \sup_{x \in \mathcal{V}} \frac{\|\mathbf{T}(x)\|}{\|x\|} < \infty\} \subseteq \mathcal{L}(\mathcal{V}) = \mathcal{L}(\mathcal{V}, \mathcal{V}).$$

Linear Operators

Definition

A map $\mathbf{T} : \mathcal{V} \to \mathcal{W}$ between two vector spaces $\mathcal{V}$ and $\mathcal{W}$ is called a linear map iff

1. $\mathbf{T}(x + y) = \mathbf{T}(x) + \mathbf{T}(y)$ and
2. $\mathbf{T}(\alpha x) = \alpha \mathbf{T}(x)$

for all $x, y \in \mathcal{V}$ and all $\alpha \in \mathbb{K}$ (e.g. $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$).

The set of all linear maps between $\mathcal{V}$ and $\mathcal{W}$ is denoted $\mathcal{L}(\mathcal{V}, \mathcal{W})$. For $\mathcal{V} = \mathcal{W}$ we talk about a linear operator on $\mathcal{V}$.

On normed vector spaces the continuous or equivalently bounded linear operators are of particular interest, i.e.

$$\mathcal{B}(\mathcal{V}) = \{\mathbf{T} \mid \|\mathbf{T}\| = \sup_{x \in \mathcal{V}} \frac{\|\mathbf{T}(x)\|}{\|x\|} < \infty\} \subseteq \mathcal{L}(\mathcal{V}) = \mathcal{L}(\mathcal{V}, \mathcal{V}).$$

Linear Operators

Definition

A map $\mathbf{T} : \mathcal{V} \to \mathcal{W}$ between two vector spaces $\mathcal{V}$ and $\mathcal{W}$ is called a linear map iff

1. $\mathbf{T}(x + y) = \mathbf{T}(x) + \mathbf{T}(y)$ and
2. $\mathbf{T}(\alpha x) = \alpha \mathbf{T}(x)$

for all $x, y \in \mathcal{V}$ and all $\alpha \in \mathbb{K}$ (e.g. $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$).

The set of all linear maps between $\mathcal{V}$ and $\mathcal{W}$ is denoted $\mathcal{L}(\mathcal{V}, \mathcal{W})$. For $\mathcal{V} = \mathcal{W}$ we talk about a linear operator on $\mathcal{V}$.

On normed vector spaces the continuous or equivalently bounded linear operators are of particular interest, i.e.

$$\mathcal{B}(\mathcal{V}) = \{\mathbf{T} \mid \|\mathbf{T}\| = \sup_{x \in \mathcal{V}} \frac{\|\mathbf{T}(x)\|}{\|x\|} < \infty\} \subseteq \mathcal{L}(\mathcal{V}) = \mathcal{L}(\mathcal{V}, \mathcal{V}).$$

Normed Spaces

Definition

A complex vector space $\mathcal{V}$ is called a normed (vector) space if there is a real valued function $\|.\|$ on $\mathcal{V}$ that satisfies ($\forall x, y \in \mathcal{V}$ and $\forall \alpha \in \mathbb{C}$):

1. $\|x\| \geq 0$
2. $\|x\| = 0 \iff x = o$
3. $\|\alpha x\| = |\alpha| \, \|x\|$
4. $\|x + y\| \leq \|x\| + \|y\|$

The function $\|.\|$ is called a norm on $\mathcal{V}$.

We have a Banach space if the topology induced by $d(x, y) = \|x - y\|$ is complete – always for finite dimensional spaces.

Normed Spaces

Definition

A complex vector space $\mathcal{V}$ is called a normed (vector) space if there is a real valued function $\|.\|$ on $\mathcal{V}$ that satisfies ($\forall x, y \in \mathcal{V}$ and $\forall \alpha \in \mathbb{C}$):

1. $\|x\| \geq 0$
2. $\|x\| = 0 \iff x = o$
3. $\|\alpha x\| = |\alpha| \, \|x\|$
4. $\|x + y\| \leq \|x\| + \|y\|$

The function $\|.\|$ is called a norm on $\mathcal{V}$.

We have a Banach space if the topology induced by $d(x, y)$ $= \|x - y\|$ is complete – always for finite dimensional spaces.

Hilbert Spaces

Definition

A complex vector space $\mathcal{H}$ is called an inner product space (or (pre-)Hilbert space) if there is a complex valued function $\langle ., . \rangle$ on $\mathcal{H} \times \mathcal{H}$ that satisfies ($\forall x, y, z \in \mathcal{H}$ and $\forall \alpha \in \mathbb{C}$):

1. $\langle x, x \rangle \geq 0$
2. $\langle x, x \rangle = 0 \iff x = o$
3. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
4. $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$
5. $\langle x, y \rangle = \overline{\langle y, x \rangle}$

The function $\langle ., . \rangle$ is called an inner product on $\mathcal{H}$.

If the topology induced by $\|x\| = \sqrt{\langle x, x \rangle}$ is complete then we have a Hilbert space – always for finite dimensional spaces.

Basis Vectors

A set of vectors x_i is said to be linearly independent iff

$$\lambda_i x_i = \sum \lambda_i x_i = 0 \quad \text{implies that} \quad \forall\, i : \lambda_i = 0$$

Two vectors in a Hilbert space are orthogonal iff $\langle x, y \rangle = 0$
An orthonormal system (base if it generates all $\mathcal{H}$) in a Hilbert space is a set of linearly independent vectors $\{b_i\}_i$ with:

$$\langle b_i, b_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{iff } i = j \\ 0 & \text{iff } i \neq j \end{cases}$$

Theorem

For a Hilbert space there exists an orthonormal basis $\{b_i\}$. The representation of each vector is unique:

$$x = \sum_i x_i b_i = \sum_i \langle x, b_i \rangle\, b_i$$

Basis Vectors

A set of vectors x_i is said to be linearly independent iff

$$\lambda_i x_i = \sum \lambda_i x_i = 0 \quad \text{implies that} \quad \forall\, i : \lambda_i = 0$$

Two vectors in a Hilbert space are orthogonal iff $\langle x, y \rangle = 0$

An orthonormal system (base if it generates all $\mathcal{H}$) in a Hilbert space is a set of linearly independent vectors $\{b_i\}_i$ with:

$$\langle b_i, b_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{iff } i = j \\ 0 & \text{iff } i \neq j \end{cases}$$

Theorem

For a Hilbert space there exists an orthonormal basis $\{b_i\}$. The representation of each vector is unique:

$$x = \sum_i x_i b_i = \sum_i \langle x, b_i \rangle\, b_i$$

Basis Vectors

A set of vectors x_i is said to be linearly independent iff

$$\lambda_i x_i = \sum \lambda_i x_i = 0 \quad \text{implies that} \quad \forall\, i : \lambda_i = 0$$

Two vectors in a Hilbert space are orthogonal iff $\langle x, y \rangle = 0$
An orthonormal system (base if it generates all $\mathcal{H}$) in a Hilbert space is a set of linearly independent vectors $\{b_i\}_i$ with:

$$\langle b_i, b_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{iff } i = j \\ 0 & \text{iff } i \neq j \end{cases}$$

Theorem

For a Hilbert space there exists an orthonormal basis $\{b_i\}$. The representation of each vector is unique:

$$x = \sum_i x_i b_i = \sum_i \langle x, b_i \rangle\, b_i$$

Basis Vectors

A set of vectors x_i is said to be linearly independent iff

$$\lambda_i x_i = \sum \lambda_i x_i = 0 \quad \text{implies that} \quad \forall\, i : \lambda_i = 0$$

Two vectors in a Hilbert space are orthogonal iff $\langle x, y \rangle = 0$
An orthonormal system (base if it generates all $\mathcal{H}$) in a Hilbert space is a set of linearly independent vectors $\{b_i\}_i$ with:

$$\langle b_i, b_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{iff } i = j \\ 0 & \text{iff } i \neq j \end{cases}$$

Theorem

For a Hilbert space there exists an orthonormal basis $\{b_i\}$. The representation of each vector is unique:

$$x = \sum_i x_i b_i = \sum_i \langle x, b_i \rangle\, b_i$$

Dual Spaces

A linear functional on a vector space $\mathcal{V}$ is a map $f : \mathcal{V} \to \mathbb{K}$ such that $f(x+y) = f(x) + f(y)$ and $f(\alpha x) = \alpha f(x)$ for all $x, y \in \mathcal{V}, \alpha \in \mathbb{K}$.

Theorem (Riesz Representation Theorem)

Every (bounded) linear functional on a Hilbert space $\mathcal{H}$ can be represented by a vector in the Hilbert space $\mathcal{H}$, such that

$$f(x) = \langle y_f | x \rangle = f_y(x)$$

The dual Hilbert space $\mathcal{H}^*$ is isomorphic to the original Hilbert space $\mathcal{H}$, e.g. for the universal Hilbert space $\ell_2(\mathbb{N})^* = \ell_2(\mathbb{N})$.

$$\ell_p(\mathbb{X}) = \left\{ (x_i)_{i\in\mathbb{X}} \mid \left(\sum_{i\in\mathbb{X}} |x_i|^2 \right)^{\frac{1}{p}} \right\}$$

Dual Spaces

A linear functional on a vector space $\mathcal{V}$ is a map $f : \mathcal{V} \to \mathbb{K}$ such that $f(x+y) = f(x) + f(y)$ and $f(\alpha x) = \alpha f(x)$ for all $x, y \in \mathcal{V}, \alpha \in \mathbb{K}$.

Theorem (Riesz Representation Theorem)

Every (bounded) linear functional on a Hilbert space $\mathcal{H}$ can be represented by a vector in the Hilbert space $\mathcal{H}$, such that

$$f(x) = \langle y_f | x \rangle = f_y(x)$$

The dual Hilbert space $\mathcal{H}^*$ is isomorphic to the original Hilbert space $\mathcal{H}$, e.g. for the universal Hilbert space $\ell_2(\mathbb{N})^* = \ell_2(\mathbb{N})$.

$$\ell_p(\mathbb{X}) = \left\{ (x_i)_{i\in\mathbb{X}} \mid \left(\sum_{i\in\mathbb{X}} |x_i|^2 \right)^{\frac{1}{p}} \right\}$$

Dual Spaces

A linear functional on a vector space $\mathcal{V}$ is a map $f : \mathcal{V} \to \mathbb{K}$ such that $f(x + y) = f(x) + f(y)$ and $f(\alpha x) = \alpha f(x)$ for all $x, y \in \mathcal{V}, \alpha \in \mathbb{K}$.

Theorem (Riesz Representation Theorem)

Every (bounded) linear functional on a Hilbert space $\mathcal{H}$ can be represented by a vector in the Hilbert space $\mathcal{H}$, such that

$$f(x) = \langle y_f | x \rangle = f_y(x)$$

The dual Hilbert space $\mathcal{H}^*$ is isomorphic to the original Hilbert space $\mathcal{H}$, e.g. for the universal Hilbert space $\ell_2(\mathbb{N})^* = \ell_2(\mathbb{N})$.

$$\ell_p(\mathbb{X}) = \left\{ (x_i)_{i \in \mathbb{X}} \mid \left(\sum_{i \in \mathbb{X}} |x_i|^2 \right)^{\frac{1}{p}} \right\}$$

Finite-Dimensional Hilbert Spaces

We represent vectors and their **transpose** using coordinates:

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}^T = (y_1, \ldots, y_n)$$

The adjoint of $\vec{x} = (x_1, \ldots, x_n)$, with $.^* = \bar{.}$ denoting complex conjugate in $\mathbb{C}$), is given by

$$\vec{x}^\dagger = \vec{x}^* = (x_1^*, \ldots, x_n^*)^T$$

The inner product is:

$$\langle \vec{y}, \vec{x} \rangle = \sum_i y_i^* x_i = \vec{y}^\dagger \vec{x}$$

Finite-Dimensional Hilbert Spaces

We represent vectors and their **transpose** using coordinates:

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}^T = (y_1, \ldots, y_n)$$

The adjoint of $\vec{x} = (x_1, \ldots, x_n)$, with $.^* = \bar{.}$ denoting complex conjugate in $\mathbb{C}$), is given by

$$\vec{x}^\dagger = \vec{x}^* = (x_1^*, \ldots, x_n^*)^T$$

The inner product is:

$$\langle \vec{y}, \vec{x} \rangle = \sum_i y_i^* x_i = \vec{y}^\dagger \vec{x}$$

Finite-Dimensional Hilbert Spaces

We represent vectors and their **transpose** using coordinates:

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}^T = (y_1, \ldots, y_n)$$

The adjoint of $\vec{x} = (x_1, \ldots, x_n)$, with $.^* = \bar{.}$ denoting complex conjugate in $\mathbb{C}$), is given by

$$\vec{x}^\dagger = \vec{x}^* = (x_1^*, \ldots, x_n^*)^T$$

The inner product is:

$$\langle \vec{y}, \vec{x} \rangle = \sum_i y_i^* x_i = \vec{y}^\dagger \vec{x}$$

Differential Equations

Discrete (Time) Dynamical Systems: Collatz

The Colltaz problem is a (one-sided) discrete time dynamical system $(\mathbf{C}, \mathbb{Z})$, which we can describe by the following transformation:

$$\mathbf{C} : \mathbb{Z} \to \mathbb{Z}$$

with

$$\mathbf{C}(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ 3 \times n + 1 & \text{otherwise.} \end{cases}$$

The unsolved question is:

Does $\exists m \in \mathbb{N}$ such that $\mathbf{C}^m(n) = 1$ for all $n \in \mathbb{N}$?

Discrete (Time) Dynamical Systems: Collatz

The Colltaz problem is a (one-sided) discrete time dynamical system $(\mathbf{C}, \mathbb{Z})$, which we can describe by the following transformation:

$$\mathbf{C} : \mathbb{Z} \to \mathbb{Z}$$

with

$$\mathbf{C}(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ 3 \times n + 1 & \text{otherwise.} \end{cases}$$

The unsolved question is:

Does $\exists m \in \mathbb{N}$ such that $\mathbf{C}^m(n) = 1$ for all $n \in \mathbb{N}$?

Continuous Dynamical Systems

A popular way to specify continuous time dynamical systems is via (ordinary) differential equations, e.g. Morris W. Hirsch, Stephen Smale, and Robert L. Devaney. *Differential Equations, Dynamical Systems and An Introduction to Chaos*. Elsevier, 2004.

The group action is interpreted as time $t \in \mathbb{R}$.

Ordinary Differential Equations

$$\begin{aligned} x_1' &= \frac{dx_1}{dt} = f_1(t, x_1, x_2, \ldots, x_n) \\ x_2' &= \frac{dx_2}{dt} = f_2(t, x_1, x_2, \ldots, x_n) \\ \ldots & \quad \ldots \quad \ldots \\ x_n' &= \frac{dx_n}{dt} = f_n(t, x_1, x_2, \ldots, x_n) \end{aligned}$$

Continuous Dynamical Systems

A popular way to specify continuous time dynamical systems is via (ordinary) differential equations, e.g. Morris W. Hirsch, Stephen Smale, and Robert L. Devaney. *Differential Equations, Dynamical Systems and An Introduction to Chaos*. Elsevier, 2004.

The group action is interpreted as time $t \in \mathbb{R}$.

Ordinary Differential Equations

$$\begin{array}{ccccc} x_1' & = & \frac{dx_1}{dt} & = & f_1(t, x_1, x_2, \ldots, x_n) \\ x_2' & = & \frac{dx_2}{dt} & = & f_2(t, x_1, x_2, \ldots, x_n) \\ \ldots & & \ldots & & \ldots \\ x_n' & = & \frac{dx_n}{dt} & = & f_n(t, x_1, x_2, \ldots, x_n) \end{array}$$

Differential

Given a function $f : \mathbb{R} \to \mathbb{R}$. We say it is differentiable at a point $t \in \mathbb{R}$ if there is a linear map $Df(t) : \mathbb{R} \to \mathbb{R}$ which approximates f at t. That is, $\forall \varepsilon > 0$ there is a neigborhood U of t such that:

$$\|f(t') - f(t) - Df(t)(t - t')\| < \varepsilon \|t - t'\| \quad \forall t' \in U$$

We also write for the differential (quotient) $Df = \frac{df}{dt}$.

We also approximate a function $f : \mathbb{R}^n \to \mathbb{R}^m$ by a linear map $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ represented by the matrix of partial derivates:

$$(Df)_{ij} = \frac{\partial f_i}{\partial t_j} \quad i = 1, \ldots, m, i = 1, \ldots, n.$$

Differential

Given a function $f : \mathbb{R} \to \mathbb{R}$. We say it is differentiable at a point $t \in \mathbb{R}$ if there is a linear map $Df(t) : \mathbb{R} \to \mathbb{R}$ which approximates f at t. That is, $\forall \varepsilon > 0$ there is a neigborhood U of t such that:

$$\|f(t') - f(t) - Df(t)(t - t')\| < \varepsilon \|t - t'\| \quad \forall t' \in U$$

We also write for the differential (quotient) $Df = \frac{df}{dt}$.

We also approximate a function $f : \mathbb{R}^n \to \mathbb{R}^m$ by a linear map $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ represented by the matrix of partial derivates:

$$(Df)_{ij} = \frac{\partial f_i}{\partial t_j} \quad i = 1, \ldots, m, i = 1, \ldots, n.$$

Euler's Number

What is Euler's e? Metafont users know, it is $e = 2.7183\ldots$
It is the unique number such that for $f(t) = e^t$ we have

$$\frac{df}{dt}(t) = f(t)$$

The exponential function is the fixed-point or eigen-function of the differential operator $\frac{d}{dt}$. One could show this via the Taylor expansion of $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ as $\frac{d}{dt}\frac{t^n}{n!} = \frac{nt^{n-1}}{n(n-1)!} = \frac{t^{n-1}}{(n-1)!}$.

The simplest differential equation one can think of is perhaps:

$$x'(t) = \frac{dx}{dt}(t) = ax(t)$$

The solution is $x(t) = ke^{at} = k\exp(at)$ for some constant k (can be determined via an initial/boundary condition, e.g. $x(0)$).

Euler's Number

What is Euler's e? Metafont users know, it is $e = 2.7183\ldots$
It is the unique number such that for $f(t) = e^t$ we have

$$\frac{df}{dt}(t) = f(t)$$

The exponential function is the fixed-point or eigen-function of the differential operator $\frac{d}{dt}$. One could show this via the Taylor expansion of $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ as $\frac{d}{dt}\frac{t^n}{n!} = \frac{nt^{n-1}}{n(n-1)!} = \frac{t^{n-1}}{(n-1)!}$.

The simplest differential equation one can think of is perhaps:

$$x'(t) = \frac{dx}{dt}(t) = ax(t)$$

The solution is $x(t) = ke^{at} = k\exp(at)$ for some constant k (can be determined via an initial/boundary condition, e.g. $x(0)$).

Euler's Number

What is Euler's e? Metafont users know, it is $e = 2.7183\ldots$
It is the unique number such that for $f(t) = e^t$ we have

$$\frac{df}{dt}(t) = f(t)$$

The exponential function is the fixed-point or eigen-function of the differential operator $\frac{d}{dt}$. One could show this via the Taylor expansion of $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ as $\frac{d}{dt}\frac{t^n}{n!} = \frac{nt^{n-1}}{n(n-1)!} = \frac{t^{n-1}}{(n-1)!}$.

The simplest differential equation one can think of is perhaps:

$$x'(t) = \frac{dx}{dt}(t) = ax(t)$$

The solution is $x(t) = ke^{at} = k\exp(at)$ for some constant k (can be determined via an initial/boundary condition, e.g. $x(0)$).

Euler's Number

What is Euler's e? Metafont users know, it is $e = 2.7183\ldots$
It is the unique number such that for $f(t) = e^t$ we have

$$\frac{df}{dt}(t) = f(t)$$

The exponential function is the fixed-point or eigen-function of the differential operator $\frac{d}{dt}$. One could show this via the Taylor expansion of $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ as $\frac{d}{dt}\frac{t^n}{n!} = \frac{nt^{n-1}}{n(n-1)!} = \frac{t^{n-1}}{(n-1)!}$.

The simplest differential equation one can think of is perhaps:

$$x'(t) = \frac{dx}{dt}(t) = ax(t)$$

The solution is $x(t) = ke^{at} = k\exp(at)$ for some constant k (can be determined via an initial/boundary condition, e.g. $x(0)$).

Ordinary Linear Differential Equations [1, p129]

Solution to ordinary differential equations via exponentation.

$$
\begin{array}{rclcl}
x_1' & = & \frac{dx_1}{dt} & = & a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\
x_2' & = & \frac{dx_2}{dt} & = & a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\
\ldots & & \ldots & & \ldots \\
x_n' & = & \frac{dx_2}{dt} & = & a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n
\end{array}
$$

Theorem

Let $\mathbf{A}$ be an $n \times n$ matrix. Then the unique solution to the initial value problem $\mathbf{x}' = \mathbf{A}\mathbf{x}$ with $\mathbf{x}(0) = \mathbf{x}_0$ is given by

$$\mathbf{x}(t) = \exp(t\mathbf{A})\mathbf{x}_0.$$

Ordinary Linear Differential Equations [1, p129]

Solution to ordinary differential equations via exponentation.

$$
\begin{array}{rcccl}
x_1' & = & \frac{dx_1}{dt} & = & a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\
x_2' & = & \frac{dx_2}{dt} & = & a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\
\ldots & & \ldots & & \ldots \\
x_n' & = & \frac{dx_2}{dt} & = & a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n
\end{array}
$$

Theorem

Let $\mathbf{A}$ *be an* $n \times n$ *matrix. Then the unique solution to the initial value problem* $\mathbf{x}' = \mathbf{A}\mathbf{x}$ *with* $\mathbf{x}(0) = \mathbf{x}_0$ *is given by*

$$\mathbf{x}(t) = \exp(t\mathbf{A})\mathbf{x}_0.$$

Computing Solutions

The exponential of a matrix $\mathbf{A}$ can be computed as:

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}$$

However this anything but an efficient way to compute it.

We can represent matrices e.g. in Jordan normal form: $\mathbf{A} = \mathbf{D} + \mathbf{N}$ where $\mathbf{D}$ is a diagonal matrix and $\mathbf{N}$ is an upper diagonal matrix which is nilpotent, i.e. $\exists m$ s.t. $\mathbf{N}^m$ vanishes. This boils down to finding the eigenvalues of $\mathbf{A}$ (via SVD).

We then have $\exp(\mathbf{A}) = \exp(\mathbf{D} + \mathbf{N}) = \exp(\mathbf{D})\exp(\mathbf{N})$ with $\exp(\text{diag}(d_1, \ldots, d_n)) = \text{diag}(\exp(d_1), \ldots, \exp(d_n))$ and $\mathbf{N}^k \neq 0$ only for finitely many terms.

Computing Solutions

The exponential of a matrix $\mathbf{A}$ can be computed as:

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}$$

However this anything but an efficient way to compute it.

