SEMINAR „PERLEN DER THEORETISCHEN INFORMATIK“

QUANTENCOMPUTER II — BJÖRN KARGE

Seminarleitung: Prof. J. Köbler, O. Beyersdorff; Humboldt-Universität Berlin

ZUSAMMENFASSUNG. Ausgehend von den Postulaten der Quantenmechanik sollen der Begriff der Quanten-TURING-Maschine eingeführt und einige in diesem Zusammenhang auftretende Komplexitätsklassen vorgestellt werden.

INHALTSVERZEICHNIS

1. ÜBERBLICK

Einführend werden die Postulate der Quantenmechanik kurz vorgestellt, wobei auf das Messungs-Postulat und den Spezialfall projektiver Messungen genauer eingegangen wird, da diese für die Formulierung des Begriffes der Quanten-TURING-Maschine (QTM) von Bedeutung sind. Danach wird der Begriff der QTM formal eingeführt, die Konstruktion einer universellen QTM vorgestellt und abschließend im Zusammenhang mit QTM auftretende Komplexitätsklassen definiert und ein Überblick über Beziehungen zu ihren „klassischen“ Gegenstücken gegeben.

Das ursprüngliche Vortragsthema („Erzeugung von Zufallszahlen mit Quantenrechnern“) wird dabei in so offensichtlicher Weise erschlagen, daß im folgenden nicht weiter ausdrücklich darauf eingegangen wird.

2. Grundlagen

Diese Ausarbeitung richtet sich an Informatik-Studenten im Hauptstudium. Ein Abriß der wichtigsten hier verwendeten mathematischen Hilfsmittel findet sich in Abschnitt 6.

Überblickshafte Einführungen zum Thema Quantencomputer gibt es, neben der unter `www.informatik.hu-berlin.de/Forschung_Lehre/algorithmenII/Lehre/WS2002-2003/Perlen` verfügbaren Ausarbeitung des Vorgängervortrages, z.B. bei [WC98], `www.qubit.org` und natürlich unter den passenden Stichworten in jeder Suchmaschine. Eine interessante Betrachtung metamathematischer Konsequenzen findet man bei [DEL99].

Die folgenden drei Postulate der Quantenmechanik werden hier nur kurz rekapituliert, ausführlicher werden sie z.B. in [NC00] diskutiert.

Postulat 1. *Zu jedem isolierten Quantensystem gibt es einen* HILBERT*-Raum* Ψ, *genannt* ***Zustandsraum****, dergestalt, daß sich jeder Zustand des Systems durch einen Einheitsvektor* $|\psi\rangle \in \Psi$, *den* ***Zustandsvektor****, vollständig beschreiben läßt. Das Skalarprodukt zweier Vektoren* $|\varphi\rangle, |\psi\rangle \in \Psi$ *ist dabei definiert als* $\langle\varphi|\psi\rangle := \langle\varphi|\,|\psi\rangle$, *wobei* $\langle\varphi| := |\varphi\rangle^*$. [1]

Ein Quantensystem mit einem zweidimensionalen Zustandsraum nennen wir ***qBit***. $\{|0\rangle, |1\rangle\}$ bezeichne immer eine Orthonormalbasis dieses Zustandsraumes. [2]

Postulat 2. *Beschreiben* $|\psi_1\rangle$ *und* $|\psi_2\rangle$ *die Zustände unabhängiger Quantensysteme mit den Zustandsräumen* Ψ_1 *bzw.* Ψ_2, *dann ist* $\Psi_1 \otimes \Psi_2$ *der* ***Zustandsraum des kombinierten Systems****, dessen Zustand durch den Zustandsvektor* $|\psi_1\rangle \otimes |\psi_2\rangle$ *beschrieben wird.*

Wichtig für unsere Betrachtungen ist dabei vor allem die Beobachtung

$$\dim \Psi_1 \otimes \Psi_2 = (\dim \Psi_1) \cdot (\dim \Psi_2).$$

Im Unterschied zu einem klassischen n-Bit-System, welches lediglich 2^n Zustände darstellen kann, ist der Zustandsraum eines n-qBit-Systems also 2^n-dimensional. Wir werden aber sehen, daß wir aufgrund des Messungspostulates nicht in der Lage sein werden, jeden Zustand des Systems zweifelsfrei zu identifizieren, ja sogar bemerken ($\rightarrow$ Beispiel 3), daß wir höchstens 2^n Zustände eines n-qBit-Systems voneinander unterscheiden können!