We can represent matrices e.g. in Jordan normal form: $\mathbf{A} = \mathbf{D} + \mathbf{N}$ where $\mathbf{D}$ is a diagonal matrix and $\mathbf{N}$ is an upper diagonal matrix which is nilpotent, i.e. $\exists m$ s.t. $\mathbf{N}^m$ vanishes. This boils down to finding the eigenvalues of $\mathbf{A}$ (via SVD).

We then have $\exp(\mathbf{A}) = \exp(\mathbf{D} + \mathbf{N}) = \exp(\mathbf{D})\exp(\mathbf{N})$ with $\exp(\text{diag}(d_1, \ldots, d_n)) = \text{diag}(\exp(d_1), \ldots, \exp(d_n))$ and $\mathbf{N}^k \neq 0$ only for finitely many terms.

Computing Solutions

The exponential of a matrix $\mathbf{A}$ can be computed as:

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}$$

However this anything but an efficient way to compute it.

We can represent matrices e.g. in Jordan normal form: $\mathbf{A} = \mathbf{D} + \mathbf{N}$ where $\mathbf{D}$ is a diagonal matrix and $\mathbf{N}$ is an upper diagonal matrix which is nilpotent, i.e. $\exists m$ s.t. $\mathbf{N}^m$ vanishes. This boils down to finding the eigenvalues of $\mathbf{A}$ (via SVD).

We then have $\exp(\mathbf{A}) = \exp(\mathbf{D} + \mathbf{N}) = \exp(\mathbf{D})\exp(\mathbf{N})$ with $\exp(\text{diag}(d_1, \ldots, d_n)) = \text{diag}(\exp(d_1), \ldots, \exp(d_n))$ and $\mathbf{N}^k \neq 0$ only for finitely many terms.

Computing Solutions

The exponential of a matrix $\mathbf{A}$ can be computed as:

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}$$

However this anything but an efficient way to compute it.

We can represent matrices e.g. in Jordan normal form: $\mathbf{A} = \mathbf{D} + \mathbf{N}$ where $\mathbf{D}$ is a diagonal matrix and $\mathbf{N}$ is an upper diagonal matrix which is nilpotent, i.e. $\exists m$ s.t. $\mathbf{N}^m$ vanishes. This boils down to finding the eigenvalues of $\mathbf{A}$ (via SVD).

We then have $\exp(\mathbf{A}) = \exp(\mathbf{D} + \mathbf{N}) = \exp(\mathbf{D})\exp(\mathbf{N})$ with $\exp(\mathrm{diag}(d_1, \ldots, d_n)) = \mathrm{diag}(\exp(d_1), \ldots, \exp(d_n))$ and $\mathbf{N}^k \neq 0$ only for finitely many terms.

Smooth Functions

We say a function f is differentiable on $\mathbb{R}$ or $U \in \mathbb{R}$ if it is differentiable at every point $t \in \mathbb{R}$ or $t \in U$. We then write $f \in C^1(\mathbb{R}) = C^1$ or $f \in C(U)$.

We can see $Df(x)$ itself as a function $\mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{nm}$.

As such we can ask if this is itself differentiable. We denote the set of p-times differentiable maps by C^p and by C^∞ the set of infinitely differentiable or smooth functions.

Note: Differentiation is primarily a real number notion. We need to introduce the notion of a differentiable manifold as a space which looks like $\mathbb{R}^m$ locally (with respect to diff. operations).

Smooth Functions

We say a function f is differentiable on $\mathbb{R}$ or $U \in \mathbb{R}$ if it is differentiable at every point $t \in \mathbb{R}$ or $t \in U$. We then write $f \in C^1(\mathbb{R}) = C^1$ or $f \in C(U)$.

We can see $Df(x)$ itself as a function $\mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{nm}$.

As such we can ask if this is itself differentiable. We denote the set of p-times differentiable maps by C^p and by C^∞ the set of infinitely differentiable or smooth functions.

Note: Differentiation is primarily a real number notion. We need to introduce the notion of a differentiable manifold as a space which looks like $\mathbb{R}^m$ locally (with respect to diff. operations).

Smooth Functions

We say a function f is differentiable on $\mathbb{R}$ or $U \in \mathbb{R}$ if it is differentiable at every point $t \in \mathbb{R}$ or $t \in U$. We then write $f \in C^1(\mathbb{R}) = C^1$ or $f \in C(U)$.

We can see $Df(x)$ itself as a function $\mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \mathbb{R}^{nm}$.

As such we can ask if this is itself differentiable. We denote the set of p-times differentiable maps by C^p and by C^∞ the set of infinitely differentiable or smooth functions.

Note: Differentiation is primarily a real number notion. We need to introduce the notion of a differentiable manifold as a space which looks like $\mathbb{R}^m$ locally (with respect to diff. operations).

Differentiable Manifolds

Definition

Let M be a topological space.

A chart (V, Φ) is a homeomorphism Φ of an open set V of M into an open set of $\mathbb{R}^m$.

Two charts (V_1, Φ_1) and (V_2, Φ_2) are said to be compatible in case $V_1 \cap V_2 = \emptyset$ or the restricted maps $\Phi_1 \circ \Phi_2^{-1}$ and $\Phi_2 \circ \Phi_1^{-1}$ are in $C^\infty(\mathbb{R}^m)$.

A atlas is a set of compatible charts that cover all of M. Two atlases are compatible if all their charts are.

A differentiable manifold is a separable, metrizisable space with an set of compatible (equivalent) atlases.

Differentiable Manifolds

Definition

Let M be a topological space.

A chart (V, Φ) is a homeomorphism Φ of an open set V of M into an open set of $\mathbb{R}^m$.

Two charts (V_1, Φ_1) and (V_2, Φ_2) are said to be compatible in case $V_1 \cap V_2 = \emptyset$ or the restricted maps $\Phi_1 \circ \Phi_2^{-1}$ and $\Phi_2 \circ \Phi_1^{-1}$ are in $C^\infty(\mathbb{R}^m)$.

A atlas is a set of compatible charts that cover all of M. Two atlases are compatible if all their charts are.

A differentiable manifold is a separable, metrizisable space with an set of compatible (equivalent) atlases.

Definition

Let M be a topological space.

A chart (V, Φ) is a homeomorphism Φ of an open set V of M into an open set of $\mathbb{R}^m$.

Two charts (V_1, Φ_1) and (V_2, Φ_2) are said to be compatible in case $V_1 \cap V_2 = \emptyset$ or the restricted maps $\Phi_1 \circ \Phi_2^{-1}$ and $\Phi_2 \circ \Phi_1^{-1}$ are in $C^\infty(\mathbb{R}^m)$.

A atlas is a set of compatible charts that cover all of M. Two atlases are compatible if all their charts are.

A differentiable manifold is a separable, metrizisable space with an set of compatible (equivalent) atlases.

Differentiable Manifolds

Definition

Let M be a topological space.

A chart (V, Φ) is a homeomorphism Φ of an open set V of M into an open set of $\mathbb{R}^m$.

Two charts (V_1, Φ_1) and (V_2, Φ_2) are said to be compatible in case $V_1 \cap V_2 = \emptyset$ or the restricted maps $\Phi_1 \circ \Phi_2^{-1}$ and $\Phi_2 \circ \Phi_1^{-1}$ are in $C^\infty(\mathbb{R}^m)$.

A atlas is a set of compatible charts that cover all of M. Two atlases are compatible if all their charts are.

A differentiable manifold is a separable, metrizisable space with an set of compatible (equivalent) atlases.

Tangents

Definition

Let $f, g \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ then we say that f is tangent to g at t iff

$$\lim_{t' \to t} \frac{\|f(t') - g(t')\|}{\|t' - t\|} = 0$$

Definition

Let M be a manifold and $m \in M$. A curve at m is a C^1 map $c : I \to M$ with an open interval in $\mathbb{R}$ containing 0 s.t. $c(0) = m$.

We say that two curves c_1 and c_2 are tangent if $\Phi \circ c_1$ and $\Phi \circ c_2$ are tangent at 0.

The tangent space $\mathbf{T}_m(M)$ of M at m is the set of (tangent) equivalent classes of curves.

Tangents

Definition

Let $f, g \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ then we say that f is tangent to g at t iff

$$\lim_{t' \to t} \frac{\|f(t') - g(t')\|}{\|t' - t\|} = 0$$

Definition

Let M be a manifold and $m \in M$. A curve at m is a C^1 map $c : I \to M$ with an open interval in $\mathbb{R}$ containing 0 s.t. $c(0) = m$.

We say that two curves c_1 and c_2 are tangent if $\Phi \circ c_1$ and $\Phi \circ c_2$ are tangent at 0.

The tangent space $\mathbf{T}_m(M)$ of M at m is the set of (tangent) equivalent classes of curves.

Tangents

Definition

Let $f, g \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ then we say that f is tangent to g at t iff

$$\lim_{t' \to t} \frac{\|f(t') - g(t')\|}{\|t' - t\|} = 0$$

Definition

Let M be a manifold and $m \in M$. A curve at m is a C^1 map $c : I \to M$ with an open interval in $\mathbb{R}$ containing 0 s.t. $c(0) = m$.

We say that two curves c_1 and c_2 are tangent if $\Phi \circ c_1$ and $\Phi \circ c_2$ are tangent at 0.

The tangent space $\mathbf{T}_m(M)$ of M at m is the set of (tangent) equivalent classes of curves.

Definition

A Lie group over a field $\mathbb{K}$ is a group G equipped with the structure of a differentiable manifold over $\mathbb{K}$ sich that

$$. \cdot . : G \times G \to G \quad \text{is differentable.}$$

Using the implicit function theorem, one can also show that $g \mapsto g^{-1}$ is differentiable (a diffeomorphism).

The fields we are typically interested are $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Definition

A Lie group over a field $\mathbb{K}$ is a group G equipped with the structure of a differentiable manifold over $\mathbb{K}$ sich that

$$. \cdot . : G \times G \to G \quad \text{is differentable.}$$

Using the implicit function theorem, one can also show that $g \mapsto g^{-1}$ is differentiable (a diffeomorphism).

The fields we are typically interested are $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Lie Groups

Definition

A Lie group over a field $\mathbb{K}$ is a group G equipped with the structure of a differentiable manifold over $\mathbb{K}$ sich that

$$. \cdot . : G \times G \to G \quad \text{is differentable.}$$

Using the implicit function theorem, one can also show that $g \mapsto g^{-1}$ is differentiable (a diffeomorphism).

The fields we are typically interested are $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Fields

Definition

A **field** is a set $\mathbb{K}$ together with two operations:

$$\text{Addition} \quad .+. : \mathbb{K} \times \mathbb{K} \mapsto \mathbb{K}$$

$$\text{Multiplication} \quad .\cdot. : \mathbb{K} \times \mathbb{K} \mapsto \mathbb{K}$$

such that

1. $\forall x, y, z \in \mathbb{K}: \; x + (y + z) = (x + y) + z$
2. $\exists\, o \in \mathbb{K}, \forall x \in \mathbb{K}: \; o + x = x$
3. $\forall\, x \in \mathbb{K}, \exists -x \in \mathbb{K}: \; x + (-x) = o$
4. $\forall x, y \in \mathbb{K}: \; x + y = x + y$
5. $\forall x, y, z \in \mathbb{K}: \; x \cdot (y \cdot z) = (x \cdot y) \cdot z$
6. $\exists\, o \neq e \in \mathbb{K}, \forall x \in \mathbb{K}: \; e \cdot x = x$
7. $\forall\, o \neq x \in \mathbb{K}, \exists x^{-1} \in \mathbb{K}: \; x \cdot x^{-1} = e$
8. $\forall x, y \in \mathbb{K}: \; x \cdot y = y \cdot x$
9. $x \cdot (y + z) = x \cdot y + x \cdot z, \quad \forall x, y, z \in \mathbb{K}$

Examples of Lie Groups

Examples of Lie groups we can mentions here:

- The additive group of the field $\mathbb{K} = \mathbb{K}^+$.
- The multiplicative group of the field $\mathbb{K}^\times$.
- The "circle" $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ or $\{e^{i\phi} \mid \phi \in [0, 2\pi)\}$.
- $GL_n(\mathbb{K})$ of invertible matrices of order n over $\mathbb{K}$.
- $SL_n(\mathbb{K})$ of matrices of order n over $\mathbb{K}$ with $\det = 1$.
- $O_n(\mathbb{K})$ orthogonal matrices over $\mathbb{K}$ of order n.
- U_n unitary matrices over $\mathbb{C}$ of order n.
- $SO_n(\mathbb{K}) = O_n(\mathbb{K}) \cap SL_n(\mathbb{K})$.
- $SU_n = U_n \cap SL_n(\mathbb{C})$.

Lie Algebras

Definition

A Lie algebra is a vector space $\mathfrak{g}$ over some field $\mathbb{K}$ together with a binary operation $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the **Lie bracket**, which satisfies the following:

Bilinearity: $\forall \alpha, \beta \in \mathbb{K}$ and $\forall x, y, z \in \mathfrak{g}$

$$[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z]$$

$$[z, \alpha x + \beta y] = \alpha[z, x] + \beta[z, y]$$

Alternating on $\mathfrak{g}$: $\forall x \in \mathfrak{g}$

$$[x, x] = 0$$

Jacobi identity: $\forall x, y, z \in \mathfrak{g}$

$$[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$$

It follows easily $\forall x, y \in \mathfrak{g}$ that $[x, y] = -[y, x]$. One could also define a associative product on an algebra $\mathfrak{g}$ and then introduce the Lie bracket as $[x, y] = xy - yx$.

Theorem

Given a Lie group G then the tangent space at the unit $\mathfrak{g} = \mathbf{T}_e G$ is a Lie algebra.

Let $g(t)$ and $h(t)$ be differentiable paths or C^1 curves on G. Assume, $g(0) = h(0) = e$ as well as $\frac{dg}{dt}(0) = \xi$ and $\frac{dh}{dt}(0) = \eta$ then we define a Lie bracket on the tangent space $\mathbf{T}_e(G)$ via

$$[\xi, \eta] = \frac{\partial^2}{\partial t \partial s}[g(t), h(s)]|_{s=t=0}$$

where $[g, h] = ghg^{-1}h^{-1}$ is the *group commutator*.

Examples of Lie Algebras

It follows easily $\forall x, y \in \mathfrak{g}$ that $[x, y] = -[y, x]$. One could also define a associative product on an algebra $\mathfrak{g}$ and then introduce the Lie bracket as $[x, y] = xy - yx$.

Theorem

Given a Lie group G then the tangent space at the unit $\mathfrak{g} = \mathbf{T}_e G$ is a Lie algebra.

Let $g(t)$ and $h(t)$ be differentiable paths or C^1 curves on G. Assume, $g(0) = h(0) = e$ as well as $\frac{dg}{dt}(0) = \xi$ and $\frac{dh}{dt}(0) = \eta$ then we define a Lie bracket on the tangent space $\mathbf{T}_e(G)$ via

$$[\xi, \eta] = \frac{\partial^2}{\partial t \partial s}[g(t), h(s)]|_{s=t=0}$$

where $[g, h] = ghg^{-1}h^{-1}$ is the *group commutator*.

Examples of Lie Algebras

It follows easily $\forall x, y \in \mathfrak{g}$ that $[x, y] = -[y, x]$. One could also define a associative product on an algebra $\mathfrak{g}$ and then introduce the Lie bracket as $[x, y] = xy - yx$.

Theorem

Given a Lie group G then the tangent space at the unit $\mathfrak{g} = \mathbf{T}_e G$ is a Lie algebra.

Let $g(t)$ and $h(t)$ be differentiable paths or C^1 curves on G. Assume, $g(0) = h(0) = e$ as well as $\frac{dg}{dt}(0) = \xi$ and $\frac{dh}{dt}(0) = \eta$ then we define a Lie bracket on the tangent space $\mathbf{T}_e(G)$ via

$$[\xi, \eta] = \frac{\partial^2}{\partial t \partial s}[g(t), h(s)]|_{s=t=0}$$

where $[g, h] = ghg^{-1}h^{-1}$ is the *group commutator*.

Prescribed Velocities

Definition

A path or curve $g(t)$ in a Lie group G with $t \in \mathbb{R}$ is called a **one-parameter subgroup** if

$$g(t+s) = g(t)g(s).$$

We denote by $g_\xi(s)$ the one-parameter sub-group with $g' = \frac{dg}{dt}(s) = \xi(s)$ – i.e. with prescribed "velocity" $\xi(s)$.

Definition

For a Lie group G and $\xi \in \mathfrak{g}$, i.e. its Lie algebra, we define:

$$\exp(\xi) = g_\xi(1)$$

Prescribed Velocities

Definition

A path or curve $g(t)$ in a Lie group G with $t \in \mathbb{R}$ is called a **one-parameter subgroup** if

$$g(t+s) = g(t)g(s).$$

We denote by $g_\xi(s)$ the one-parameter sub-group with $g' = \frac{dg}{dt}(s) = \xi(s)$ – i.e. with prescribed "velocity" $\xi(s)$.

Definition

For a Lie group G and $\xi \in \mathfrak{g}$, i.e. its Lie algebra, we define:

$$\exp(\xi) = g_\xi(1)$$

Prescribed Velocities

Definition

A path or curve $g(t)$ in a Lie group G with $t \in \mathbb{R}$ is called a **one-parameter subgroup** if

$$g(t+s) = g(t)g(s).$$

We denote by $g_\xi(s)$ the one-parameter sub-group with $g' = \frac{dg}{dt}(s) = \xi(s)$ – i.e. with prescribed "velocity" $\xi(s)$.

Definition

For a Lie group G and $\xi \in \mathfrak{g}$, i.e. its Lie algebra, we define:

$$\exp(\xi) = g_\xi(1)$$

Exponentation

Theorem

The exponential map $\exp : \mathfrak{g} \to G$ *maps a neighbourhood of zero in the tangent algebra* $\mathfrak{g} = \mathbf{T}_e(G)$ *diffeomorphically onto a neighbourhood of the identity in* G*.*

Theorem

Let $\mathfrak{g} = \mathfrak{a}_1 \oplus \ldots \oplus \mathfrak{a}_k$ *be a decomposition of a Lie algebra as direct sum, then* $\xi_1 + \ldots + \xi_k \mapsto \exp(\xi_1) \ldots \exp(\xi_k)$ *maps a neighbourhood of zero in* $\mathfrak{g}$ *diffeomorphically onto a neighbourhood of the identity in* G*.*

If G is the group of invertible elements in an associative algebra (e.g. of non-singular matrices), then

$$\exp(\xi) = \sum_{n=0}^{\infty} \frac{\xi^n}{n!}.$$

Exponentation

Theorem

The exponential map $\exp : \mathfrak{g} \to G$ *maps a neighbourhood of zero in the tangent algebra* $\mathfrak{g} = \mathbf{T}_e(G)$ *diffeomorphically onto a neighbourhood of the identity in* G.

Theorem

Let $\mathfrak{g} = \mathfrak{a}_1 \oplus \ldots \oplus \mathfrak{a}_k$ *be a decomposition of a Lie algebra as direct sum, then* $\xi_1 + \ldots + \xi_k \mapsto \exp(\xi_1) \ldots \exp(\xi_k)$ *maps a neighbourhood of zero in* $\mathfrak{g}$ *diffeomorphically onto a neighbourhood of the identity in* G.

If G is the group of invertible elements in an associative algebra (e.g. of non-singular matrices), then

$$\exp(\xi) = \sum_{n=0}^{\infty} \frac{\xi^n}{n!}.$$

Exponentation

Theorem

The exponential map $\exp : \mathfrak{g} \to G$ *maps a neighbourhood of zero in the tangent algebra* $\mathfrak{g} = \mathbf{T}_e(G)$ *diffeomorphically onto a neighbourhood of the identity in G.*

Theorem

Let $\mathfrak{g} = \mathfrak{a}_1 \oplus \ldots \oplus \mathfrak{a}_k$ *be a decomposition of a Lie algebra as direct sum, then* $\xi_1 + \ldots + \xi_k \mapsto \exp(\xi_1) \ldots \exp(\xi_k)$ *maps a neighbourhood of zero in* $\mathfrak{g}$ *diffeomorphically onto a neighbourhood of the identity in G.*

If G is the group of invertible elements in an associative algebra (e.g. of non-singular matrices), then

$$\exp(\xi) = \sum_{n=0}^{\infty} \frac{\xi^n}{n!}.$$

Stochastic Dynamics

Discrete Time Markov Chains (DTCM)

Definition

A discrete time Markov chain (DTMC) on S is defined via a stochastic matrix $\mathbf{P}$, i.e. an $r \times r$ (square) matrix with entries $0 \leq p_{ij} \leq 1$ and such that all row sums are equal to one, i.e.

$$\sum_j p_{ij} = 1.$$

This defines a discrete linear dynamical system:

Phase group: $\mathbb{Z}$ or $\mathbb{N}$,

Phase space: $\mathbb{R}^r$,

Group action: $\pi(n, x) = x \cdot \mathbf{P}^n$.

Discrete Time Markov Chains (DTCM)

Definition

A discrete time Markov chain (DTMC) on S is defined via a stochastic matrix $\mathbf{P}$, i.e. an $r \times r$ (square) matrix with entries $0 \leq p_{ij} \leq 1$ and such that all row sums are equal to one, i.e.

$$\sum_j p_{ij} = 1.$$

This defines a discrete linear dynamical system:

Phase group: $\mathbb{Z}$ or $\mathbb{N}$,

Phase space: $\mathbb{R}^r$,

Group action: $\pi(n, x) = x \cdot \mathbf{P}^n$.

Memoryless Property of DTMC

Let I be a finite (or maybe countable) set. Each $i \in I$ is called a state or index. Given a probability space $(\Omega, \sigma, \mathbb{P})$ a random variable is a map $X : \Omega \to I$.

A sequence of random variables X_n is a Markov Chain if

$$\begin{aligned} &\mathbb{P}(X_{n+1} = i+1 \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \\ &= \mathbb{P}(X_{n+1} = i+1 \mid X_n = i_n) = \\ &= p_{i_n, i_{n+1}} \mathbb{P}(X_n = i_n) \end{aligned}$$

The probability $\mathbb{P}(i \to^n j)$ of reaching state (actually index) j from i in exactly n steps is given by $p_{ij}^{(n)}$ i.e. the entry in row i and column j of $\mathbf{P}^n$.

Memoryless Property of DTMC

Let I be a finite (or maybe countable) set. Each $i \in I$ is called a state or index. Given a probability space $(\Omega, \sigma, \mathbb{P})$ a random variable is a map $X : \Omega \to I$.

A sequence of random variables X_n is a Markov Chain if

$$\begin{aligned} \mathbb{P}(X_{n+1} = i+1 \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \\ = \quad & \mathbb{P}(X_{n+1} = i+1 \mid X_n = i_n) = \\ = \quad & p_{i_n, i_{n+1}} \mathbb{P}(X_n = i_n) \end{aligned}$$

The probability $\mathbb{P}(i \to^n j)$ of reaching state (actually index) j from i in exactly n steps is given by $p_{ij}^{(n)}$ i.e. the entry in row i and column j of $\mathbf{P}^n$.

Memoryless Property of DTMC

Let I be a finite (or maybe countable) set. Each $i \in I$ is called a state or index. Given a probability space $(\Omega, \sigma, \mathbb{P})$ a random variable is a map $X : \Omega \to I$.

A sequence of random variables X_n is a Markov Chain if

$$\begin{aligned} \mathbb{P}(X_{n+1} = i+1 \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \\ = \quad \mathbb{P}(X_{n+1} = i+1 \mid X_n = i_n) = \\ = \quad p_{i_n, i_{n+1}} \mathbb{P}(X_n = i_n) \end{aligned}$$

The probability $\mathbb{P}(i \to^n j)$ of reaching state (actually index) j from i in exactly n steps is given by $p_{ij}^{(n)}$ i.e. the entry in row i and column j of $\mathbf{P}^n$.

Properties

Definition

Given a DTMC with transition matrix $\mathbf{P}$. A state i is said to be

recurrent if $\mathbb{P}(i \rightarrow^n i$ for infinitely many n$\} = 1$

transient if $\mathbb{P}(i \rightarrow^n i$ for infinitely many n$\} = 0$

Definition

A DTMC with transition matrix $\mathbf{P}$ is called

ergodic or irreducible: if $\forall i, j \; \exists n$ such that $P_{ij}^n > 0$.

regular: if $\exists n$ such that $\forall i, j$ we have $P_{ij}^n > 0$.

Properties

Definition

Given a DTMC with transition matrix **P**. A state i is said to be

recurrent if $\mathbb{P}(i \rightarrow^n i \text{ for infinitely many n}\} = 1$

transient if $\mathbb{P}(i \rightarrow^n i \text{ for infinitely many n}\} = 0$

Definition

A DTMC with transition matrix **P** is called

ergodic or irreducible: if $\forall i, j \; \exists n$ such that $P_{ij}^n > 0$.

regular: if $\exists n$ such that $\forall i, j$ we have $P_{ij}^n > 0$.

Long Run Behaviour

Theorem

Given a DTMC with transition matrix $\mathbf{P}$. If it is regular and v an arbitrary probability vector. Then

$$\lim_{n \to \infty} v\mathbf{P}^n = w$$

where w is the unique probability vector for $\mathbf{P}$.

Theorem

Given a DTMC with transition matrix $\mathbf{P}$. Assume $\mathbf{P}$ is ergodic. Let $\mathbf{A}_n$ be the matrix defined by:

$$\mathbf{A}_n = \frac{\mathbf{I} + \mathbf{P} + \ldots + \mathbf{P}^n}{n+1}$$

then $\mathbf{A}_n \to \mathbf{W}$ where $\mathbf{W}$ is a matrix all of whose rows are equal to the unique vector w for $\mathbf{P}$.

Long Run Behaviour

Theorem

Given a DTMC with transition matrix $\mathbf{P}$. If it is regular and v an arbitrary probability vector. Then

$$\lim_{n \to \infty} v\mathbf{P}^n = w$$

where w is the unique probability vector for $\mathbf{P}$.

Theorem

Given a DTMC with transition matrix $\mathbf{P}$. Assume $\mathbf{P}$ is ergodic. Let $\mathbf{A}_n$ be the matrix defined by:

$$\mathbf{A}_n = \frac{\mathbf{I} + \mathbf{P} + \ldots + \mathbf{P}^n}{n+1}$$

then $\mathbf{A}_n \to \mathbf{W}$ where $\mathbf{W}$ is a matrix all of whose rows are equal to the unique vector w for $\mathbf{P}$.

Definition

A continuous time Markov chain (CTMC) on $S = \{s_1, \ldots, s_r\}$ is defined via an $r \times r$ (square) generator or Q-matrix $\mathbf{Q} = (q_{ij})$ specifying the rates going from an index or state i to an index or state j and which fullfills:

1. $0 \leq -q_{ii} < \infty$ for all i
2. $q_{ij} \geq 0$ for all $i \neq j$
3. $\sum_j q_{ij} = 0$ for all i

Continuous Time Markov Chains (CTMC)

Definition

A continuous time Markov chain (CTMC) on $S = \{s_1, \ldots, s_r\}$ is defined via an $r \times r$ (square) generator or Q-matrix $\mathbf{Q} = (q_{ij})$ specifying the rates going from an index or state i to an index or state j and which fullfills:

1. $0 \leq -q_{ii} < \infty$ for all i
2. $q_{ij} \geq 0$ for all $i \neq j$
3. $\sum_j q_{ij} = 0$ for all i

Computing Transition Probabilities

Again we use exponentation to get the transition probabilities.

$$\mathbf{P}(t) = \exp(t\mathbf{Q}) = \sum_{k=0}^{\infty} \frac{(t\mathbf{Q})^k}{k!}$$

This gives the unique solutions to the forward equations

$$\frac{d}{dt}P(t) = \mathbf{P}(t)\mathbf{Q} \quad \text{with } \mathbf{P}(0) = \mathbf{I}$$

and the backward equation

$$\frac{d}{dt}P(t) = \mathbf{Q}\mathbf{P}(t) \quad \text{with } \mathbf{P}(0) = \mathbf{I}$$

and fulfills the (semi-)group property:

$$\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t).$$

Computing Transition Probabilities

Again we use exponentation to get the transition probabilities.

$$\mathbf{P}(t) = \exp(t\mathbf{Q}) = \sum_{k=0}^{\infty} \frac{(t\mathbf{Q})^k}{k!}$$

This gives the unique solutions to the forward equations

$$\frac{d}{dt}P(t) = \mathbf{P}(t)\mathbf{Q} \text{ with } \mathbf{P}(0) = \mathbf{I}$$

and the backward equation

$$\frac{d}{dt}P(t) = \mathbf{Q}\mathbf{P}(t) \text{ with } \mathbf{P}(0) = \mathbf{I}$$

and fulfills the (semi-)group property:

$$\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t).$$

Dynamical Systems in Physics

Classical Mechanics (in 10min)

Consider point particles (no volume) with mass m. The position of a particle is given in some coordinates q_i.

The velocity of the particle is given by

$$v_i = \frac{dq_i}{dt} = \dot{q}_i$$

its acceleration is given by

$$a_i = \frac{dv}{dt} = \frac{d^2 q_i}{dt^2} = \ddot{q}_i$$

its momentum is defined as

$$p_i = m\dot{q}_i = m\frac{dq_i}{dt}$$

Classical Mechanics (in 10min)

Consider point particles (no volume) with mass m. The position of a particle is given in some coordinates q_i.

The velocity of the particle is given by

$$v_i = \frac{dq_i}{dt} = \dot{q}_i$$

its acceleration is given by

$$a_i = \frac{dv}{dt} = \frac{d^2 q_i}{dt^2} = \ddot{q}_i$$

its momentum is defined as

$$p_i = m\dot{q}_i = m\frac{dq_i}{dt}$$

Classical Mechanics (in 10min)

Consider point particles (no volume) with mass m. The position of a particle is given in some coordinates q_i.

The velocity of the particle is given by

$$v_i = \frac{dq_i}{dt} = \dot{q}_i$$

its acceleration is given by

$$a_i = \frac{dv}{dt} = \frac{d^2 q_i}{dt^2} = \ddot{q}_i$$

its momentum is defined as

$$p_i = m\dot{q}_i = m\frac{dq_i}{dt}$$

Classical Mechanics (in 10min)

Consider point particles (no volume) with mass m. The position of a particle is given in some coordinates q_i.

The velocity of the particle is given by

$$v_i = \frac{dq_i}{dt} = \dot{q}_i$$

its acceleration is given by

$$a_i = \frac{dv}{dt} = \frac{d^2 q_i}{dt^2} = \ddot{q}_i$$

its momentum is defined as

$$p_i = m\dot{q}_i = m\frac{dq_i}{dt}$$

Lagrange Formalism

Describe the dynamics of a mechanical system via the Lagrange function or Lagrangian

$$L(q_1, q_2, \ldots, q_s, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_s, t)$$

the action is defined as $S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)dt$.

The Principle of Least Action then implies the Lagrange equations which give the dynamics:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0$$

L.D. Landau and E.M. Lifschitz. *Mechanik*. Akademie-Verlag, Berlin, 1981.

Lagrange Formalism

Describe the dynamics of a mechanical system via the Lagrange function or Lagrangian

$$L(q_1, q_2, \ldots, q_s, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_s, t)$$

the action is defined as $S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)dt$.

The Principle of Least Action then implies the Lagrange equations which give the dynamics:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0$$

L.D. Landau and E.M. Lifschitz. *Mechanik*. Akademie-Verlag, Berlin, 1981.

Lagrange Examples

Single Particle:

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

Pendulum: length l, angle ϕ, mass m, gravitational constant g

$$L = \frac{m}{2} l^2 \dot{\phi}^2 + mgl\cos(\phi)$$

Double Pendulum: angles ϕ_1 and ϕ_2, lengths l_1 and l_2, masses m_1 and m_2 [2, p13]:

$$\begin{aligned} L = {} & \frac{m_1 + m_2}{2} l_1^2 \dot{\phi}_1^2 + \frac{m_2}{2} l_2^2 \dot{\phi}_2^2 + \\ & m_2 l_1 l_2 \dot{\phi}_1 \dot{\phi}_2 \cos(\phi_1 - \phi_2) + \\ & (m_1 + m_2) g l_1 \cos(\phi_1) + m_2 g l_2 \cos(\phi_2) \end{aligned}$$

Lagrange Examples

Single Particle:

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

Pendulum: length l, angle ϕ, mass m, gravitational constant g

$$L = \frac{m}{2}l^2\dot{\phi}^2 + mgl\cos(\phi)$$

Double Pendulum: angles ϕ_1 and ϕ_2, lengths l_1 and l_2, masses m_1 and m_2 [2, p13]:

$$\begin{aligned} L = {} & \frac{m_1 + m_2}{2} l_1^2 \dot{\phi}_1^2 + \frac{m_2}{2} l_2^2 \dot{\phi}_2^2 + \\ & m_2 l_1 l_2 \dot{\phi}_1 \dot{\phi}_2 \cos(\phi_1 - \phi_2) + \\ & (m_1 + m_2) g l_1 \cos(\phi_1) + m_2 g l_2 \cos(\phi_2) \end{aligned}$$

Lagrange Examples

Single Particle:

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

Pendulum: length l, angle ϕ, mass m, gravitational constant g

$$L = \frac{m}{2} l^2 \dot{\phi}^2 + mgl\cos(\phi)$$

Double Pendulum: angles ϕ_1 and ϕ_2, lengths l_1 and l_2, masses m_1 and m_2 [2, p13]:

$$\begin{aligned} L \quad = \quad & \frac{m_1 + m_2}{2} l_1^2 \dot{\phi}_1^2 + \frac{m_2}{2} l_2^2 \dot{\phi}_2^2 + \\ & m_2 l_1 l_2 \dot{\phi}_1 \dot{\phi}_2 \cos(\phi_1 - \phi_2) + \\ & (m_1 + m_2) g l_1 \cos(\phi_1) + m_2 g l_2 \cos(\phi_2) \end{aligned}$$

Hamiltonian Formalism

Describe the dynamics of a mechanical system via the Hamilton function or Hamiltonian:

$$H(p_i, q_i, t) = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i, t)$$

The dynamics of the system is then described via the Hamiltonian or canonical equations:

$$\dot{q}_i = \frac{dq}{dt} = \frac{\partial H}{\partial p_i}$$

$$\dot{q}_i = \frac{dq}{dt} = -\frac{\partial H}{\partial q_i}$$

L.D. Landau and E.M. Lifschitz. *Mechanik*. Akademie-Verlag, Berlin, 1981.

Hamiltonian Formalism

Describe the dynamics of a mechanical system via the Hamilton function or Hamiltonian:

$$H(p_i, q_i, t) = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i, t)$$

The dynamics of the system is then described via the Hamiltonian or canonical equations:

$$\dot{q}_i = \frac{dq}{dt} = \frac{\partial H}{\partial p_i}$$

$$\dot{q}_i = \frac{dq}{dt} = -\frac{\partial H}{\partial q_i}$$

L.D. Landau and E.M. Lifschitz. *Mechanik*. Akademie-Verlag, Berlin, 1981.

Hamiltonian Examples

Single Particle:

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)$$

Particle in Field:

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U(x, y, z)$$

Pendulum: with $p_\phi = ml^2\dot{\phi}$ and $\dot{\phi} = \frac{p_\phi}{ml^2}$

$$H = \frac{p_\phi^2}{2ml^2} - mgl\cos(\phi)$$

Hamiltonian Examples

Single Particle:

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)$$

Particle in Field:

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U(x, y, z)$$

Pendulum: with $p_\phi = ml^2\dot{\phi}$ and $\dot{\phi} = \frac{p_\phi}{ml^2}$

$$H = \frac{p_\phi^2}{2ml^2} - mgl\cos(\phi)$$

Hamiltonian Examples

Single Particle:

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)$$

Particle in Field:

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U(x, y, z)$$

Pendulum: with $p_\phi = ml^2\dot{\phi}$ and $\dot{\phi} = \frac{p_\phi}{ml^2}$

$$H = \frac{p_\phi^2}{2ml^2} - mgl\cos(\phi)$$

Quantum Mechanics (in 20min)

Arguably, **physics** is ultimately about explaining experiments and forecasting measurement results.

Observables: Entities which are (actually) measured when an experiment is conducted on a system.

State: Entities which completely describe (or model) the system we are interested in.

Measurement establishes a relation between states and observables of a given system. Dynamics describes how observables and/or the state changes over time.

Related Questions: What is our knowledge of what? How do we obtain this information? What is a description on how the system changes?

Quantum Mechanics (in 20min)

Arguably, **physics** is ultimately about explaining experiments and forecasting measurement results.

Observables: Entities which are (actually) measured when an experiment is conducted on a system.

State: Entities which completely describe (or model) the system we are interested in.

Measurement establishes a relation between states and observables of a given system. Dynamics describes how observables and/or the state changes over time.

Related Questions: What is our knowledge of what? How do we obtain this information? What is a description on how the system changes?

Quantum Mechanics (in 20min)

Arguably, **physics** is ultimately about explaining experiments and forecasting measurement results.

Observables: Entities which are (actually) measured when an experiment is conducted on a system.

State: Entities which completely describe (or model) the system we are interested in.

Measurement establishes a relation between states and observables of a given system. Dynamics describes how observables and/or the state changes over time.

Related Questions: What is our knowledge of what? How do we obtain this information? What is a description on how the system changes?

Quantum Mechanics (in 20min)

Arguably, **physics** is ultimately about explaining experiments and forecasting measurement results.

Observables: Entities which are (actually) measured when an experiment is conducted on a system.

State: Entities which completely describe (or model) the system we are interested in.

Measurement establishes a relation between states and observables of a given system. Dynamics describes how observables and/or the state changes over time.

Related Questions: What is our knowledge of what? How do we obtain this information? What is a description on how the system changes?

Quantum Mechanics (in 20min)

Arguably, **physics** is ultimately about explaining experiments and forecasting measurement results.

Observables: Entities which are (actually) measured when an experiment is conducted on a system.

State: Entities which completely describe (or model) the system we are interested in.

Measurement establishes a relation between states and observables of a given system. Dynamics describes how observables and/or the state changes over time.

Related Questions: What is our knowledge of what? How do we obtain this information? What is a description on how the system changes?

Postulates for Quantum Mechanics (ca 1950)

- The quantum **state** of a (free) particle is described by a (normalised) complex valued function:

$$\vec{\psi} \in L^2(x) \text{ i.e. } \int |\psi(x)|^2 dx = 1$$

- Two quantum states can be **superimposed**, i.e.

$$\alpha_1 \vec{\psi}_1 + \alpha_2 \vec{\psi}_2$$

- Any **observable** A is represented by a linear, self-adjoint operator **A** on $L^2(x)$.
- **Possible** measurement results: Eigenvalues of **A**, representing the observable A:

$$\mathbf{A}\vec{\phi}_i = \lambda_i \vec{\phi}_i$$

- **Probability** to measure λ_n if the system is in state $\vec{\psi} = \sum_i \psi_i \vec{\phi}_i$ is

$$Pr(A = \lambda_n, \vec{\psi}) = \|\vec{\psi}_n\|^2$$

Postulates for Quantum Mechanics (ca 1950)

- The quantum **state** of a (free) particle is described by a (normalised) complex valued function:

$$\vec{\psi} \in L^2(x) \text{ i.e. } \int |\psi(x)|^2 dx = 1$$

- Two quantum states can be **superimposed**, i.e.

$$\alpha_1 \vec{\psi}_1 + \alpha_2 \vec{\psi}_2$$

- Any **observable** A is represented by a linear, self-adjoint operator $\mathbf{A}$ on $L^2(x)$.
- **Possible** measurement results: Eigenvalues of $\mathbf{A}$, representing the observable A:

$$\mathbf{A}\vec{\phi}_i = \lambda_i \vec{\phi}_i$$

- **Probability** to measure λ_n if the system is in state $\vec{\psi} = \sum_i \psi_i \vec{\phi}_i$ is

$$Pr(A = \lambda_n, \vec{\psi}) = \|\vec{\psi}_n\|^2$$

Postulates for Quantum Mechanics (ca 1950)

- The quantum **state** of a (free) particle is described by a (normalised) complex valued function:

$$\vec{\psi} \in L^2(x) \text{ i.e. } \int |\psi(x)|^2 dx = 1$$

- Two quantum states can be **superimposed**, i.e.

$$\alpha_1 \vec{\psi}_1 + \alpha_2 \vec{\psi}_2$$

- Any **observable** A is represented by a linear, self-adjoint operator **A** on $L^2(x)$.
- **Possible** measurement results: Eigenvalues of **A**, representing the observable A:

$$\mathbf{A}\vec{\phi}_i = \lambda_i \vec{\phi}_i$$

- **Probability** to measure λ_n if the system is in state $\vec{\psi} = \sum_i \psi_i \vec{\phi}_i$ is

$$Pr(A = \lambda_n, \vec{\psi}) = \|\vec{\psi}_n\|^2$$

Postulates for Quantum Mechanics (ca 1950)

- The quantum **state** of a (free) particle is described by a (normalised) complex valued function:

$$\vec{\psi} \in L^2(x) \text{ i.e. } \int |\psi(x)|^2 dx = 1$$

- Two quantum states can be **superimposed**, i.e.

$$\alpha_1 \vec{\psi}_1 + \alpha_2 \vec{\psi}_2$$

- Any **observable** A is represented by a linear, self-adjoint operator **A** on $L^2(x)$.
- **Possible** measurement results: Eigenvalues of **A**, representing the observable A:

$$\mathbf{A}\vec{\phi}_i = \lambda_i \vec{\phi}_i$$

- **Probability** to measure λ_n if the system is in state $\vec{\psi} = \sum_i \psi_i \vec{\phi}_i$ is

$$Pr(A = \lambda_n, \vec{\psi}) = \|\vec{\psi}_n\|^2$$

Postulates for Quantum Mechanics (ca 1950)

- The quantum **state** of a (free) particle is described by a (normalised) complex valued function:

$$\vec{\psi} \in L^2(x) \text{ i.e. } \int |\psi(x)|^2 dx = 1$$

- Two quantum states can be **superimposed**, i.e.

$$\alpha_1 \vec{\psi}_1 + \alpha_2 \vec{\psi}_2$$

- Any **observable** A is represented by a linear, self-adjoint operator **A** on $L^2(x)$.
- **Possible** measurement results: Eigenvalues of **A**, representing the observable A:

$$\mathbf{A}\vec{\phi}_i = \lambda_i \vec{\phi}_i$$

- **Probability** to measure λ_n if the system is in state $\vec{\psi} = \sum_i \psi_i \vec{\phi}_i$ is

$$Pr(A = \lambda_n, \vec{\psi}) = \|\vec{\psi}_n\|^2$$

Postulates for Quantum Mechanics

Observables and **states** of a system are represented by *hermitian* (i.e. self-adjoint) elements *a* of a C*-algebra $\mathcal{A}$ and by *states* *w* (i.e. normalised linear functionals) over this algebra.

Possible results of **measurements** of an observable *a* are given by the *spectrum* $\mathrm{Sp}(a)$ of an observable. Their probability distribution in a certain state *w* is given by the probability measure $\mu(w)$ induced by the state *w* on $\mathrm{Sp}(a)$.

Walter Thirring: *Quantum Mathematical Physics*, 2nd ed. Springer Verlag, 2002

Postulates for Quantum Mechanics

Observables and **states** of a system are represented by *hermitian* (i.e. self-adjoint) elements *a* of a C*-algebra $\mathcal{A}$ and by *states* *w* (i.e. normalised linear functionals) over this algebra.

Possible results of **measurements** of an observable *a* are given by the *spectrum* $\mathrm{Sp}(a)$ of an observable. Their probability distribution in a certain state *w* is given by the probability measure $\mu(w)$ induced by the state *w* on $\mathrm{Sp}(a)$.

Walter Thirring: *Quantum Mathematical Physics*, 2nd ed. Springer Verlag, 2002

Postulates for Quantum Mechanics

Observables and **states** of a system are represented by *hermitian* (i.e. self-adjoint) elements *a* of a C*-algebra $\mathcal{A}$ and by *states* *w* (i.e. normalised linear functionals) over this algebra.

Possible results of **measurements** of an observable *a* are given by the *spectrum* $\mathrm{Sp}(a)$ of an observable. Their probability distribution in a certain state *w* is given by the probability measure $\mu(w)$ induced by the state *w* on $\mathrm{Sp}(a)$.

Walter Thirring: *Quantum Mathematical Physics*, 2nd ed. Springer Verlag, 2002

Quantum Mathematics

Quantum physics is often/sometimes counter-intuitive.

However, the standard mathematical model of (closed) quantum systems is relatively simple and just requires some basic **(complex) linear algebra**.

- The information describing the state of an (isolated) quantum mechanical system is represented mathematically by a (normalised) vector in a complex vector Hilbert space $\mathcal{H}$.
- An observable are represented mathematically by a self-adjoint matrix (operator) $\mathbf{A}$ acting on $\mathcal{H}$.

Two states can be combined to form a new state $\alpha |x\rangle + \beta |y\rangle$ as long as $|\alpha|^2 + |\beta|^2 = 1$ (Superposition).

Quantum Mathematics

Quantum physics is often/sometimes counter-intuitive.

However, the standard mathematical model of (closed) quantum systems is relatively simple and just requires some basic **(complex) linear algebra**.

- The information describing the state of an (isolated) quantum mechanical system is represented mathematically by a (normalised) vector in a complex vector Hilbert space $\mathcal{H}$.
- An observable are represented mathematically by a self-adjoint matrix (operator) **A** acting on $\mathcal{H}$.

Two states can be combined to form a new state $\alpha |x\rangle + \beta |y\rangle$ as long as $|\alpha|^2 + |\beta|^2 = 1$ (Superposition).

Quantum Mathematics

Quantum physics is often/sometimes counter-intuitive.

However, the standard mathematical model of (closed) quantum systems is relatively simple and just requires some basic **(complex) linear algebra**.

- The information describing the state of an (isolated) quantum mechanical system is represented mathematically by a (normalised) vector in a complex vector Hilbert space $\mathcal{H}$.
- An observable are represented mathematically by a self-adjoint matrix (operator) $\mathbf{A}$ acting on $\mathcal{H}$.

Two states can be combined to form a new state $\alpha|x\rangle + \beta|y\rangle$ as long as $|\alpha|^2 + |\beta|^2 = 1$ (Superposition).

Quantum Mathematics

Quantum physics is often/sometimes counter-intuitive.

However, the standard mathematical model of (closed) quantum systems is relatively simple and just requires some basic **(complex) linear algebra**.

- The information describing the state of an (isolated) quantum mechanical system is represented mathematically by a (normalised) vector in a complex vector Hilbert space $\mathcal{H}$.
- An observable are represented mathematically by a self-adjoint matrix (operator) **A** acting on $\mathcal{H}$.

Two states can be combined to form a new state $\alpha |x\rangle + \beta |y\rangle$ as long as $|\alpha|^2 + |\beta|^2 = 1$ (Superposition).

Quantum Mathematics

Quantum physics is often/sometimes counter-intuitive.

However, the standard mathematical model of (closed) quantum systems is relatively simple and just requires some basic **(complex) linear algebra**.

- The information describing the state of an (isolated) quantum mechanical system is represented mathematically by a (normalised) vector in a complex vector Hilbert space $\mathcal{H}$.
- An observable are represented mathematically by a self-adjoint matrix (operator) $\mathbf{A}$ acting on $\mathcal{H}$.

Two states can be combined to form a new state $\alpha |x\rangle + \beta |y\rangle$ as long as $|\alpha|^2 + |\beta|^2 = 1$ (Superposition).

Quantum States and Notation

The state of a QM system is usually denoted by $|x\rangle \in \mathcal{H}$. The *inner product* $\langle x|y\rangle$ of two vectors in $\mathcal{H}$ – which is describing the *angle* between them – is very important in QM.

P.A.M. Dirac "invented" the Bra-Ket Notation based on the following simple facts:

Typewriters had no sub-scripts $\vec{x}_i$
Hilbert spaces have inner product

Simply "take inner product appart" to denote vectors in $\mathcal{H}$:

$$\langle x_i, y_j\rangle = \langle x_i|y_j\rangle = \langle i|\,|j\rangle$$

Quantum States and Notation

The state of a QM system is usually denoted by $|x\rangle \in \mathcal{H}$. The inner product $\langle x|y\rangle$ of two vectors in $\mathcal{H}$ – which is describing the *angle* between them – is very important in QM.

P.A.M. Dirac "invented" the Bra-Ket Notation based on the following simple facts:

Typewriters had no sub-scripts $\vec{x}_i$
Hilbert spaces have inner product

Simply "take inner product appart" to denote vectors in $\mathcal{H}$:

$$\langle x_i, y_j \rangle = \langle x_i | y_j \rangle = \langle i | \, | j \rangle$$

Quantum States and Notation

The state of a QM system is usually denoted by $|x\rangle \in \mathcal{H}$. The inner product $\langle x|y\rangle$ of two vectors in $\mathcal{H}$ – which is describing the *angle* between them – is very important in QM.

P.A.M. Dirac "invented" the Bra-Ket Notation based on the following simple facts:

Typewriters had no sub-scripts $\vec{x}_i$
Hilbert spaces have inner product

Simply "take inner product appart" to denote vectors in $\mathcal{H}$:

$$\langle x_i, y_j \rangle = \langle x_i | y_j \rangle = \langle i| \, |j\rangle$$

Quantum States and Notation

The state of a QM system is usually denoted by $|x\rangle \in \mathcal{H}$. The inner product $\langle x|y\rangle$ of two vectors in $\mathcal{H}$ – which is describing the *angle* between them – is very important in QM.

P.A.M. Dirac "invented" the Bra-Ket Notation based on the following simple facts:

Typewriters had no sub-scripts $\vec{x}_i$
Hilbert spaces have inner product

Simply "take inner product appart" to denote vectors in $\mathcal{H}$:

$$\left\langle x_i, y_j \right\rangle = \left\langle x_i | y_j \right\rangle = \langle i| \, |j\rangle$$

Conventions

Physical Convention:

$$\langle x | \alpha y \rangle = \alpha \langle x | y \rangle$$

Mathematical Convention:

$$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$$

Linear in first or second argument? In mathematics we have:

$$\langle x, \alpha y \rangle = \overline{\langle \alpha y, x \rangle} = \bar{\alpha} \overline{\langle y, x \rangle} = \bar{\alpha} \langle x, y \rangle$$

Conventions

Physical Convention:

$$\langle x|\alpha y\rangle = \alpha \langle x|y\rangle$$

Mathematical Convention:

$$\langle \alpha x, y\rangle = \alpha \langle x, y\rangle$$

Linear in first or second argument? In mathematics we have:

$$\langle x, \alpha y\rangle = \overline{\langle \alpha y, x\rangle} = \bar{\alpha}\overline{\langle y, x\rangle} = \bar{\alpha} \langle x, y\rangle$$

Conventions

Physical Convention:

$$\langle x | \alpha y \rangle = \alpha \langle x | y \rangle$$

Mathematical Convention:

$$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$$

Linear in first or second argument? In mathematics we have:

$$\langle x, \alpha y \rangle = \overline{\langle \alpha y, x \rangle} = \bar{\alpha} \overline{\langle y, x \rangle} = \bar{\alpha} \langle x, y \rangle$$

Quantum Measurement

- The expected result of measuring $\mathbf{A}$ of a system in state $|x\rangle \in \mathcal{H}$ is given by:

$$\langle A \rangle_x = \langle x | \mathbf{A} | x \rangle = \langle x | | \mathbf{A} x \rangle$$

- The only possible results are eigenvalues λ_i of $\mathbf{A}$.
- The probability of measuring λ_n in state $|x\rangle$ is

$$Pr(A = \lambda_n, x) = \langle x | \mathbf{P}_n | x \rangle$$

with $\mathbf{P}_n$ the (orthogonal) projection (s.t. $\mathbf{A} = \sum_i \lambda_i \mathbf{P}_i$)

$$\mathbf{P}_n = \sum_{j=1}^{d(n)} |\lambda_n, j\rangle \langle \lambda_n, j|$$

Quantum Measurement

- The expected result of measuring **A** of a system in state $|x\rangle \in \mathcal{H}$ is given by:

$$\langle A \rangle_x = \langle x | \mathbf{A} | x \rangle = \langle x | | \mathbf{A} x \rangle$$

- The only possible results are eigenvalues λ_i of **A**.
- The probability of measuring λ_n in state $|x\rangle$ is

$$Pr(A = \lambda_n, x) = \langle x | \mathbf{P}_n | x \rangle$$

with $\mathbf{P}_n$ the (orthogonal) projection (s.t. $\mathbf{A} = \sum_i \lambda_i \mathbf{P}_i$)

$$\mathbf{P}_n = \sum_{j=1}^{d(n)} |\lambda_n, j\rangle \langle \lambda_n, j|$$

Quantum Measurement

- The expected result of measuring **A** of a system in state $|x\rangle \in \mathcal{H}$ is given by:

$$\langle A \rangle_x = \langle x | \mathbf{A} | x \rangle = \langle x | | \mathbf{A} x \rangle$$

- The only possible results are eigenvalues λ_i of **A**.
- The probability of measuring λ_n in state $|x\rangle$ is

$$Pr(A = \lambda_n, x) = \langle x | \mathbf{P}_n | x \rangle$$

with $\mathbf{P}_n$ the (orthogonal) projection (s.t. $\mathbf{A} = \sum_i \lambda_i \mathbf{P}_i$)

$$\mathbf{P}_n = \sum_{j=1}^{d(n)} |\lambda_n, j\rangle \langle \lambda_n, j|$$

Quantum Dynamics

- The **dynamics** of a closed system is described by the Schrödinger Equation:

$$i\hbar \frac{d\,|x\rangle}{dt} = \mathbf{H}\,|x\rangle$$

for the (self-adjoint) Hamiltonian **H**.

- The **solution** is a unitary operator $\mathbf{U}_t = \exp(it\mathbf{H})$.

Theorem

For any self-adjoint operator **A** *the operator*

$$\exp(i\mathbf{A}) = e^{i\mathbf{A}} = \sum_{n=0}^{\infty} \frac{(i\mathbf{A})^n}{n!}$$

is a unitary operator.

Quantum Dynamics

- The **dynamics** of a closed system is described by the Schrödinger Equation:

$$i\hbar \frac{d\,|x\rangle}{dt} = \mathbf{H}\,|x\rangle$$

for the (self-adjoint) Hamiltonian **H**.

- The **solution** is a unitary operator $\mathbf{U}_t = \exp(it\mathbf{H})$.

Theorem

For any self-adjoint operator **A** *the operator*

$$\exp(i\mathbf{A}) = e^{i\mathbf{A}} = \sum_{n=0}^{\infty} \frac{(i\mathbf{A})^n}{n!}$$

is a unitary operator.

Quantum Dynamics

- The **dynamics** of a closed system is described by the Schrödinger Equation:

$$i\hbar \frac{d\,|x\rangle}{dt} = \mathbf{H}\,|x\rangle$$

  for the (self-adjoint) Hamiltonian **H**.
- The **solution** is a unitary operator $\mathbf{U}_t = \exp(it\mathbf{H})$.

Theorem

For any self-adjoint operator **A** *the operator*

$$\exp(i\mathbf{A}) = e^{i\mathbf{A}} = \sum_{n=0}^{\infty} \frac{(i\mathbf{A})^n}{n!}$$

is a unitary operator.

Adjoint Operator

For a matrix $\mathbf{A} = (\mathbf{A}_{ij})$ its **transpose** matrix $\mathbf{A}^T$ is defined as

$$(\mathbf{A}^T_{ij}) = (\mathbf{A}_{ji})$$

the conjugate matrix $\mathbf{A}^*$ is defined by

$$(\mathbf{A}^*_{ij}) = (\mathbf{A}_{ij})^*$$

and the adjoint matrix $\mathbf{A}^\dagger$ is given via

$$(\mathbf{A}^\dagger_{ij}) = (\mathbf{A}^*_{ji}) \quad \text{or} \quad \mathbf{A}^\dagger = (\mathbf{A}^*)^T$$

Notation: In **mathematics** the adjoint operator is usually denoted by $\mathbf{A}^*$ and defined implicitly via:

$$\langle \mathbf{A}(x), y \rangle = \langle x, \mathbf{A}^*(y) \rangle \quad \text{or} \quad \langle \mathbf{A}^\dagger x | y \rangle = \langle x, \mathbf{A} y \rangle$$

Adjoint Operator

For a matrix $\mathbf{A} = (\mathbf{A}_{ij})$ its **transpose** matrix $\mathbf{A}^T$ is defined as

$$(\mathbf{A}_{ij}^T) = (\mathbf{A}_{ji})$$

the conjugate matrix $\mathbf{A}^*$ is defined by

$$(\mathbf{A}_{ij}^*) = (\mathbf{A}_{ij})^*$$

and the adjoint matrix $\mathbf{A}^\dagger$ is given via

$$(\mathbf{A}_{ij}^\dagger) = (\mathbf{A}_{ji}^*) \text{ or } \mathbf{A}^\dagger = (\mathbf{A}^*)^T$$

Notation: In **mathematics** the adjoint operator is usually denoted by $\mathbf{A}^*$ and defined implicitly via:

$$\langle \mathbf{A}(x), y \rangle = \langle x, \mathbf{A}^*(y) \rangle \quad \text{or} \quad \langle \mathbf{A}^\dagger x | y \rangle = \langle x, \mathbf{A} y \rangle$$

Adjoint Operator

For a matrix $\mathbf{A} = (\mathbf{A}_{ij})$ its **transpose** matrix $\mathbf{A}^T$ is defined as

$$(\mathbf{A}^T_{ij}) = (\mathbf{A}_{ji})$$

the conjugate matrix $\mathbf{A}^*$ is defined by

$$(\mathbf{A}^*_{ij}) = (\mathbf{A}_{ij})^*$$

and the adjoint matrix $\mathbf{A}^\dagger$ is given via

$$(\mathbf{A}^\dagger_{ij}) = (\mathbf{A}^*_{ji}) \text{ or } \mathbf{A}^\dagger = (\mathbf{A}^*)^T$$

Notation: In **mathematics** the adjoint operator is usually denoted by $\mathbf{A}^*$ and defined implicitly via:

$$\langle \mathbf{A}(x), y\rangle = \langle x, \mathbf{A}^*(y)\rangle \text{ or } \langle \mathbf{A}^\dagger x | y\rangle = \langle x, \mathbf{A}y\rangle$$

Adjoint Operator

For a matrix $\mathbf{A} = (\mathbf{A}_{ij})$ its **transpose** matrix $\mathbf{A}^T$ is defined as

$$(\mathbf{A}^T_{ij}) = (\mathbf{A}_{ji})$$

the conjugate matrix $\mathbf{A}^*$ is defined by

$$(\mathbf{A}^*_{ij}) = (\mathbf{A}_{ij})^*$$

and the adjoint matrix $\mathbf{A}^\dagger$ is given via

$$(\mathbf{A}^\dagger_{ij}) = (\mathbf{A}^*_{ji}) \text{ or } \mathbf{A}^\dagger = (\mathbf{A}^*)^T$$

Notation: In **mathematics** the adjoint operator is usually denoted by $\mathbf{A}^*$ and defined implicitly via:

$$\langle \mathbf{A}(x), y \rangle = \langle x, \mathbf{A}^*(y) \rangle \text{ or } \langle \mathbf{A}^\dagger x | y \rangle = \langle x, \mathbf{A} y \rangle$$

Unitary Operators

A square matrix/operator **U** is called **unitary** if

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{I} = \mathbf{U}\mathbf{U}^\dagger$$

That means **U**'s inverse is $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. It also implies that **U** is **invertible** and the inverse is easy to compute.

The postulates of **Quantum Mechanics** require that the **time evolution** to a quantum state, e.g. a qubit, are implemented via a unitary operator (as long as there is no measurement).

The unitary evolution of an (isolated) quantum state/system is a mathematical consequence of being a solution of the Schrödinger equation for some Hamiltonian operator **H**.

Unitary Operators

A square matrix/operator $\mathbf{U}$ is called **unitary** if

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{I} = \mathbf{U}\mathbf{U}^\dagger$$

That means $\mathbf{U}$'s inverse is $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. It also implies that $\mathbf{U}$ is **invertible** and the inverse is easy to compute.

The postulates of **Quantum Mechanics** require that the **time evolution** to a quantum state, e.g. a qubit, are implemented via a unitary operator (as long as there is no measurement).

The unitary evolution of an (isolated) quantum state/system is a mathematical consequence of being a solution of the Schrödinger equation for some Hamiltonian operator $\mathbf{H}$.

Unitary Operators

A square matrix/operator $\mathbf{U}$ is called **unitary** if

$$\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I} = \mathbf{U}\mathbf{U}^{\dagger}$$

That means $\mathbf{U}$'s inverse is $\mathbf{U}^{\dagger} = \mathbf{U}^{-1}$. It also implies that $\mathbf{U}$ is **invertible** and the inverse is easy to compute.

The postulates of **Quantum Mechanics** require that the **time evolution** to a quantum state, e.g. a qubit, are implemented via a unitary operator (as long as there is no measurement).

The unitary evolution of an (isolated) quantum state/system is a mathematical consequence of being a solution of the Schrödinger equation for some Hamiltonian operator $\mathbf{H}$.

Unitary Operators

A square matrix/operator **U** is called **unitary** if

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{I} = \mathbf{U}\mathbf{U}^\dagger$$

That means **U**'s inverse is $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. It also implies that **U** is **invertible** and the inverse is easy to compute.

The postulates of **Quantum Mechanics** require that the **time evolution** to a quantum state, e.g. a qubit, are implemented via a unitary operator (as long as there is no measurement).

The unitary evolution of an (isolated) quantum state/system is a mathematical consequence of being a solution of the Schrödinger equation for some Hamiltonian operator **H**.

Self Adjoint Operators

An operator **A** is called **self-adjoint** or **hermitian** iff

$$\mathbf{A} = \mathbf{A}^\dagger$$

The postulates of **Quantum Mechanics** require that a quantum **observable** A is represented by a self-adjoint operator **A**.

Possible measurement results are **eigenvalues** λ_i of **A** defined as

$$\mathbf{A}|i\rangle = \lambda_i |i\rangle \quad \text{or} \quad \mathbf{A}\vec{a}_i = \lambda_i \vec{a}_i$$

Probability to observe λ_k in state $|x\rangle = \sum_i \alpha_i |i\rangle$ is

$$Pr(A = \lambda_k, |x\rangle) = |\alpha_k|^2$$

Self Adjoint Operators

An operator **A** is called **self-adjoint** or **hermitian** iff

$$\mathbf{A} = \mathbf{A}^\dagger$$

The postulates of **Quantum Mechanics** require that a quantum **observable** A is represented by a self-adjoint operator **A**.

Possible measurement results are **eigenvalues** λ_i of **A** defined as

$$\mathbf{A}|i\rangle = \lambda_i |i\rangle \quad \text{or} \quad \mathbf{A}\vec{a}_i = \lambda_i \vec{a}_i$$

Probability to observe λ_k in state $|x\rangle = \sum_i \alpha_i |i\rangle$ is

$$Pr(A = \lambda_k, |x\rangle) = |\alpha_k|^2$$

Self Adjoint Operators

An operator $\mathbf{A}$ is called **self-adjoint** or **hermitian** iff

$$\mathbf{A} = \mathbf{A}^\dagger$$

The postulates of **Quantum Mechanics** require that a quantum **observable** A is represented by a self-adjoint operator $\mathbf{A}$.

Possible measurement results are **eigenvalues** λ_i of $\mathbf{A}$ defined as

$$\mathbf{A}\left|i\right\rangle = \lambda_i \left|i\right\rangle \quad \text{or} \quad \mathbf{A}\vec{a}_i = \lambda_i \vec{a}_i$$

Probability to observe λ_k in state $\left|x\right\rangle = \sum_i \alpha_i \left|i\right\rangle$ is

$$Pr(A = \lambda_k, \left|x\right\rangle) = |\alpha_k|^2$$

Self Adjoint Operators

An operator **A** is called **self-adjoint** or **hermitian** iff

$$\mathbf{A} = \mathbf{A}^\dagger$$

The postulates of **Quantum Mechanics** require that a quantum **observable** A is represented by a self-adjoint operator **A.**

Possible measurement results are **eigenvalues** λ_i of **A** defined as

$$\mathbf{A}\,|i\rangle = \lambda_i\,|i\rangle \quad \text{or} \quad \mathbf{A}\vec{a}_i = \lambda_i \vec{a}_i$$

Probability to observe λ_k in state $|x\rangle = \sum_i \alpha_i\,|i\rangle$ is

$$Pr(A = \lambda_k, |x\rangle) = |\alpha_k|^2$$

Projections
An operator $\mathbf{P}$ on $\mathbb{C}^n$ is called **projection** (or **idempotent**) iff

$$\mathbf{P}^2 = \mathbf{PP} = \mathbf{P}$$

Orthogonal Projection
An operator $\mathbf{P}$ on $\mathbb{C}^n$ is called **(orthogonal) projection** iff

$$\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^\dagger$$

We say that an (orthogonal) projection $\mathbf{P}$ projects **onto** its image space $\mathbf{P}(\mathbb{C}^n)$, which is always a linear sub-spaces of $\mathbb{C}^n$.

Birkhoff-von Neumann: Projection on a Hilbert space form an ortho-lattice which gives rise to non-classical a “quantum logic”.

Projections

An operator **P** on $\mathbb{C}^n$ is called **projection** (or **idempotent**) iff

$$\mathbf{P}^2 = \mathbf{P}\mathbf{P} = \mathbf{P}$$

Orthogonal Projection

An operator **P** on $\mathbb{C}^n$ is called **(orthogonal) projection** iff

$$\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^\dagger$$

We say that an (orthogonal) projection **P** projects **onto** its image space $\mathbf{P}(\mathbb{C}^n)$, which is always a linear sub-spaces of $\mathbb{C}^n$.

Birkhoff-von Neumann: Projection on a Hilbert space form an ortho-lattice which gives rise to non-classical a "quantum logic".

Projections

An operator **P** on $\mathbb{C}^n$ is called **projection** (or **idempotent**) iff

$$\mathbf{P}^2 = \mathbf{P}\mathbf{P} = \mathbf{P}$$

Orthogonal Projection

An operator **P** on $\mathbb{C}^n$ is called **(orthogonal) projection** iff

$$\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^\dagger$$

We say that an (orthogonal) projection **P** projects **onto** its image space $\mathbf{P}(\mathbb{C}^n)$, which is always a linear sub-spaces of $\mathbb{C}^n$.

Birkhoff-von Neumann: Projection on a Hilbert space form an ortho-lattice which gives rise to non-classical a "quantum logic".

Projections

An operator **P** on $\mathbb{C}^n$ is called **projection** (or **idempotent**) iff

$$\mathbf{P}^2 = \mathbf{P}\mathbf{P} = \mathbf{P}$$

Orthogonal Projection

An operator **P** on $\mathbb{C}^n$ is called **(orthogonal) projection** iff

$$\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^\dagger$$

We say that an (orthogonal) projection **P** projects **onto** its image space $\mathbf{P}(\mathbb{C}^n)$, which is always a linear sub-spaces of $\mathbb{C}^n$.

Birkhoff-von Neumann: Projection on a Hilbert space form an ortho-lattice which gives rise to non-classical a “quantum logic”.

Spectral Theorem

In the bra-ket notation we can represent a projection onto the sub-space generated by $|x\rangle$ by the outer product $\mathbf{P}_x = |x\rangle\langle x|$.

Theorem

A self-adjoint operator $\mathbf{A}$ *(on a finite dimensional Hilbert space, e.g.* $\mathbb{C}^n$*) can be represented uniquely as a linear combination*

$$\mathbf{A} = \sum_i \lambda_i \mathbf{P}_i$$

with $\lambda_i \in \mathbb{R}$ *and* $\mathbf{P}_i$ *the (orthogonal) projection onto the eigen-space generated by the eigen-vector* $|i\rangle$*, i.e.* $\mathbf{P}_i = |i\rangle\langle i|$

In the degenerate case we had to consider: $P_i = \sum_{j=1}^{d(n)} |i_j\rangle\langle i_j|$.

In the bra-ket notation we can represent a projection onto the sub-space generated by $|x\rangle$ by the outer product $\mathbf{P}_x = |x\rangle\langle x|$.

Theorem

A self-adjoint operator $\mathbf{A}$ *(on a finite dimensional Hilbert space, e.g.* $\mathbb{C}^n$*) can be represented uniquely as a linear combination*

$$\mathbf{A} = \sum_i \lambda_i \mathbf{P}_i$$

with $\lambda_i \in \mathbb{R}$ *and* $\mathbf{P}_i$ *the (orthogonal) projection onto the eigen-space generated by the eigen-vector* $|i\rangle$*, i.e.* $\mathbf{P}_i = |i\rangle\langle i|$

In the degenerate case we had to consider: $P_i = \sum_{j=1}^{d(n)} |i_j\rangle\langle i_j|$.

Spectral Theorem

In the bra-ket notation we can represent a projection onto the sub-space generated by $|x\rangle$ by the outer product $\mathbf{P}_x = |x\rangle\langle x|$.

Theorem

A self-adjoint operator $\mathbf{A}$ *(on a finite dimensional Hilbert space, e.g.* $\mathbb{C}^n$*) can be represented uniquely as a linear combination*

$$\mathbf{A} = \sum_i \lambda_i \mathbf{P}_i$$

with $\lambda_i \in \mathbb{R}$ *and* $\mathbf{P}_i$ *the (orthogonal) projection onto the eigen-space generated by the eigen-vector* $|i\rangle$*, i.e.* $\mathbf{P}_i = |i\rangle\langle i|$

In the degenerate case we had to consider: $P_i = \sum_{j=1}^{d(n)} |i_j\rangle\langle i_j|$.

Measurement Process

If we perform a measurement of the observable represented by:

$$\mathbf{A} = \sum_i \lambda_i |i\rangle\langle i|$$

with eigen-values λ_i and eigen-vectors $|i\rangle$ in a state $|x\rangle$ we have to decompose the state according to the observable, i.e.

$$|x\rangle = \sum_i \mathbf{P}_i |x\rangle = \sum_i |i\rangle\langle i|x\rangle = \sum_i \langle i|x\rangle |i\rangle = \sum_i \alpha_i |i\rangle$$

With probability $|\alpha_i|^2 = |\langle i|x\rangle|^2$ two things happen

- The measurement instrument will the **display** λ_i.
- The state $|x\rangle$ **collapses** to $|i\rangle$.

Measurement Process

If we perform a measurement of the observable represented by:

$$\mathbf{A} = \sum_i \lambda_i |i\rangle\langle i|$$

with eigen-values λ_i and eigen-vectors $|i\rangle$ in a state $|x\rangle$ we have to decompose the state according to the observable, i.e.

$$|x\rangle = \sum_i \mathbf{P}_i |x\rangle = \sum_i |i\rangle\langle i|x\rangle = \sum_i \langle i|x\rangle |i\rangle = \sum_i \alpha_i |i\rangle$$

With probability $|\alpha_i|^2 = |\langle i|x\rangle|^2$ two things happen

- The measurement instrument will the **display** λ_i.
- The state $|x\rangle$ **collapses** to $|i\rangle$.

Measurement Process

If we perform a measurement of the observable represented by:

$$\mathbf{A} = \sum_i \lambda_i |i\rangle\langle i|$$

with eigen-values λ_i and eigen-vectors $|i\rangle$ in a state $|x\rangle$ we have to decompose the state according to the observable, i.e.

$$|x\rangle = \sum_i \mathbf{P}_i |x\rangle = \sum_i |i\rangle\langle i|x\rangle = \sum_i \langle i|x\rangle |i\rangle = \sum_i \alpha_i |i\rangle$$

With probability $|\alpha_i|^2 = |\langle i|x\rangle|^2$ two things happen

- The measurement instrument will the **display** λ_i.
- The state $|x\rangle$ **collapses** to $|i\rangle$.

Measurement Process

If we perform a measurement of the observable represented by:

$$\mathbf{A} = \sum_i \lambda_i |i\rangle\langle i|$$

with eigen-values λ_i and eigen-vectors $|i\rangle$ in a state $|x\rangle$ we have to decompose the state according to the observable, i.e.

$$|x\rangle = \sum_i \mathbf{P}_i |x\rangle = \sum_i |i\rangle\langle i|x\rangle = \sum_i \langle i|x\rangle |i\rangle = \sum_i \alpha_i |i\rangle$$

With probability $|\alpha_i|^2 = |\langle i|x\rangle|^2$ two things happen

- The measurement instrument will the **display** λ_i.
- The state $|x\rangle$ **collapses** to $|i\rangle$.

Spectrum

The set of eigen-values $\{\lambda_1, \lambda_2, \ldots\}$ of an operator **A** is called its spectrum $\sigma(\mathbf{A})$.

$$\sigma(\mathbf{A}) = \{\lambda \mid \lambda\mathbf{I} - \mathbf{A} \text{ is not invertible}\}$$

It is possible that for an eigen-value λ_i in the equation

$$\mathbf{A}\,|i\rangle = \lambda_i\,|i\rangle$$

we may have more than one eigen-vector $|i\rangle$, i.e. the dimension of the eigen-space $d(n) > 1$. We will not consider these **degenerate** cases here.

Spectrum

The set of eigen-values $\{\lambda_1, \lambda_2, \ldots\}$ of an operator **A** is called its spectrum $\sigma(\mathbf{A})$.

$$\sigma(\mathbf{A}) = \{\lambda \mid \lambda\mathbf{I} - \mathbf{A} \text{ is not invertible}\}$$

It is possible that for an eigen-value λ_i in the equation

$$\mathbf{A}\,|i\rangle = \lambda_i\,|i\rangle$$

we may have more than one eigen-vector $|i\rangle$, i.e. the dimension of the eigen-space $d(n) > 1$. We will not consider these **degenerate** cases here.

Heisenberg and Schrödinger Picture

Describe the dynamics in terms of observables or states.

In particular if we consider not just pure (isolated) states, i.e. vectors in a Hilbert space, but instead probabilistic states which are repesented by density matrices.

A density matrix $\rho \in \mathcal{B}(\mathcal{H})$ is a Hermitian semi-positive definite matrix or operator with trace$(\rho) = 1$. Note that a given pure state $|\psi\rangle$ can also be represented with density matrix $|\psi\rangle\langle\psi|$.

The quantum dynamics can be described as for **A** observable and ρ state (as density matrix/operator).

Schrödinger Picture: $\varrho_t = \mathbf{U}(t)\varrho_0\mathbf{U}^*(t)$.

Heisenberg Picture: $\mathbf{A}_t = \mathbf{U}^*(t)\mathbf{A}_0\mathbf{U}(t)$.

Heisenberg and Schrödinger Picture

Describe the dynamics in terms of observables or states.

In particular if we consider not just pure (isolated) states, i.e. vectors in a Hilbert space, but instead probabilistic states which are repesented by density matrices.

A density matrix $\rho \in \mathcal{B}(\mathcal{H})$ is a Hermitian semi-positive definite matrix or operator with trace$(\rho) = 1$. Note that a given pure state $|\psi\rangle$ can also be represented with density matrix $|\psi\rangle\langle\psi|$.

The quantum dynamics can be described as for **A** observable and ρ state (as density matrix/operator).

Schrödinger Picture: $\varrho_t = \mathbf{U}(t)\varrho_0\mathbf{U}^*(t)$.

Heisenberg Picture: $\mathbf{A}_t = \mathbf{U}^*(t)\mathbf{A}_0\mathbf{U}(t)$.

Heisenberg and Schrödinger Picture

Describe the dynamics in terms of observables or states.

In particular if we consider not just pure (isolated) states, i.e. vectors in a Hilbert space, but instead probabilistic states which are repesented by density matrices.

A density matrix $\rho \in \mathcal{B}(\mathcal{H})$ is a Hermitian semi-positive definite matrix or operator with trace$(\rho) = 1$. Note that a given pure state $|\psi\rangle$ can also be represented with density matrix $|\psi\rangle\langle\psi|$.

The quantum dynamics can be described as for **A** observable and ρ state (as density matrix/operator).

Schrödinger Picture: $\varrho_t = \mathbf{U}(t)\varrho_0\mathbf{U}^*(t)$.

Heisenberg Picture: $\mathbf{A}_t = \mathbf{U}^*(t)\mathbf{A}_0\mathbf{U}(t)$.

Heisenberg and Schrödinger Picture

Describe the dynamics in terms of observables or states.

In particular if we consider not just pure (isolated) states, i.e. vectors in a Hilbert space, but instead probabilistic states which are repesented by density matrices.

A density matrix $\rho \in \mathcal{B}(\mathcal{H})$ is a Hermitian semi-positive definite matrix or operator with trace$(\rho) = 1$. Note that a given pure state $|\psi\rangle$ can also be represented with density matrix $|\psi\rangle\langle\psi|$.

The quantum dynamics can be described as for $\mathbf{A}$ observable and ρ state (as density matrix/operator).

Schrödinger Picture: $\varrho_t = \mathbf{U}(t)\varrho_0\mathbf{U}^*(t)$.

Heisenberg Picture: $\mathbf{A}_t = \mathbf{U}^*(t)\mathbf{A}_0\mathbf{U}(t)$.

Reversibility

Quantum Dynamics
For unitary transformations describing qubit dynamics:

$$\mathbf{U}^{\dagger} = \mathbf{U}^{-1}$$

The quantum dynamics is **invertible** or **reversible**

Quantum Measurement
For projection operators involved in quantum measurement:

$$\mathbf{P}^{\dagger} \neq \mathbf{P}^{-1}$$

The quantum measurement is not **reversible.** However

$$\mathbf{P}^{2} = \mathbf{P}$$

The quantum measurement is **idempotent.**

Reversibility

Quantum Dynamics
For unitary transformations describing qubit dynamics:

$$\mathbf{U}^{\dagger} = \mathbf{U}^{-1}$$

The quantum dynamics is **invertible** or **reversible**

Quantum Measurement
For projection operators involved in quantum measurement:

$$\mathbf{P}^{\dagger} \neq \mathbf{P}^{-1}$$

The quantum measurement is not **reversible.** However

$$\mathbf{P}^{2} = \mathbf{P}$$

The quantum measurement is **idempotent.**

Reversibility

Quantum Dynamics
For unitary transformations describing qubit dynamics:

$$\mathbf{U}^{\dagger} = \mathbf{U}^{-1}$$

The quantum dynamics is **invertible** or **reversible**

Quantum Measurement
For projection operators involved in quantum measurement:

$$\mathbf{P}^{\dagger} \neq \mathbf{P}^{-1}$$

The quantum measurement is not **reversible.** However

$$\mathbf{P}^{2} = \mathbf{P}$$

The quantum measurement is **idempotent.**

Dynamics of Programs

pWhile – Syntax I

Full programs contain optional variable declarations:

$$
\begin{array}{lcl}
P & ::= & \textbf{begin } S \textbf{ end} \\
 & \mid & \textbf{var } D \textbf{ begin } S \textbf{ end}
\end{array}
$$

Declarations are of the form:

$$
\begin{array}{lcl}
r & ::= & \textbf{bool} \\
 & \mid & \textbf{int} \\
 & \mid & \{\, c_1, \ldots, c_n \,\} \\
 & \mid & \{\, c_1 \textbf{ .. } c_n \,\} \\
D & ::= & v \textbf{ : } r \\
 & \mid & v \textbf{ : } r \textbf{ ; } D
\end{array}
$$

with c_i (integer) constants and r denoting ranges.

pWhile – Syntax I

Full programs contain optional variable declarations:

$$\begin{array}{lcl} P & ::= & \textbf{begin } S \textbf{ end} \\ & | & \textbf{var } D \textbf{ begin } S \textbf{ end} \end{array}$$

Declarations are of the form:

$$\begin{array}{lcl} r & ::= & \textbf{bool} \\ & | & \textbf{int} \\ & | & \{\, c_1, \ldots, c_n \,\} \\ & | & \{\, c_1 \,..\, c_n \,\} \\ D & ::= & v : r \\ & | & v : r \,;\, D \end{array}$$

with c_i (integer) constants and r denoting ranges.

pWhile – Syntax II

The syntax of statements S is as follows:

$$
\begin{array}{rcl}
S & ::= & \textbf{stop} \\
 & | & \textbf{skip} \\
 & | & v \; := a \\
 & | & v \; ?= r \\
 & | & S_1 \, ; \;\; S_2 \\
 & | & \textbf{choose } p_1 : S_1 \textbf{ or } p_2 : S_2 \textbf{ ro} \\
 & | & \textbf{if } b \textbf{ then } S_1 \textbf{ else } S_2 \textbf{ fi} \\
 & | & \textbf{while } b \textbf{ do } S \textbf{ od}
\end{array}
$$

Where the p_i are constants, representing choice probabilities.

pWhile – Syntax II

The syntax of statements S is as follows:

$$
\begin{array}{rcl}
S & ::= & [\mathbf{stop}]^\ell \\
 & | & [\mathbf{skip}]^\ell \\
 & | & [v := a]^\ell \\
 & | & [v \mathrel{?{=}} r]^\ell \\
 & | & S_1 ;\ S_2 \\
 & | & \mathbf{choose}^\ell\ p_1 : S_1\ \mathbf{or}\ p_2 : S_2\ \mathbf{ro} \\
 & | & \mathbf{if}\ [b]^\ell\ \mathbf{then}\ S_1\ \mathbf{else}\ S_2\ \mathbf{fi} \\
 & | & \mathbf{while}\ [b]^\ell\ \mathbf{do}\ S\ \mathbf{od}
\end{array}
$$

Where the p_i are constants, representing choice probabilities.

Evaluation of Expressions

$$\sigma \ni \textbf{State} = \textbf{Var} \to \textbf{Z} \uplus \textbf{T}$$

To illustrate approach consider only finite sub-range of **Z**.

Evaluation $\mathcal{E}$ of expressions e in state σ:

$$\begin{aligned} \mathcal{E}(n)\sigma &= n \\ \mathcal{E}(v)\sigma &= \sigma(v) \\ \mathcal{E}(a_1 \odot a_2)\sigma &= \mathcal{E}(a_1)\sigma \odot \mathcal{E}(a_2)\sigma \end{aligned}$$

$$\begin{aligned} \mathcal{E}(\textbf{true})\sigma &= \textbf{tt} \\ \mathcal{E}(\textbf{false})\sigma &= \textbf{ff} \\ \mathcal{E}(\textbf{not } b)\sigma &= \neg\mathcal{E}(b)\sigma \\ \ldots &= \ldots \end{aligned}$$

Evaluation of Expressions

$$\sigma \ni \textbf{State} = \textbf{Var} \to \textbf{Z} \uplus \textbf{T}$$

To illustrate approach consider only finite sub-range of **Z.**
Evaluation $\mathcal{E}$ of expressions e in state σ:

$$\begin{aligned} \mathcal{E}(n)\sigma &= n \\ \mathcal{E}(v)\sigma &= \sigma(v) \\ \mathcal{E}(a_1 \odot a_2)\sigma &= \mathcal{E}(a_1)\sigma \odot \mathcal{E}(a_2)\sigma \end{aligned}$$

$$\begin{aligned} \mathcal{E}(\textbf{true})\sigma &= \textbf{tt} \\ \mathcal{E}(\textbf{false})\sigma &= \textbf{ff} \\ \mathcal{E}(\textbf{not } b)\sigma &= \neg\mathcal{E}(b)\sigma \\ \ldots &= \ldots \end{aligned}$$

Evaluation of Expressions

$$\sigma \ni \textbf{State} = \textbf{Var} \rightarrow \textbf{Z} \uplus \textbf{T}$$

To illustrate approach consider only finite sub-range of **Z.**
Evaluation $\mathcal{E}$ of expressions e in state σ:

$$\begin{aligned} \mathcal{E}(n)\sigma &= n \\ \mathcal{E}(v)\sigma &= \sigma(v) \\ \mathcal{E}(a_1 \odot a_2)\sigma &= \mathcal{E}(a_1)\sigma \odot \mathcal{E}(a_2)\sigma \end{aligned}$$

$$\begin{aligned} \mathcal{E}(\textbf{true})\sigma &= \textbf{tt} \\ \mathcal{E}(\textbf{false})\sigma &= \textbf{ff} \\ \mathcal{E}(\textbf{not } b)\sigma &= \neg\mathcal{E}(b)\sigma \\ \ldots &= \ldots \end{aligned}$$

pWhile – SOS Semantics I

R0 $\langle \textbf{skip}, \sigma \rangle \Rightarrow_1 \langle \textbf{stop}, \sigma \rangle$

R1 $\langle \textbf{stop}, \sigma \rangle \Rightarrow_1 \langle \textbf{stop}, \sigma \rangle$

R2 $\langle \mathtt{v} := e, \sigma \rangle \Rightarrow_1 \langle \textbf{stop}, \sigma[v \mapsto \mathcal{E}(e)\sigma] \rangle$

R3 $\langle \mathtt{v} \,?\!= r, \sigma \rangle \Rightarrow_{\frac{1}{|r|}} \langle \textbf{stop}, \sigma[v \mapsto r_i \in r] \rangle$

R4$_1$ $$\frac{\langle S_1, \sigma \rangle \Rightarrow_p \langle S_1', \sigma' \rangle}{\langle S_1; S_2, \sigma \rangle \Rightarrow_p \langle S_1'; S_2, \sigma' \rangle}$$

R4$_2$ $$\frac{\langle S_1, \sigma \rangle \Rightarrow_p \langle \textbf{stop}, \sigma' \rangle}{\langle S_1; S_2, \sigma \rangle \Rightarrow_p \langle S_2, \sigma' \rangle}$$

pWhile – SOS Semantics II

R5$_1$	$\langle$**choose** $p_1 : S_1$ **or** $p_2 : S_2, \sigma\rangle \Rightarrow_{p_1} \langle S_1, \sigma\rangle$	
R5$_2$	$\langle$**choose** $p_1 : S_1$ **or** $p_2 : S_2, \sigma\rangle \Rightarrow_{p_2} \langle S_2, \sigma\rangle$	
R6$_1$	$\langle$**if** b **then** S_1 **else** $S_2, \sigma\rangle \Rightarrow_1 \langle S_1, \sigma\rangle$	if $\mathcal{E}(b)\sigma = \mathbf{tt}$
R6$_2$	$\langle$**if** b **then** S_1 **else** $S_2, \sigma\rangle \Rightarrow_1 \langle S_2, \sigma\rangle$	if $\mathcal{E}(b)\sigma = \mathbf{ff}$
R7$_1$	$\langle$**while** b **do** $S, \sigma\rangle \Rightarrow_1 \langle S;$ **while** b **do** $S, \sigma\rangle$	if $\mathcal{E}(b)\sigma = \mathbf{tt}$
R7$_2$	$\langle$**while** b **do** $S, \sigma\rangle \Rightarrow_1 \langle$**stop**$, \sigma\rangle$	if $\mathcal{E}(b)\sigma = \mathbf{ff}$

Factorial

```
var
  m : {0..2};
  n : {0..2};

begin
m := 1;
while (n>1) do
  m := m*n;
  n := n-1;
od;
stop; # looping
end
```

Multi Variable State

The problem we first consider is how to describe distributions over the cartesian product in order to represent the probabilities that two or more variables have certain values.

As we have $\mathcal{D}(S) \subseteq \mathcal{V}(S)$ we investigate $\mathcal{V}(S \times S)$. In order to understand the relation between $\mathcal{V}(S)$ and $\mathcal{V}(S \times S)$ and in general $\mathcal{V}(S^n)$ we need to consider the tensor product.

Essential for the further treatment is the fact (more later) that

$$\mathcal{V}(S \times S) = \mathcal{V}(S) \otimes \mathcal{V}(S)$$

Multi Variable State

The problem we first consider is how to describe distributions over the cartesian product in order to represent the probabilities that two or more variables have certain values.

As we have $\mathcal{D}(S) \subseteq \mathcal{V}(S)$ we investigate $\mathcal{V}(S \times S)$. In order to understand the relation between $\mathcal{V}(S)$ and $\mathcal{V}(S \times S)$ and in general $\mathcal{V}(S^n)$ we need to consider the tensor product.

Essential for the further treatment is the fact (more later) that

$$\mathcal{V}(S \times S) = \mathcal{V}(S) \otimes \mathcal{V}(S)$$

Multi Variable State

The problem we first consider is how to describe distributions over the cartesian product in order to represent the probabilities that two or more variables have certain values.

As we have $\mathcal{D}(S) \subseteq \mathcal{V}(S)$ we investigate $\mathcal{V}(S \times S)$. In order to understand the relation between $\mathcal{V}(S)$ and $\mathcal{V}(S \times S)$ and in general $\mathcal{V}(S^n)$ we need to consider the tensor product.

Essential for the further treatment is the fact (more later) that

$$\mathcal{V}(S \times S) = \mathcal{V}(S) \otimes \mathcal{V}(S)$$

Kronecker Product

Given a $n \times m$ **matrix A** and a $k \times l$ matrix **B**:

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_{11} & \dots & b_{1l} \\ \vdots & \ddots & \vdots \\ b_{k1} & \dots & b_{kl} \end{pmatrix}$$

The **tensor** or **Kronecker product** $\mathbf{A} \otimes \mathbf{B}$ is a $nk \times ml$ matrix:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1m}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & \dots & a_{nm}\mathbf{B} \end{pmatrix}$$

Special cases are **square matrices** ($n = m$ and $k = l$) and **vectors** (row $n = k = 1$, column $m = l = 1$).

Kronecker Product

Given a $n \times m$ **matrix A** and a $k \times l$ matrix **B**:

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_{11} & \dots & b_{1l} \\ \vdots & \ddots & \vdots \\ b_{k1} & \dots & b_{kl} \end{pmatrix}$$

The **tensor** or **Kronecker product** $\mathbf{A} \otimes \mathbf{B}$ is a $nk \times ml$ matrix:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1m}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & \dots & a_{nm}\mathbf{B} \end{pmatrix}$$

Special cases are **square matrices** ($n = m$ and $k = l$) and **vectors** (row $n = k = 1$, column $m = l = 1$).

Kronecker Product

Given a $n \times m$ **matrix A** and a $k \times l$ matrix **B**:

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_{11} & \dots & b_{1l} \\ \vdots & \ddots & \vdots \\ b_{k1} & \dots & b_{kl} \end{pmatrix}$$

The **tensor** or **Kronecker product** $\mathbf{A} \otimes \mathbf{B}$ is a $nk \times ml$ matrix:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1m}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & \dots & a_{nm}\mathbf{B} \end{pmatrix}$$

Special cases are **square matrices** ($n = m$ and $k = l$) and **vectors** (row $n = k = 1$, column $m = l = 1$).

Tensor Product Properties

The tensor product of n linear operators $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ is associative (but in general not commutative) and has e.g. the following properties:

1. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n) \cdot (\mathbf{B}_1 \otimes \ldots \otimes \mathbf{B}_n) = \mathbf{A}_1 \cdot \mathbf{B}_1 \otimes \ldots \otimes \mathbf{A}_n \cdot \mathbf{B}_n$
2. $\mathbf{A}_1 \otimes \ldots \otimes (\alpha \mathbf{A}_i) \otimes \ldots \otimes \mathbf{A}_n = \alpha(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n)$
3. $\mathbf{A}_1 \otimes \ldots \otimes (\mathbf{A}_i + \mathbf{B}_i) \otimes \ldots \otimes \mathbf{A}_n =$
   $= (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n) + (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{B}_i \otimes \ldots \otimes \mathbf{A}_n)$
4. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n)^* = \mathbf{A}_1^* \otimes \ldots \otimes \mathbf{A}_n^*$
5. $\|\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n\| = \|\mathbf{A}_1\| \ldots \|\mathbf{A}_n\|$

Tensor Product Properties

The tensor product of n linear operators $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ is associative (but in general not commutative) and has e.g. the following properties:

1. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n) \cdot (\mathbf{B}_1 \otimes \ldots \otimes \mathbf{B}_n) = \mathbf{A}_1 \cdot \mathbf{B}_1 \otimes \ldots \otimes \mathbf{A}_n \cdot \mathbf{B}_n$
2. $\mathbf{A}_1 \otimes \ldots \otimes (\alpha \mathbf{A}_i) \otimes \ldots \otimes \mathbf{A}_n = \alpha(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n)$
3. $\mathbf{A}_1 \otimes \ldots \otimes (\mathbf{A}_i + \mathbf{B}_i) \otimes \ldots \otimes \mathbf{A}_n =$
   $= (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n) + (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{B}_i \otimes \ldots \otimes \mathbf{A}_n)$
4. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n)^* = \mathbf{A}_1^* \otimes \ldots \otimes \mathbf{A}_n^*$
5. $\|\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n\| = \|\mathbf{A}_1\| \ldots \|\mathbf{A}_n\|$

Tensor Product Properties

The tensor product of n linear operators $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ is associative (but in general not commutative) and has e.g. the following properties:

1. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n) \cdot (\mathbf{B}_1 \otimes \ldots \otimes \mathbf{B}_n) = \mathbf{A}_1 \cdot \mathbf{B}_1 \otimes \ldots \otimes \mathbf{A}_n \cdot \mathbf{B}_n$
2. $\mathbf{A}_1 \otimes \ldots \otimes (\alpha \mathbf{A}_i) \otimes \ldots \otimes \mathbf{A}_n = \alpha(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n)$
3. $\mathbf{A}_1 \otimes \ldots \otimes (\mathbf{A}_i + \mathbf{B}_i) \otimes \ldots \otimes \mathbf{A}_n = $
   $= (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n) + (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{B}_i \otimes \ldots \otimes \mathbf{A}_n)$
4. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n)^* = \mathbf{A}_1^* \otimes \ldots \otimes \mathbf{A}_n^*$
5. $\|\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n\| = \|\mathbf{A}_1\| \ldots \|\mathbf{A}_n\|$

Tensor Product Properties

The tensor product of n linear operators $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ is associative (but in general not commutative) and has e.g. the following properties:

1. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n) \cdot (\mathbf{B}_1 \otimes \ldots \otimes \mathbf{B}_n) = \mathbf{A}_1 \cdot \mathbf{B}_1 \otimes \ldots \otimes \mathbf{A}_n \cdot \mathbf{B}_n$
2. $\mathbf{A}_1 \otimes \ldots \otimes (\alpha \mathbf{A}_i) \otimes \ldots \otimes \mathbf{A}_n = \alpha(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n)$
3. $\mathbf{A}_1 \otimes \ldots \otimes (\mathbf{A}_i + \mathbf{B}_i) \otimes \ldots \otimes \mathbf{A}_n =$
   $= (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n) + (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{B}_i \otimes \ldots \otimes \mathbf{A}_n)$
4. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n)^* = \mathbf{A}_1^* \otimes \ldots \otimes \mathbf{A}_n^*$
5. $\|\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n\| = \|\mathbf{A}_1\| \ldots \|\mathbf{A}_n\|$

Tensor Product Properties

The tensor product of n linear operators $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ is associative (but in general not commutative) and has e.g. the following properties:

1. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n) \cdot (\mathbf{B}_1 \otimes \ldots \otimes \mathbf{B}_n) = \mathbf{A}_1 \cdot \mathbf{B}_1 \otimes \ldots \otimes \mathbf{A}_n \cdot \mathbf{B}_n$
2. $\mathbf{A}_1 \otimes \ldots \otimes (\alpha \mathbf{A}_i) \otimes \ldots \otimes \mathbf{A}_n = \alpha(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n)$
3. $\mathbf{A}_1 \otimes \ldots \otimes (\mathbf{A}_i + \mathbf{B}_i) \otimes \ldots \otimes \mathbf{A}_n =$
   $= (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n) + (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{B}_i \otimes \ldots \otimes \mathbf{A}_n)$
4. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n)^* = \mathbf{A}_1^* \otimes \ldots \otimes \mathbf{A}_n^*$
5. $\|\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n\| = \|\mathbf{A}_1\| \ldots \|\mathbf{A}_n\|$

Tensor Product Properties

The tensor product of n linear operators $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ is associative (but in general not commutative) and has e.g. the following properties:

1. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n) \cdot (\mathbf{B}_1 \otimes \ldots \otimes \mathbf{B}_n) = \mathbf{A}_1 \cdot \mathbf{B}_1 \otimes \ldots \otimes \mathbf{A}_n \cdot \mathbf{B}_n$
2. $\mathbf{A}_1 \otimes \ldots \otimes (\alpha \mathbf{A}_i) \otimes \ldots \otimes \mathbf{A}_n = \alpha(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n)$
3. $\mathbf{A}_1 \otimes \ldots \otimes (\mathbf{A}_i + \mathbf{B}_i) \otimes \ldots \otimes \mathbf{A}_n =$
   $= (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_i \otimes \ldots \otimes \mathbf{A}_n) + (\mathbf{A}_1 \otimes \ldots \otimes \mathbf{B}_i \otimes \ldots \otimes \mathbf{A}_n)$
4. $(\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n)^* = \mathbf{A}_1^* \otimes \ldots \otimes \mathbf{A}_n^*$
5. $\|\mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_n\| = \|\mathbf{A}_1\| \ldots \|\mathbf{A}_n\|$

Tensor Product Base

Every vector space has an algebraic base $\{\mathbf{e}_i\}$

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots$$

This allows to specify vectors via coordinates $\mathbf{x} = (x_1, x_2, \dots)$. Base vectors are in this context simply of the form

$$\mathbf{e}_i = (e_{i1}, e_{i2}, \dots) \quad \text{with } e_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{otherwise} \end{cases}$$

The tensor product space $\mathcal{V} \otimes \mathcal{W}$ can be seen as generated by (formal) tensors of the form $\mathbf{v}_i \otimes \mathbf{w}_j$ with in $\mathbf{v}_i \in \mathcal{V}$ and $\mathbf{w} \in \mathcal{W}$ base vectors.

Tensor Product Base

Every vector space has an algebraic base $\{\mathbf{e}_i\}$

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots$$

This allows to specify vectors via coordinates $\mathbf{x} = (x_1, x_2, \ldots)$.
Base vectors are in this context simply of the form

$$\mathbf{e}_i = (e_{i1}, e_{i2}, \ldots) \quad \text{with } e_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{otherwise} \end{cases}$$

The tensor product space $\mathcal{V} \otimes \mathcal{W}$ can be seen as generated by (formal) tensors of the form $\mathbf{v}_i \otimes \mathbf{w}_j$ with in $\mathbf{v}_i \in \mathcal{V}$ and $\mathbf{w} \in \mathcal{W}$ base vectors.

Tensor Product Base

Every vector space has an algebraic base $\{\mathbf{e}_i\}$

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots$$

This allows to specify vectors via coordinates $\mathbf{x} = (x_1, x_2, \ldots)$. Base vectors are in this context simply of the form

$$\mathbf{e}_i = (e_{i1}, e_{i2}, \ldots) \quad \text{with } e_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{otherwise} \end{cases}$$

The tensor product space $\mathcal{V} \otimes \mathcal{W}$ can be seen as generated by (formal) tensors of the form $\mathbf{v}_i \otimes \mathbf{w}_j$ with in $\mathbf{v}_i \in \mathcal{V}$ and $\mathbf{w} \in \mathcal{W}$ base vectors.

Tensor Product Base

Every vector space has an algebraic base $\{\mathbf{e}_i\}$

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots$$

This allows to specify vectors via coordinates $\mathbf{x} = (x_1, x_2, \ldots)$. Base vectors are in this context simply of the form

$$\mathbf{e}_i = (e_{i1}, e_{i2}, \ldots) \quad \text{with } e_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{otherwise} \end{cases}$$

The tensor product space $\mathcal{V} \otimes \mathcal{W}$ can be seen as generated by (formal) tensors of the form $\mathbf{v}_i \otimes \mathbf{w}_j$ with in $\mathbf{v}_i \in \mathcal{V}$ and $\mathbf{w} \in \mathcal{W}$ base vectors.

Classical and Probabilistic State

We have (always) a finite number v of variables.

Classical state $\sigma \in$ **State** given by:

$$\sigma \in \textbf{State} = (\textbf{Var} \rightarrow \textbf{Value}) = \textbf{Value}^v$$

For each variable we assume also a finite range of values.

Probabilistic state **d** of a single variable is a distribution over possible values of the variable.

$$\mathbf{d} \in \mathcal{V}(\textbf{Value}) = \{ (x_c)_{c \in \textbf{Value}} \mid x_i \in \mathbb{R} \}$$

Classical and Probabilistic State

We have (always) a finite number v of variables.

Classical state $\sigma \in$ **State** given by:

$$\sigma \in \textbf{State} = (\textbf{Var} \rightarrow \textbf{Value}) = \textbf{Value}^v$$

For each variable we assume also a finite range of values.

Probabilistic state **d** of a single variable is a distribution over possible values of the variable.

$$\mathbf{d} \in \mathcal{V}(\textbf{Value}) = \{ (x_c)_{c \in \textbf{Value}} \mid x_i \in \mathbb{R} \}$$

Classical and Probabilistic State

We have (always) a finite number v of variables.

Classical state $\sigma \in$ **State** given by:

$$\sigma \in \textbf{State} = (\textbf{Var} \rightarrow \textbf{Value}) = \textbf{Value}^v$$

For each variable we assume also a finite range of values.

Probabilistic state **d** of a single variable is a distribution over possible values of the variable.

$$\mathbf{d} \in \mathcal{V}(\textbf{Value}) = \{ (x_c)_{c \in \textbf{Value}} \mid x_i \in \mathbb{R} \}$$

Classical and Probabilistic State

We have (always) a finite number v of variables.

Classical state $\sigma \in$ **State** given by:

$$\sigma \in \textbf{State} = (\textbf{Var} \to \textbf{Value}) = \textbf{Value}^v$$

For each variable we assume also a finite range of values.

Probabilistic state **d** of a single variable is a distribution over possible values of the variable.

$$\mathbf{d} \in \mathcal{V}(\textbf{Value}) = \{ (x_c)_{c \in \textbf{Value}} \mid x_i \in \mathbb{R} \}$$

States and Tensor Products

For finite ranges we can represent distributions over cartesian product as an element in the tensor product in $\mathcal{V}(\textbf{Value})^{\otimes v}$.

Probabilistic state **d** of a all variables together

$$\begin{aligned} \mathbf{d} &\in \mathcal{V}(\textbf{Var} \to \textbf{Value}) = \\ &= \mathcal{V}(\textbf{Value}_1 \times \textbf{Value}_2 \times \ldots \times \textbf{Value}_v) = \\ &= \mathcal{V}(\textbf{Value}_1) \otimes \mathcal{V}(\textbf{Value}_2) \otimes \ldots \otimes \mathcal{V}(\textbf{Value}_v) \end{aligned}$$

For infinite value ranges we would need to consider measures. Product measures exist, for example, by Fubini's Theorem.

States and Tensor Products

For finite ranges we can represent distributions over cartesian product as an element in the tensor product in $\mathcal{V}(\textbf{Value})^{\otimes v}$.

Probabilistic state **d** of a all variables together

$$\begin{aligned} \mathbf{d} &\in \mathcal{V}(\textbf{Var} \rightarrow \textbf{Value}) = \\ &= \mathcal{V}(\textbf{Value}_1 \times \textbf{Value}_2 \times \ldots \times \textbf{Value}_v) = \\ &= \mathcal{V}(\textbf{Value}_1) \otimes \mathcal{V}(\textbf{Value}_2) \otimes \ldots \otimes \mathcal{V}(\textbf{Value}_v) \end{aligned}$$

For infinite value ranges we would need to consider measures. Product measures exist, for example, by Fubini's Theorem.

States and Tensor Products

For finite ranges we can represent distributions over cartesian product as an element in the tensor product in $\mathcal{V}(\mathbf{Value})^{\otimes v}$.

Probabilistic state **d** of a all variables together

$$\begin{aligned} \mathbf{d} &\in \mathcal{V}(\mathbf{Var} \to \mathbf{Value}) = \\ &= \mathcal{V}(\mathbf{Value}_1 \times \mathbf{Value}_2 \times \ldots \times \mathbf{Value}_v) = \\ &= \mathcal{V}(\mathbf{Value}_1) \otimes \mathcal{V}(\mathbf{Value}_2) \otimes \ldots \otimes \mathcal{V}(\mathbf{Value}_v) \end{aligned}$$

For infinite value ranges we would need to consider measures. Product measures exist, for example, by Fubini's Theorem.

States and Tensor Products

For finite ranges we can represent distributions over cartesian product as an element in the tensor product in $\mathcal{V}(\textbf{Value})^{\otimes v}$.

Probabilistic state **d** of a all variables together

$$\begin{aligned} \mathbf{d} &\in \mathcal{V}(\textbf{Var} \to \textbf{Value}) = \\ &= \mathcal{V}(\textbf{Value}_1 \times \textbf{Value}_2 \times \ldots \times \textbf{Value}_v) = \\ &= \mathcal{V}(\textbf{Value}_1) \otimes \mathcal{V}(\textbf{Value}_2) \otimes \ldots \otimes \mathcal{V}(\textbf{Value}_v) \end{aligned}$$

For infinite value ranges we would need to consider measures. Product measures exist, for example, by Fubini's Theorem.

States and Tensor Products

For finite ranges we can represent distributions over cartesian product as an element in the tensor product in $\mathcal{V}(\textbf{Value})^{\otimes v}$.

Probabilistic state **d** of a all variables together

$$
\begin{aligned}
\mathbf{d} &\in \mathcal{V}(\textbf{Var} \to \textbf{Value}) = \\
&= \mathcal{V}(\textbf{Value}_1 \times \textbf{Value}_2 \times \ldots \times \textbf{Value}_v) = \\
&= \mathcal{V}(\textbf{Value}_1) \otimes \mathcal{V}(\textbf{Value}_2) \otimes \ldots \otimes \mathcal{V}(\textbf{Value}_v)
\end{aligned}
$$

For infinite value ranges we would need to consider measures. Product measures exist, for example, by Fubini's Theorem.

States and Tensor Products

For finite ranges we can represent distributions over cartesian product as an element in the tensor product in $\mathcal{V}(\textbf{Value})^{\otimes v}$.

Probabilistic state **d** of a all variables together

$$
\begin{aligned}
\mathbf{d} \quad &\in \quad \mathcal{V}(\textbf{Var} \to \textbf{Value}) = \\
&= \quad \mathcal{V}(\textbf{Value}_1 \times \textbf{Value}_2 \times \ldots \times \textbf{Value}_v) = \\
&= \quad \mathcal{V}(\textbf{Value}_1) \otimes \mathcal{V}(\textbf{Value}_2) \otimes \ldots \otimes \mathcal{V}(\textbf{Value}_v)
\end{aligned}
$$

For infinite value ranges we would need to consider measures. Product measures exist, for example, by Fubini's Theorem.

Probabilistic Control Flow

Consider the following (labelled) program:

1: **while** $[z < 100]^1$ **do**
2: **choose**^{2} $\frac{1}{3}$: $[x := 3]^3$ **or** $\frac{2}{3}$: $[x := 1]^4$ **ro**
3: **od**
4: $[\textbf{stop}]^5$

Its probabilistic control flow is given by:

$$flow(P) = \{\langle 1, 1, \underline{2}\rangle, \langle 1, 1, 5\rangle, \langle 2, \frac{1}{3}, 3\rangle, \langle 2, \frac{2}{3}, 4\rangle, \langle 3, 1, 1\rangle, \langle 4, 1, 1\rangle\}.$$

Probabilistic Control Flow

Consider the following (labelled) program:

1: **while** $[z < 100]^1$ **do**
2: **choose**^{2} $\frac{1}{3}$: $[\mathbf{x}:=3]^3$ **or** $\frac{2}{3}$: $[\mathbf{x}:=1]^4$ **ro**
3: **od**
4: $[\mathbf{stop}]^5$

Its probabilistic control flow is given by:

$$\mathit{flow}(P) = \{\langle 1, 1, \underline{2} \rangle, \langle 1, 1, 5 \rangle, \langle 2, \frac{1}{3}, 3 \rangle, \langle 2, \frac{2}{3}, 4 \rangle, \langle 3, 1, 1 \rangle, \langle 4, 1, 1 \rangle\}.$$

Init — First Statement

$$
\begin{aligned}
init([\textbf{skip}]^{\ell}) &= \ell \\
init([\textbf{stop}]^{\ell}) &= \ell \\
init([\mathrm{v} := e]^{\ell}) &= \ell \\
init([\mathrm{v} \,?{=}\, e]^{\ell}) &= \ell \\
init(S_1 ; S_2) &= init(S_1) \\
init(\textbf{choose}^{\ell}\ p_1 : S_1\ \textbf{or}\ p_2 : S_2) &= \ell \\
init(\textbf{if}\ [b]^{\ell}\ \textbf{then}\ S_1\ \textbf{else}\ S_2) &= \ell \\
init(\textbf{while}\ [b]^{\ell}\ \textbf{do}\ S) &= \ell
\end{aligned}
$$

Final — Last Statements

$$
\begin{aligned}
\mathit{final}([\textbf{skip}]^\ell) &= \{\ell\} \\
\mathit{final}([\textbf{stop}]^\ell) &= \{\ell\} \\
\mathit{final}([\mathrm{v} := e]^\ell) &= \{\ell\} \\
\mathit{final}([\mathrm{v} \mathrel{?}= e]^\ell) &= \{\ell\} \\
\mathit{final}(S_1 ; S_2) &= \mathit{final}(S_2) \\
\mathit{final}(\textbf{choose}^\ell\ p_1 : S_1\ \textbf{or}\ p_2 : S_2) &= \mathit{final}(S_1) \cup \mathit{final}(S_2) \\
\mathit{final}(\textbf{if}\ [b]^\ell\ \textbf{then}\ S_1\ \textbf{else}\ S_2) &= \mathit{final}(S_1) \cup \mathit{final}(S_2) \\
\mathit{final}(\textbf{while}\ [b]^\ell\ \textbf{do}\ S) &= \{\ell\}
\end{aligned}
$$

Flow I — Control Transfer

$$
\begin{aligned}
\mathit{flow}([\mathbf{skip}]^{\ell}) &= \emptyset \\
\mathit{flow}([\mathbf{stop}]^{\ell}) &= \{\langle \ell, 1, \ell \rangle\} \\
\mathit{flow}([\mathrm{v} := e]^{\ell}) &= \emptyset \\
\mathit{flow}([\mathrm{v} ?= e]^{\ell}) &= \emptyset \\
\mathit{flow}(S_1 ; S_2) &= \mathit{flow}(S_1) \cup \mathit{flow}(S_2) \cup \\
&\cup \; \{(\ell, 1, \mathit{init}(S_2)) \mid \ell \in \mathit{final}(S_1)\}
\end{aligned}
$$

Flow ||— Control Transfer

$$
\begin{aligned}
\mathit{flow}(\textbf{choose}^{\ell}\ p_1 : S_1\ \textbf{or}\ p_2 : S_2) &= \mathit{flow}(S_1) \cup \mathit{flow}(S_2) \cup \\
&\cup\ \{(\ell, p_1, \mathit{init}(S_1)), (\ell, p_2, \mathit{init}(S_2))\} \\
\mathit{flow}(\textbf{if}\ [b]^{\ell}\ \textbf{then}\ S_1\ \textbf{else}\ S_2) &= \mathit{flow}(S_1) \cup \mathit{flow}(S_2) \cup \\
&\cup\ \{(\ell, 1, \underline{\mathit{init}(S_1)}), (\ell, 1, \mathit{init}(S_2))\} \\
\mathit{flow}(\textbf{while}\ [b]^{\ell}\ \textbf{do}\ S) &= \mathit{flow}(S) \cup \\
&\cup\ \{(\ell, 1, \underline{\mathit{init}(S)})\} \\
&\cup\ \{(\ell', 1, \ell) \mid \ell' \in \mathit{final}(S)\}
\end{aligned}
$$

Collecting Semantics

The collecting semantics of a program P is given by:

$$\mathbf{T}(P) = \sum_{\langle i, p_{ij}, j\rangle \in \mathcal{F}(P)} p_{ij} \cdot \mathbf{T}(\ell_i, \ell_j)$$

i.e. as a linear operator on $\mathcal{V}(\mathbf{Value})^{\otimes v} \otimes \mathcal{V}(\mathbf{Lab})$.

Local effects $\mathbf{T}(\ell_i, \ell_j)$: Data Update $\mathbf{N}$ + Control Step $\mathbf{M}$

$$\mathbf{T}(\ell_i, \ell_j) = \mathbf{N}_i \otimes \mathbf{M}_{ij} = \mathbf{N}_{i1} \otimes \mathbf{N}_{i2} \otimes \ldots \otimes \mathbf{N}_{iv} \otimes \mathbf{M}_{ij}$$

Collecting Semantics

The collecting semantics of a program P is given by:

$$\mathbf{T}(P) = \sum_{\langle i, p_{ij}, j\rangle \in \mathcal{F}(P)} p_{ij} \cdot \mathbf{T}(\ell_i, \ell_j)$$

i.e. as a linear operator on $\mathcal{V}(\mathbf{Value})^{\otimes v} \otimes \mathcal{V}(\mathbf{Lab})$.

Local effects $\mathbf{T}(\ell_i, \ell_j)$: Data Update $\mathbf{N}$ + Control Step $\mathbf{M}$

$$\mathbf{T}(\ell_i, \ell_j) = \mathbf{N}_i \otimes \mathbf{M}_{ij} = \mathbf{N}_{i1} \otimes \mathbf{N}_{i2} \otimes \ldots \otimes \mathbf{N}_{iv} \otimes \mathbf{M}_{ij}$$

Local Transfer Operators

$$
\begin{array}{rcll}
\mathbf{T}(\ell_1,\ell_2) & = & \mathbf{I}\otimes\mathbf{E}(\ell_1,\ell_2) & \text{for } [\textbf{skip}]^{\ell_1} \\
\mathbf{T}(\ell,\ell) & = & \mathbf{I}\otimes\mathbf{E}(\ell,\ell) & \text{for } [\textbf{stop}]^{\ell} \\
\mathbf{T}(\ell_1,\ell_2) & = & \mathbf{U}(v\leftarrow e)\otimes\mathbf{E}(\ell_1,\ell_2) & \text{for } [\mathrm{v} := e]^{\ell_1} \\
\mathbf{T}(\ell_1,\ell_2) & = & \left(\frac{1}{|r|}\sum_{c\in r}\mathbf{U}(v := c)\right)\otimes\mathbf{E}(\ell_1,\ell_2) & \text{for } [\mathrm{v} \mathrel{?{=}} r]^{\ell_1} \\
\mathbf{T}(\ell,\ell_k) & = & \mathbf{I}\otimes\mathbf{E}(\ell,\ell_k) & \text{for } [\textbf{choose}]^{\ell} \\
\mathbf{T}(\ell,\underline{\ell_t}) & = & \mathbf{P}(b=\mathbf{tt})\otimes\mathbf{E}(\ell,\ell_t) & \text{for } [b]^{\ell} \\
\mathbf{T}(\ell,\ell_f) & = & \mathbf{P}(b=\mathbf{ff})\otimes\mathbf{E}(\ell,\ell_f) & \text{for } [b]^{\ell}
\end{array}
$$

Trivial Operators

Matrix Units – Represent a single transition

$$(\mathbf{E}(m,n))_{ij} = \begin{cases} 1 & \text{if } m = i \wedge n = j \\ 0 & \text{otherwise.} \end{cases}$$

Identity – Represents “no change” transition

$$(\mathbf{I})_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Tests Operators and Filters

Select a certain value $c \in$ **Value**:

$$(\mathbf{P}(c))_{ij} = \begin{cases} 1 & \text{if } i = c = j \\ 0 & \text{otherwise.} \end{cases}$$

Select a certain classical state $\sigma \in$ **State**:

$$\mathbf{P}(\sigma) = \bigotimes_{i=1}^{v} \mathbf{P}(\sigma(\mathrm{v}_i))$$

Select states where expression $e = a \mid b$ evaluates to c:

$$\mathbf{P}(e = c) = \sum_{\mathcal{E}(e)\sigma = c} \mathbf{P}(\sigma)$$

Tests Operators and Filters

Select a certain value $c \in$ **Value**:

$$(\mathbf{P}(c))_{ij} = \begin{cases} 1 & \text{if } i = c = j \\ 0 & \text{otherwise.} \end{cases}$$

Select a certain classical state $\sigma \in$ **State**:

$$\mathbf{P}(\sigma) = \bigotimes_{i=1}^{v} \mathbf{P}(\sigma(\mathrm{v}_i))$$

Select states where expression $e = a \mid b$ evaluates to c:

$$\mathbf{P}(e = c) = \sum_{\mathcal{E}(e)\sigma = c} \mathbf{P}(\sigma)$$

Tests Operators and Filters

Select a certain value $c \in$ **Value**:

$$(\mathbf{P}(c))_{ij} = \begin{cases} 1 & \text{if } i = c = j \\ 0 & \text{otherwise.} \end{cases}$$

Select a certain classical state $\sigma \in$ **State**:

$$\mathbf{P}(\sigma) = \bigotimes_{i=1}^{v} \mathbf{P}(\sigma(\mathrm{v}_i))$$

Select states where expression $e = a \mid b$ evaluates to c:

$$\mathbf{P}(e = c) = \sum_{\mathcal{E}(e)\sigma = c} \mathbf{P}(\sigma)$$

Selection via Projections

Filtering out *relevant* configurations, i.e. only those which fulfill a certain condition. Use diagonal matrix **P**:

$$(\mathbf{P})_{ii} = \begin{cases} 1 & \text{if condition holds for } C_i \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{pmatrix}^t \cdot \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ C_2 \\ C_3 \\ 0 \\ C_5 \\ 0 \end{pmatrix}^t$$

Selection via Projections

Filtering out *relevant* configurations, i.e. only those which fulfill a certain condition. Use diagonal matrix **P**:

$$(\mathbf{P})_{ii} = \begin{cases} 1 & \text{if condition holds for } C_i \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{pmatrix}^t \cdot \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ C_2 \\ C_3 \\ 0 \\ C_5 \\ 0 \end{pmatrix}^t$$

Update Operators

For all initial values change to constant $c \in \textbf{Value}$:

$$(\mathbf{U}(c))_{ij} = \begin{cases} 1 & \text{if } j = c \\ 0 & \text{otherwise.} \end{cases}$$

Set value of the kth variable $\mathrm{v}_k \in \textbf{Var}$ to constant $c \in \textbf{Value}$:

$$\mathbf{U}(\mathrm{v}_k \leftarrow c) = \left(\bigotimes_{i=1}^{k-1} \mathbf{I}\right) \otimes \mathbf{U}(c) \otimes \left(\bigotimes_{i=k+1}^{v} \mathbf{I}\right)$$

Set value of variable $\mathrm{v}_k \in \textbf{Var}$ to value given by $e = a \mid b$:

$$\mathbf{U}(\mathrm{v}_k \leftarrow e) = \sum_c \mathbf{P}(e = c)\mathbf{U}(\mathrm{v}_k \leftarrow c)$$

Update Operators

For all initial values change to constant $c \in \mathbf{Value}$:

$$(\mathbf{U}(c))_{ij} = \begin{cases} 1 & \text{if } j = c \\ 0 & \text{otherwise.} \end{cases}$$

Set value of the kth variable $\mathrm{v}_k \in \mathbf{Var}$ to constant $c \in \mathbf{Value}$:

$$\mathbf{U}(\mathrm{v}_k \leftarrow c) = \left(\bigotimes_{i=1}^{k-1} \mathbf{I} \right) \otimes \mathbf{U}(c) \otimes \left(\bigotimes_{i=k+1}^{v} \mathbf{I} \right)$$

Set value of variable $\mathrm{v}_k \in \mathbf{Var}$ to value given by $e = a \mid b$:

$$\mathbf{U}(\mathrm{v}_k \leftarrow e) = \sum_c \mathbf{P}(e = c)\mathbf{U}(\mathrm{v}_k \leftarrow c)$$

Update Operators

For all initial values change to constant $c \in \textbf{Value}$:

$$(\mathbf{U}(c))_{ij} = \begin{cases} 1 & \text{if } j = c \\ 0 & \text{otherwise.} \end{cases}$$

Set value of the kth variable $\mathrm{v}_k \in \textbf{Var}$ to constant $c \in \textbf{Value}$:

$$\mathbf{U}(\mathrm{v}_k \leftarrow c) = \left(\bigotimes_{i=1}^{k-1} \mathbf{I} \right) \otimes \mathbf{U}(c) \otimes \left(\bigotimes_{i=k+1}^{v} \mathbf{I} \right)$$

Set value of variable $\mathrm{v}_k \in \textbf{Var}$ to value given by $e = a \mid b$:

$$\mathbf{U}(\mathrm{v}_k \leftarrow e) = \sum_c \mathbf{P}(e = c)\mathbf{U}(\mathrm{v}_k \leftarrow c)$$

Program Approximation

Problems and Solutions

A general approach towards problems and attempts to solve them could be described as follows:

- If the problem is to difficult
  - formulate a simplified version,
  - try to solve this easy problem.
- Investigate the realtion between the true and the approximate solution.

We know that program analysis is a hard (undecidible) problem.

Problems and Solutions

A general approach towards problems and attempts to solve them could be described as follows:

- If the problem is to difficult
  - formulate a simplified version,
  - try to solve this easy problem.
- Investigate the realtion between the true and the approximate solution.

We know that program analysis is a hard (undecidible) problem.

Problems and Solutions

A general approach towards problems and attempts to solve them could be described as follows:

- If the problem is to difficult
  - formulate a simplified version,
  - try to solve this easy problem.
- Investigate the realtion between the true and the approximate solution.

We know that program analysis is a hard (undecidible) problem.

Problems and Solutions

A general approach towards problems and attempts to solve them could be described as follows:

- If the problem is to difficult
  - formulate a simplified version,
  - try to solve this easy problem.
- Investigate the realtion between the true and the approximate solution.

We know that program analysis is a hard (undecidible) problem.

Problems and Solutions

A general approach towards problems and attempts to solve them could be described as follows:

- If the problem is to difficult
  - formulate a simplified version,
  - try to solve this easy problem.
- Investigate the realtion between the true and the approximate solution.

We know that program analysis is a hard (undecidible) problem.

Problems and Solutions

A general approach towards problems and attempts to solve them could be described as follows:

- If the problem is to difficult
  - formulate a simplified version,
  - try to solve this easy problem.
- Investigate the realtion between the true and the approximate solution.

We know that program analysis is a hard (undecidible) problem.

Possible aims of Static Program Analysis:

- Safe Approximations:
  Correct under all circumstances.
- Good/Close Estimates:
  Fix it (at runtime) if there is a problem.

With modern computer architectures some compile time tasks (type checking, threading, etc.) become runtime features.

A possible application could support Speculative Evaluation.

Probabilistic Program Analysis

Possible aims of Static Program Analysis:

- Safe Approximations:
  Correct under all circumstances.
- Good/Close Estimates:
  Fix it (at runtime) if there is a problem.

With modern computer architectures some compile time tasks (type checking, threading, etc.) become runtime features.

A possible application could support Speculative Evaluation.

Possible aims of Static Program Analysis:

- Safe Approximations:
  Correct under all circumstances.
- Good/Close Estimates:
  Fix it (at runtime) if there is a problem.

With modern computer architectures some compile time tasks (type checking, threading, etc.) become runtime features.

A possible application could support Speculative Evaluation.

Probabilistic Program Analysis

Possible aims of Static Program Analysis:

- Safe Approximations:
  Correct under all circumstances.
- Good/Close Estimates:
  Fix it (at runtime) if there is a problem.

With modern computer architectures some compile time tasks (type checking, threading, etc.) become runtime features.

A possible application could support Speculative Evaluation.

Probabilistic Program Analysis

Possible aims of Static Program Analysis:

- Safe Approximations:
  Correct under all circumstances.
- Good/Close Estimates:
  Fix it (at runtime) if there is a problem.

With modern computer architectures some compile time tasks (type checking, threading, etc.) become runtime features.

A possible application could support Speculative Evaluation.

Semantical Abstraction

Consider a Concrete Domain $\mathcal{C}$ and an Abstract Domain $\mathcal{D}$:

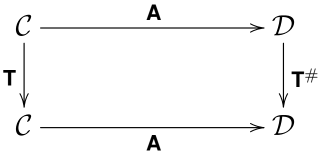
$$
\begin{array}{ccc}
\mathcal{C} & \xrightarrow{\mathbf{A}} & \mathcal{D} \\
\mathbf{T} \downarrow & & \downarrow \mathbf{T}^{\#} \\
\mathcal{C} & \xrightarrow[\mathbf{A}]{} & \mathcal{D}
\end{array}
$$

With an abstraction $\mathbf{A} : \mathbf{C} \rightarrow \mathbf{D}$ and a concretisation $\mathbf{G} : \mathbf{D} \rightarrow \mathbf{C}$:

$$\mathbf{T}^{\#} = \mathbf{GTA}$$

Semantical Abstraction

Consider a Concrete Domain $\mathcal{C}$ and an Abstract Domain $\mathcal{D}$:

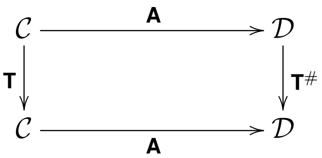


With an abstraction $\mathbf{A} : \mathbf{C} \rightarrow \mathbf{D}$ and a concretisation $\mathbf{G} : \mathbf{D} \rightarrow \mathbf{C}$:

$$\mathbf{T}^{\#} = \mathbf{GTA}$$

Semantical Abstraction

Consider a Concrete Domain $\mathcal{C}$ and an Abstract Domain $\mathcal{D}$:



With an abstraction $\mathbf{A} : \mathbf{C} \to \mathbf{D}$ and a concretisation $\mathbf{G} : \mathbf{D} \to \mathbf{C}$:

$$\mathbf{T}^{\#} = \mathbf{GTA}$$

Semantical Abstraction

Consider a Concrete Domain $\mathcal{C}$ and an Abstract Domain $\mathcal{D}$:

$$\begin{array}{ccc} \mathcal{C} & \xrightarrow{\mathbf{A}} & \mathcal{D} \\ \mathbf{T} \downarrow & & \downarrow \mathbf{T}^{\#} \\ \mathcal{C} & \xrightarrow[\mathbf{A}]{} & \mathcal{D} \end{array}$$

With an abstraction $\mathbf{A} : \mathbf{C} \rightarrow \mathbf{D}$ and a concretisation $\mathbf{G} : \mathbf{D} \rightarrow \mathbf{C}$:

$$\mathbf{T}^{\#} = \mathbf{GTA}$$

Abstract Interpretation: $(\mathbf{A}, \mathbf{G})$ form a Galois Connection.

Semantical Abstraction

Consider a Concrete Domain $\mathcal{C}$ and an Abstract Domain $\mathcal{D}$:

$$
\begin{array}{ccc}
\mathcal{C} & \xrightarrow{\mathbf{A}} & \mathcal{D} \\
\mathbf{T} \downarrow & & \downarrow \mathbf{T}^{\#} \\
\mathcal{C} & \xrightarrow[\mathbf{A}]{} & \mathcal{D}
\end{array}
$$

With an abstraction $\mathbf{A} : \mathbf{C} \rightarrow \mathbf{D}$ and a concretisation $\mathbf{G} : \mathbf{D} \rightarrow \mathbf{C}$:

$$\mathbf{T}^{\#} = \mathbf{GTA}$$

Probabilistic Abst.Int.: $(\mathbf{A}, \mathbf{G})$ Moore-Penrose Pseudo-Inverse.

Galois Connections

Definition

Let $\mathcal{C} = (\mathcal{C}, \leq)$ and $\mathcal{D} = (\mathcal{D}, \sqsubseteq)$ be two partially ordered set. If there are two functions $\alpha : \mathcal{C} \to \mathcal{D}$ and $\gamma : \mathcal{D} \to \mathcal{C}$ such that for all $c \in \mathcal{C}$ and all $d \in \mathcal{D}$:

$$c \leq_{\mathcal{C}} \gamma(d) \text{ iff } \alpha(c) \sqsubseteq d,$$

then $(\mathcal{C}, \alpha, \gamma, \mathcal{D})$ form a Galois connection.

Moore-Penrose Pseudo-Inverse I

Definition

Let $\mathcal{C}$ and $\mathcal{D}$ be two Hilbert spaces and $\mathbf{A} : \mathcal{C} \to \mathcal{D}$ a bounded linear map. A bounded linear map $\mathbf{A}^\dagger = \mathbf{G} : \mathcal{D} \to \mathcal{C}$ is the Moore-Penrose pseudo-inverse of $\mathbf{A}$ iff

(i) $\mathbf{A} \circ \mathbf{G} = \mathbf{P}_A$,

(ii) $\mathbf{G} \circ \mathbf{A} = \mathbf{P}_G$,

where $\mathbf{P}_A$ and $\mathbf{P}_G$ denote orthogonal projections onto the ranges of $\mathbf{A}$ and $\mathbf{G}$.

(Orthogonal) Projections – Idempotents

On finite dimensional vector (Hilbert) spaces we have an inner product $\langle ., . \rangle$. This allows us to define an adjoint via:

$$\langle \mathbf{A}(x), y \rangle = \langle x, \mathbf{A}^*(y) \rangle$$

- An operator $\mathbf{A}$ is self-adjoint if $\mathbf{A} = \mathbf{A}^*$.
- An operator $\mathbf{A}$ is positive, i.e. $\mathbf{A} \sqsupseteq 0$, if there exists an operator $\mathbf{B}$ such that $\mathbf{A} = \mathbf{B}^*\mathbf{B}$.
- An (orthogonal) projection is a self-adjoint $\mathbf{E}$ with $\mathbf{EE} = \mathbf{E}$.

Projections identify (closed) sub-spaces $Y_{\mathbf{E}} = \{\mathbf{E}x \mid x \in \mathcal{V}\}$.

(Orthogonal) Projections – Idempotents

On finite dimensional vector (Hilbert) spaces we have an inner product $\langle ., . \rangle$. This allows us to define an adjoint via:

$$\langle \mathbf{A}(x), y \rangle = \langle x, \mathbf{A}^*(y) \rangle$$

- An operator $\mathbf{A}$ is self-adjoint if $\mathbf{A} = \mathbf{A}^*$.
- An operator $\mathbf{A}$ is positive, i.e. $\mathbf{A} \sqsupseteq 0$, if there exists an operator $\mathbf{B}$ such that $\mathbf{A} = \mathbf{B}^*\mathbf{B}$.
- An (orthogonal) projection is a self-adjoint $\mathbf{E}$ with $\mathbf{EE} = \mathbf{E}$.

Projections identify (closed) sub-spaces $Y_\mathbf{E} = \{\mathbf{E}x \mid x \in \mathcal{V}\}$.

(Orthogonal) Projections – Idempotents

On finite dimensional vector (Hilbert) spaces we have an inner product $\langle ., . \rangle$. This allows us to define an adjoint via:

$$\langle \mathbf{A}(x), y \rangle = \langle x, \mathbf{A}^*(y) \rangle$$

- An operator $\mathbf{A}$ is self-adjoint if $\mathbf{A} = \mathbf{A}^*$.
- An operator $\mathbf{A}$ is positive, i.e. $\mathbf{A} \sqsupseteq 0$, if there exists an operator $\mathbf{B}$ such that $\mathbf{A} = \mathbf{B}^*\mathbf{B}$.
- An (orthogonal) projection is a self-adjoint $\mathbf{E}$ with $\mathbf{EE} = \mathbf{E}$.

Projections identify (closed) sub-spaces $Y_{\mathbf{E}} = \{\mathbf{E}x \mid x \in \mathcal{V}\}$.

(Orthogonal) Projections – Idempotents

On finite dimensional vector (Hilbert) spaces we have an inner product $\langle ., . \rangle$. This allows us to define an adjoint via:

$$\langle \mathbf{A}(x), y \rangle = \langle x, \mathbf{A}^*(y) \rangle$$

- An operator $\mathbf{A}$ is self-adjoint if $\mathbf{A} = \mathbf{A}^*$.
- An operator $\mathbf{A}$ is positive, i.e. $\mathbf{A} \sqsupseteq 0$, if there exists an operator $\mathbf{B}$ such that $\mathbf{A} = \mathbf{B}^*\mathbf{B}$.
- An (orthogonal) projection is a self-adjoint $\mathbf{E}$ with $\mathbf{EE} = \mathbf{E}$.

Projections identify (closed) sub-spaces $Y_\mathbf{E} = \{\mathbf{E}x \mid x \in \mathcal{V}\}$.

(Orthogonal) Projections – Idempotents

On finite dimensional vector (Hilbert) spaces we have an inner product $\langle ., . \rangle$. This allows us to define an adjoint via:

$$\langle \mathbf{A}(x), y \rangle = \langle x, \mathbf{A}^*(y) \rangle$$

- An operator $\mathbf{A}$ is self-adjoint if $\mathbf{A} = \mathbf{A}^*$.
- An operator $\mathbf{A}$ is positive, i.e. $\mathbf{A} \sqsupseteq 0$, if there exists an operator $\mathbf{B}$ such that $\mathbf{A} = \mathbf{B}^*\mathbf{B}$.
- An (orthogonal) projection is a self-adjoint $\mathbf{E}$ with $\mathbf{EE} = \mathbf{E}$.

Projections identify (closed) sub-spaces $Y_{\mathbf{E}} = \{\mathbf{E}x \mid x \in \mathcal{V}\}$.

Moore-Penrose Pseudo-Inverse II

Definition

An operator $\mathbf{A} \in \mathcal{B}(\mathcal{H})$ is Moore-Penrose invertible if there exists an element $\mathbf{G} \in \mathcal{B}(\mathcal{H})$ such that:

(i) $\mathbf{AGA} = \mathbf{A}$,

(ii) $\mathbf{GAG} = \mathbf{G}$,

(iii) $(\mathbf{AG})^* = \mathbf{AG}$,

(iv) $(\mathbf{GA})^* = \mathbf{GA}$.

If it exists $\mathbf{G} = \mathbf{A}^\dagger$ is called Moore-Penrose pseudo-inverse.

Galois Connection II

Definition

Let $\mathcal{C} = (\mathcal{C}, \leq_{\mathcal{C}})$ and $\mathcal{D} = (\mathcal{D}, \leq_{\mathcal{D}})$ be two partially ordered sets with two order-preserving functions $\alpha : \mathcal{C} \mapsto \mathcal{D}$ and $\gamma : \mathcal{D} \mapsto \mathcal{C}$. Then $(\mathcal{C}, \alpha, \gamma, \mathcal{D})$ form a Galois connection iff

(i) $\alpha \circ \gamma$ is reductive i.e. $\forall d \in \mathcal{D}$, $\alpha \circ \gamma(d) \leq_{\mathcal{D}} d$,

(ii) $\gamma \circ \alpha$ is extensive i.e. $\forall c \in \mathcal{C}$, $c \leq_{\mathcal{C}} \gamma \circ \alpha(c)$.

Proposition

Let $(\mathcal{C}, \alpha, \gamma, \mathcal{D})$ be a Galois connection. Then α and γ are quasi-inverse, i.e.

(i) $\alpha \circ \gamma \circ \alpha = \alpha$

(ii) $\gamma \circ \alpha \circ \gamma = \gamma$

Galois Connection II

Definition

Let $\mathcal{C} = (\mathcal{C}, \leq_{\mathcal{C}})$ and $\mathcal{D} = (\mathcal{D}, \leq_{\mathcal{D}})$ be two partially ordered sets with two order-preserving functions $\alpha : \mathcal{C} \mapsto \mathcal{D}$ and $\gamma : \mathcal{D} \mapsto \mathcal{C}$. Then $(\mathcal{C}, \alpha, \gamma, \mathcal{D})$ form a Galois connection iff

(i) $\alpha \circ \gamma$ is reductive i.e. $\forall d \in \mathcal{D}$, $\alpha \circ \gamma(d) \leq_{\mathcal{D}} d$,

(ii) $\gamma \circ \alpha$ is extensive i.e. $\forall c \in \mathcal{C}$, $c \leq_{\mathcal{C}} \gamma \circ \alpha(c)$.

Proposition

Let $(\mathcal{C}, \alpha, \gamma, \mathcal{D})$ be a Galois connection. Then α and γ are quasi-inverse, i.e.

(i) $\alpha \circ \gamma \circ \alpha = \alpha$

(ii) $\gamma \circ \alpha \circ \gamma = \gamma$

Examples of Abstractions

Parity Abstraction operator on $\mathcal{V}(\{1,\ldots,n\})$ (with n even):

$$\mathbf{A}_p = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix} \qquad \mathbf{A}_p^\dagger = \begin{pmatrix} \frac{2}{n} & 0 & \frac{2}{n} & 0 & \ldots & 0 \\ 0 & \frac{2}{n} & 0 & \frac{2}{n} & \ldots & \frac{2}{n} \end{pmatrix}$$

Examples of Abstractions

Parity Abstraction operator on $\mathcal{V}(\{1, \ldots, n\})$ (with n even):

$$\mathbf{A}_p = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix} \qquad \mathbf{A}_p^\dagger = \begin{pmatrix} \frac{2}{n} & 0 & \frac{2}{n} & 0 & \ldots & 0 \\ 0 & \frac{2}{n} & 0 & \frac{2}{n} & \ldots & \frac{2}{n} \end{pmatrix}$$

Examples of Abstractions

Sign Abstraction operator on $\mathcal{V}(\{-n, \ldots, 0, \ldots, n\})$:

$$\mathbf{A}_s = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{A}_s^\dagger = \begin{pmatrix} \frac{1}{n} & \ldots & \frac{1}{n} & 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & \frac{1}{n} & \ldots & \frac{1}{n} \end{pmatrix}$$

Examples of Abstractions

Sign Abstraction operator on $\mathcal{V}(\{-n, \ldots, 0, \ldots, n\})$:

$$\mathbf{A}_s = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{A}_s^\dagger = \begin{pmatrix} \frac{1}{n} & \ldots & \frac{1}{n} & 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & \frac{1}{n} & \ldots & \frac{1}{n} \end{pmatrix}$$

Lifting of an extraction function $\alpha : \mathcal{C} \mapsto \mathcal{D}$

Power Set lifting to an abstraction function $\tilde{\alpha} : \mathcal{P}(\mathcal{C}) \to \mathcal{P}(\mathcal{D})$

$$\tilde{\alpha}(\{c_1, c_2, \ldots\}) = \{\alpha(c_1), \alpha(c_2), \ldots\}$$

Vector Space lifting to an abstraction function $\vec{\alpha} : \mathcal{V}(\mathcal{C}) \to \mathcal{V}(\mathcal{D})$

$$\vec{\alpha}(p_1 \cdot \vec{c}_1 + p_2 \cdot \vec{c}_2 + \ldots) = p_i \cdot \alpha(c_1) + p_2 \cdot \alpha(c_2) \ldots$$

Support Set: **supp** : $\mathcal{V}(\mathcal{C}) \to \mathcal{P}(\mathcal{C})$

$$\textbf{supp}(\vec{x}) = \{c_i \mid \langle c_i, p_i \rangle \in \vec{x} \text{ and } p_i \neq 0\}$$

Uniform Distribution: **vec** : $\mathcal{P}(\mathcal{C}) \to \mathcal{V}(\mathcal{C})$

$$\textbf{vec}(\tilde{x}) = \{\langle c_i, 1/|\tilde{x}| \rangle\}$$

Lifting of an extraction function $\alpha : \mathcal{C} \mapsto \mathcal{D}$

Power Set lifting to an abstraction function $\tilde{\alpha} : \mathcal{P}(\mathcal{C}) \to \mathcal{P}(\mathcal{D})$

$$\tilde{\alpha}(\{c_1, c_2, \ldots\}) = \{\alpha(c_1), \alpha(c_2), \ldots\}$$

Vector Space lifting to an abstraction function $\vec{\alpha} : \mathcal{V}(\mathcal{C}) \to \mathcal{V}(\mathcal{D})$

$$\vec{\alpha}(p_1 \cdot \vec{c}_1 + p_2 \cdot \vec{c}_2 + \ldots) = p_i \cdot \alpha(c_1) + p_2 \cdot \alpha(c_2) \ldots$$

Support Set: **supp** : $\mathcal{V}(\mathcal{C}) \to \mathcal{P}(\mathcal{C})$

$$\textbf{supp}(\vec{x}) = \left\{ c_i \mid \langle c_i, p_i \rangle \in \vec{x} \text{ and } p_i \neq 0 \right\}$$

Uniform Distribution: **vec** : $\mathcal{P}(\mathcal{C}) \to \mathcal{V}(\mathcal{C})$

$$\textbf{vec}(\tilde{x}) = \{\langle c_i, 1/|\tilde{x}| \rangle\}$$

Relation between Abstractions [PPDP00]

Proposition

Let $\vec{\alpha}$ be a probabilistic abstraction function and let $\vec{\gamma}$ be its Moore-Penrose pseudo-inverse.

Then $\vec{\gamma} \circ \vec{\alpha}$ is extensive with respect to the inclusion on the support sets of vectors in $\mathcal{V}(\mathcal{C})$, i.e. $\forall \vec{x} \in \mathcal{V}(\mathcal{C})$,

$$\mathbf{supp}(\vec{x}) \subseteq \mathbf{supp}(\vec{\gamma} \circ \vec{\alpha}(\vec{x})).$$

Least Square Approximation

Given a linear equation

$$x\mathbf{A} = b$$

it has either (i) a (unique) solution $\bar{x}$, or (ii) the residual

$$r_x = b - x\mathbf{A}$$

is non-zero for all x.

The (unique) least-square solution $\bar{x}$, i.e. for which the residual $\|b - \bar{x}\mathbf{A}\|$ is minimal, can be obtained using the Moore-Penrose pseudo-inverse:

$$\bar{x} = b\mathbf{A}^\dagger$$

Least Square Approximation

Given a linear equation

$$x\mathbf{A} = b$$

it has either (i) a (unique) solution $\bar{x}$, or (ii) the residual

$$r_x = b - x\mathbf{A}$$

is non-zero for all x.

The (unique) least-square solution $\bar{x}$, i.e. for which the residual $\|b - \bar{x}\mathbf{A}\|$ is minimal, can be obtained using the Moore-Penrose pseudo-inverse:

$$\bar{x} = b\mathbf{A}^{\dagger}$$

Abstract LOS Semantics

Moore-Penrose Pseudo-Inverse of a Tensor Product is simply

$$(\mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \ldots \otimes \mathbf{A}_n)^\dagger = \mathbf{A}_1^\dagger \otimes \mathbf{A}_2^\dagger \otimes \ldots \otimes \mathbf{A}_n^\dagger$$

Via linearity we can construct $\mathbf{T}^\#$ in the same way as $\mathbf{T}$, i.e

$$\mathbf{T}^\#(P) = \sum_{\langle i, p_{ij}, j\rangle \in \mathcal{F}(P)} p_{ij} \cdot \mathbf{T}^\#(\ell_i, \ell_j)$$

with local abstraction of individual variables:

$$\mathbf{T}^\#(\ell_i, \ell_j) = (\mathbf{A}_1^\dagger \mathbf{N}_{i1} \mathbf{A}_1) \otimes (\mathbf{A}_2^\dagger \mathbf{N}_{i2} \mathbf{A}_2) \otimes \ldots \otimes (\mathbf{A}_v^\dagger \mathbf{N}_{iv} \mathbf{A}_v) \otimes \mathbf{M}_{ij}$$

Abstract LOS Semantics

Moore-Penrose Pseudo-Inverse of a Tensor Product is simply

$$(\mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \ldots \otimes \mathbf{A}_n)^\dagger = \mathbf{A}_1^\dagger \otimes \mathbf{A}_2^\dagger \otimes \ldots \otimes \mathbf{A}_n^\dagger$$

Via linearity we can construct $\mathbf{T}^\#$ in the same way as $\mathbf{T}$, i.e

$$\mathbf{T}^\#(P) = \sum_{\langle i, p_{ij}, j\rangle \in \mathcal{F}(P)} p_{ij} \cdot \mathbf{T}^\#(\ell_i, \ell_j)$$

with local abstraction of individual variables:

$$\mathbf{T}^\#(\ell_i, \ell_j) = (\mathbf{A}_1^\dagger \mathbf{N}_{i1} \mathbf{A}_1) \otimes (\mathbf{A}_2^\dagger \mathbf{N}_{i2} \mathbf{A}_2) \otimes \ldots \otimes (\mathbf{A}_v^\dagger \mathbf{N}_{iv} \mathbf{A}_v) \otimes \mathbf{M}_{ij}$$

Proof Argument

$$\begin{aligned}
\mathbf{T}^{\#} &= \mathbf{A}^{\dagger}\mathbf{T}\mathbf{A} \\
&= \mathbf{A}^{\dagger}\Big(\sum_{i,j}\mathbf{T}(i,j)\Big)\mathbf{A} \\
&= \sum_{i,j}\mathbf{A}^{\dagger}\mathbf{T}(i,j)\mathbf{A} \\
&= \sum_{i,j}\Big(\bigotimes_{k}\mathbf{A}_k\otimes\mathbf{I}\Big)^{\dagger}\mathbf{T}(i,j)\Big(\bigotimes_{k}\mathbf{A}_k\otimes\mathbf{I}\Big) \\
&= \sum_{i,j}\Big(\bigotimes_{k}\mathbf{A}_k\otimes\mathbf{I}\Big)^{\dagger}\Big(\bigotimes_{k}\mathbf{N}_{ik}\otimes\mathbf{M}_{ij}\Big)\Big(\bigotimes_{k}\mathbf{A}_k\otimes\mathbf{I}\Big) \\
&= \sum_{i,j}\Big(\bigotimes_{k}(\mathbf{A}_k^{\dagger}\mathbf{N}_{ik}\mathbf{A}_k)\otimes\mathbf{M}_{ij}\Big)
\end{aligned}$$

Proof Argument

$$
\begin{aligned}
\mathbf{T}^{\#} &= \mathbf{A}^{\dagger}\mathbf{T}\mathbf{A} \\
&= \mathbf{A}^{\dagger}\Big(\sum_{i,j}\mathbf{T}(i,j)\Big)\mathbf{A} \\
&= \sum_{i,j}\mathbf{A}^{\dagger}\mathbf{T}(i,j)\mathbf{A} \\
&= \sum_{i,j}\Big(\bigotimes_k \mathbf{A}_k \otimes \mathbf{I}\Big)^{\dagger}\mathbf{T}(i,j)\Big(\bigotimes_k \mathbf{A}_k \otimes \mathbf{I}\Big) \\
&= \sum_{i,j}\Big(\bigotimes_k \mathbf{A}_k \otimes \mathbf{I}\Big)^{\dagger}\Big(\bigotimes_k \mathbf{N}_{ik} \otimes \mathbf{M}_{ij}\Big)\Big(\bigotimes_k \mathbf{A}_k \otimes \mathbf{I}\Big) \\
&= \sum_{i,j}\Big(\bigotimes_k (\mathbf{A}_k^{\dagger}\mathbf{N}_{ik}\mathbf{A}_k) \otimes \mathbf{M}_{ij}\Big)
\end{aligned}
$$

Proof Argument

$$
\begin{aligned}
\mathbf{T}^{\#} &= \mathbf{A}^{\dagger}\mathbf{T}\mathbf{A} \\
&= \mathbf{A}^{\dagger}(\sum_{i,j} \mathbf{T}(i,j))\mathbf{A} \\
&= \sum_{i,j} \mathbf{A}^{\dagger}\mathbf{T}(i,j)\mathbf{A} \\
&= \sum_{i,j} (\bigotimes_{k} \mathbf{A}_k \otimes \mathbf{I})^{\dagger}\mathbf{T}(i,j)(\bigotimes_{k} \mathbf{A}_k \otimes \mathbf{I}) \\
&= \sum_{i,j} (\bigotimes_{k} \mathbf{A}_k \otimes \mathbf{I})^{\dagger}(\bigotimes_{k} \mathbf{N}_{ik} \otimes \mathbf{M}_{ij})(\bigotimes_{k} \mathbf{A}_k \otimes \mathbf{I}) \\
&= \sum_{i,j} (\bigotimes_{k} (\mathbf{A}_k^{\dagger}\mathbf{N}_{ik}\mathbf{A}_k) \otimes \mathbf{M}_{ij})
\end{aligned}
$$

Proof Argument

$$
\begin{aligned}
\mathbf{T}^{\#} &= \mathbf{A}^{\dagger}\mathbf{T}\mathbf{A} \\
&= \mathbf{A}^{\dagger}(\sum_{i,j} \mathbf{T}(i,j))\mathbf{A} \\
&= \sum_{i,j} \mathbf{A}^{\dagger}\mathbf{T}(i,j)\mathbf{A} \\
&= \sum_{i,j} (\bigotimes_{k} \mathbf{A}_k \otimes \mathbf{I})^{\dagger}\mathbf{T}(i,j)(\bigotimes_{k} \mathbf{A}_k \otimes \mathbf{I}) \\
&= \sum_{i,j} (\bigotimes_{k} \mathbf{A}_k \otimes \mathbf{I})^{\dagger}(\bigotimes_{k} \mathbf{N}_{ik} \otimes \mathbf{M}_{ij})(\bigotimes_{k} \mathbf{A}_k \otimes \mathbf{I}) \\
&= \sum_{i,j} (\bigotimes_{k} (\mathbf{A}_k^{\dagger}\mathbf{N}_{ik}\mathbf{A}_k) \otimes \mathbf{M}_{ij})
\end{aligned}
$$

Proof Argument

$$
\begin{aligned}
\mathbf{T}^{\#} &= \mathbf{A}^{\dagger}\mathbf{T}\mathbf{A} \\
&= \mathbf{A}^{\dagger}\Big(\sum_{i,j}\mathbf{T}(i,j)\Big)\mathbf{A} \\
&= \sum_{i,j}\mathbf{A}^{\dagger}\mathbf{T}(i,j)\mathbf{A} \\
&= \sum_{i,j}\Big(\bigotimes_k \mathbf{A}_k \otimes \mathbf{I}\Big)^{\dagger}\mathbf{T}(i,j)\Big(\bigotimes_k \mathbf{A}_k \otimes \mathbf{I}\Big) \\
&= \sum_{i,j}\Big(\bigotimes_k \mathbf{A}_k \otimes \mathbf{I}\Big)^{\dagger}\Big(\bigotimes_k \mathbf{N}_{ik} \otimes \mathbf{M}_{ij}\Big)\Big(\bigotimes_k \mathbf{A}_k \otimes \mathbf{I}\Big) \\
&= \sum_{i,j}\Big(\bigotimes_k (\mathbf{A}_k^{\dagger}\mathbf{N}_{ik}\mathbf{A}_k) \otimes \mathbf{M}_{ij}\Big)
\end{aligned}
$$

Proof Argument

$$
\begin{aligned}
\mathbf{T}^{\#} &= \mathbf{A}^{\dagger}\mathbf{T}\mathbf{A} \\
&= \mathbf{A}^{\dagger}\Big(\sum_{i,j}\mathbf{T}(i,j)\Big)\mathbf{A} \\
&= \sum_{i,j}\mathbf{A}^{\dagger}\mathbf{T}(i,j)\mathbf{A} \\
&= \sum_{i,j}\Big(\bigotimes_{k}\mathbf{A}_k\otimes\mathbf{I}\Big)^{\dagger}\mathbf{T}(i,j)\Big(\bigotimes_{k}\mathbf{A}_k\otimes\mathbf{I}\Big) \\
&= \sum_{i,j}\Big(\bigotimes_{k}\mathbf{A}_k\otimes\mathbf{I}\Big)^{\dagger}\Big(\bigotimes_{k}\mathbf{N}_{ik}\otimes\mathbf{M}_{ij}\Big)\Big(\bigotimes_{k}\mathbf{A}_k\otimes\mathbf{I}\Big) \\
&= \sum_{i,j}\Big(\bigotimes_{k}(\mathbf{A}_k^{\dagger}\mathbf{N}_{ik}\mathbf{A}_k)\otimes\mathbf{M}_{ij}\Big)
\end{aligned}
$$

Example: Factorial

$[m \leftarrow 1]^1$;
while $[n > 1]^2$ **do**
    $[m \leftarrow m \times n]^3$;
    $[n \leftarrow n - 1]^4$
od
$[\textbf{stop}]^5$

Example: Factorial

1: $[m \leftarrow 1]^1$;
2: **while** $[n > 1]^2$ **do**
3: $\quad [m \leftarrow m \times n]^3$;
4: $\quad [n \leftarrow n - 1]^4$
5: **od**
6: $[\textbf{stop}]^5$

$$
\begin{aligned}
\mathbf{T} &= \mathbf{U}(m \leftarrow 1) \otimes \mathbf{E}(1,2) \\
&+ \mathbf{P}(n > 1) \otimes \mathbf{E}(2,3) \\
&+ \mathbf{P}(n \leq 1) \otimes \mathbf{E}(2,5) \\
&+ \mathbf{U}(m \leftarrow m \times n) \otimes \mathbf{E}(3,4) \\
&+ \mathbf{U}(n \leftarrow n - 1) \otimes \mathbf{E}(4,2) \\
&+ \mathbf{I} \otimes \mathbf{E}(5,5)
\end{aligned}
$$

The abstract versions of the local filters and updates, e.g. $\mathbf{P}^{\#}(n > 1), \mathbf{U}^{\#}(m \leftarrow m \times n), \mathbf{U}^{\#}(n \leftarrow n - 1)$ etc. justify our previous ad hoc analysis.

Example: Factorial

1: $[m \leftarrow 1]^1$;
2: **while** $[n > 1]^2$ **do**
3: $[m \leftarrow m \times n]^3$;
4: $[n \leftarrow n - 1]^4$
5: **od**
6: $[\textbf{stop}]^5$

$$
\begin{aligned}
\mathbf{T}^{\#} &= \mathbf{U}^{\#}(m \leftarrow 1) \otimes \mathbf{E}(1,2) \\
&+ \mathbf{P}^{\#}(n > 1) \otimes \mathbf{E}(2,3) \\
&+ \mathbf{P}^{\#}(n \leq 1) \otimes \mathbf{E}(2,5) \\
&+ \mathbf{U}^{\#}(m \leftarrow m \times n) \otimes \mathbf{E}(3,4) \\
&+ \mathbf{U}^{\#}(n \leftarrow n - 1) \otimes \mathbf{E}(4,2) \\
&+ \mathbf{I}^{\#} \otimes \mathbf{E}(5,5)
\end{aligned}
$$

The abstract versions of the local filters and updates, e.g. $\mathbf{P}^{\#}(n > 1), \mathbf{U}^{\#}(m \leftarrow m \times n), \mathbf{U}^{\#}(n \leftarrow n - 1)$ etc. justify our previous ad hoc analysis.

Example: Factorial

1: $[m \leftarrow 1]^1$;
2: **while** $[n > 1]^2$ **do**
3: $\quad [m \leftarrow m \times n]^3$;
4: $\quad [n \leftarrow n - 1]^4$
5: **od**
6: $[\mathbf{stop}]^5$

$$\begin{aligned} \mathbf{T}^{\#} &= \mathbf{U}^{\#}(m \leftarrow 1) \otimes \mathbf{E}(1,2) \\ &+ \mathbf{P}^{\#}(n > 1) \otimes \mathbf{E}(2,3) \\ &+ \mathbf{P}^{\#}(n \leq 1) \otimes \mathbf{E}(2,5) \\ &+ \mathbf{U}^{\#}(m \leftarrow m \times n) \otimes \mathbf{E}(3,4) \\ &+ \mathbf{U}^{\#}(n \leftarrow n - 1) \otimes \mathbf{E}(4,2) \\ &+ \mathbf{I}^{\#} \otimes \mathbf{E}(5,5) \end{aligned}$$

The abstract versions of the local filters and updates, e.g. $\mathbf{P}^{\#}(n > 1), \mathbf{U}^{\#}(m \leftarrow m \times n), \mathbf{U}^{\#}(n \leftarrow n - 1)$ etc. justify our previous ad hoc analysis.

Abstract Semantics

Abstraction: $\mathbf{A} = \mathbf{A}_p \otimes \mathbf{I}$, i.e. m abstract (parity) but n concrete.

$$
\begin{aligned}
\mathbf{T}^{\#} \quad = \quad & \mathbf{U}^{\#}(m \leftarrow 1) \otimes \mathbf{E}(1,2) \\
+ \quad & \mathbf{P}^{\#}(n > 1) \otimes \mathbf{E}(2,3) \\
+ \quad & \mathbf{P}^{\#}(n \leq 1) \otimes \mathbf{E}(2,5) \\
+ \quad & \mathbf{U}^{\#}(m \leftarrow m \times n) \otimes \mathbf{E}(3,4) \\
+ \quad & \mathbf{U}^{\#}(n \leftarrow n - 1) \otimes \mathbf{E}(4,2) \\
+ \quad & \mathbf{I}^{\#} \otimes \mathbf{E}(5,5)
\end{aligned}
$$

Abstract Semantics

$$\mathbf{U}^{\#}(m \leftarrow i) =$$

$$= \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & & \ldots & 1 \end{pmatrix}$$

Abstract Semantics

$$\mathbf{U}^{\#}(n \leftarrow n-1) =$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

Abstract Semantics

$\mathbf{P}^{\#}(n > 1))$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

$$\mathbf{P}^{\#}(n \leq 1) =$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Abstract Semantics

$$\mathbf{U}^{\#}(m \leftarrow m \times n) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} +$$

$$+ \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \ddots \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \ddots \end{pmatrix}$$

Implementation of concrete and abstract semantics of Factorial using **octave**. Ranges: $n \in \{1, 2, \max\}$ and $m \in \{1, 2, \max!\}$.

n	$\dim(\mathbf{T}(F))$	$\dim(\mathbf{T}^{\#}(F))$
2	45	30
3	140	40
4	625	50
5	3630	60
6	25235	70
7	201640	80
8	1814445	90
9	18144050	100

Using uniform initial distributions $\mathbf{d}_0$ for n and m.

Implementation

Implementation of concrete and abstract semantics of Factorial using **octave**. Ranges: $n \in \{1, 2, \max\}$ and $m \in \{1, 2, \max!\}$.

n	$\dim(\mathbf{T}(F))$	$\dim(\mathbf{T}^{\#}(F))$
2	45	30
3	140	40
4	625	50
5	3630	60
6	25235	70
7	201640	80
8	1814445	90
9	18144050	100

Using uniform initial distributions $\mathbf{d_0}$ for n and m.

Scaleablity

The abstract probabilities for m being **even** or **odd** when we execute the abstract program for various maximal n values are:

n	**even**	**odd**
10	0.81818	0.18182
100	0.98019	0.019802
1000	0.99800	0.0019980
10000	0.99980	0.00019998

The End

Bibliography

- Morris W. Hirsch, Stephen Smale, and Robert L. Devaney.
  Differential Equations, Dynamical Systems and An Introduction to Chaos.
  Elsevier, 2004.
- L.D. Landau and E.M. Lifschitz.
  Mechanik.
  Akademie-Verlag, Berlin, 1981.
- Barry Simon.
  Representation of Finite and Compact Groups, volume 10 of *Graduate Studies in Mathematics*.
  AMS, 1996.