Postulat 3. *Die* ***zeitliche Entwicklung eines Quantensystems*** *wird durch die* SCHRÖDINGER*-****Gleichung***

$$\mathrm{i}\,\hbar|\dot{\psi}\rangle = H\,|\psi\rangle$$

beschrieben,[3] *wobei der* HAMILTON*-****Operator*** H *eine vom betrachteten System abhängige, hermitesche Abbildung und* $\hbar$ *die* PLANCK*'sche Konstante ist.*

---

[1] Die verschiedenen Notationen werden im Text und zusammengefaßt in Abschnitt 7 erläutert

[2] dabei ist $|0\rangle$ nicht mit dem Nullvektor 0 zu verwechseln

[3] das ist eine verkürzte Schreibweise für die Differentialgleichung $|\psi(t)\rangle' = (1/\,\mathrm{i}\,\hbar)H(t)\,|\psi(t)\rangle$

Dieses Postulat mag den Anschein erwecken, lediglich eine zusätzliche Eigenschaft von Quantensystemen zu sein. In der Praxis hat es jedoch fatale Auswirkungen, denn es ist in der Regel nicht möglich, H für realisierbare Quantensysteme (z.B. mit Temperaturen größer als $0°K$) exakt zu bestimmen. Man kann die Entwicklung eines solchen Systems also nie korrekt vorhergesagen, was den Einsatz von Fehlerkorrekturverfahren bei der Interpretation der Rechenergebnisse von Quantencomputern unerläßlich macht.

Gemeinhin wird jedoch davon ausgegangen, daß man die meisten Algorithmen nachträglich unempfindlicher gegen solche Fehler machen kann, weshalb diese grundsätzliche Schwäche von Quantencomputern bei der Analyse von Algorithmen in der Regel vernachlässigt wird.

3. Messung

Man kann den Vorgang der Beobachtung eines physikalischen Systems allgemein als „qualitative Messung“ auffassen. Bei einer Messung gewinnen wir Informationen über ein System, was nicht anders möglich ist, als das wir mit diesem interagieren. Jede Messung, die an einem System vorgenommen wird, bedeuetet also eine Einwirkung auf das System, indem beispielsweise Energie hinzugefügt oder entzogen wird.

Beispiel 1. Um zu erkennen, ob eine Tür geöffnet oder geschlossen ist, kann man überprüfen, ob man ein Objekt (mit positiver Masse) durch die von der Tür in geschlossenem Zustand begrenzte Fläche bewegen kann. Ist die Tür geschlossen, so wird die Tür beim Aufprall des Objektes eine Veränderung erfahren. Ist die Tür geöffnet, so wird sie durch die Messung nicht in Mitleidenschaft gezogen.

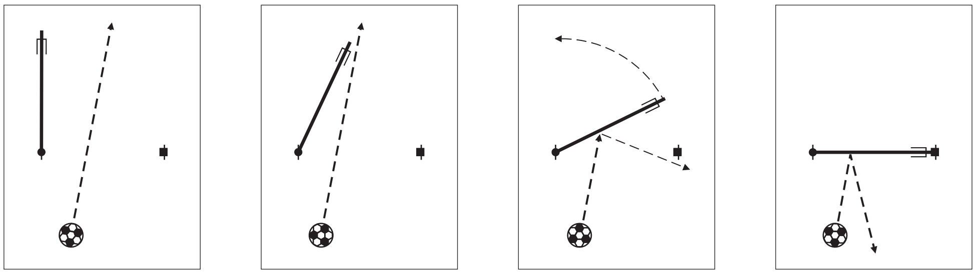


Abb: Qualitative Messung einer Tür mit Hilfe eines Fußballs

Unter Umständen ist es sogar möglich, daß die Tür durch diese Art von Messung geöffnet oder sogar zerstört wird. Insbesondere wird eine (gut geölte) nur halb geöffnete Tür nach einer solchen Messung vollständig geöffnet sein, wenn nur der Einfallswinkel des Objektes geeignet gewählt wurde. ◁

Wir werden daher davon ausgehen, daß jede Messung mathematisch als auf das System einwirkender Operator aufgefaßt werden kann, dessen Wirkung davon abhängt, in welchem Zustand sich das System befindet. Dabei vereinfachen wir den Begriff der Messung zudem auf die Frage „in welchem der Zustände a_1, a_2, ..., a_n befindet sich das System?“.

Um zu erreichen, daß Messungen eindeutig interpretiert werden können, fordern wir, daß zwei Meßergebnisse a und b in einer Messung nicht gleichzeitig auftreten können, wodurch die Menge der durch eine Messung bestimmbaren Zustände des Systems eine Menge von Elementarereignissen bildet, deren Wahrscheinlichkeiten sich zu Eins summieren. Diese Überlegungen, untermauert und ergänzt durch Beobachtungen in physikalischen Experimenten, formalisieren wir durch folgendes

Postulat 4. *Sei A eine Menge möglicher Meßergebnisse (Ereignisse) und $(M_a)_{a\in A}$ eine Familie von linearen **Messoperatoren** auf dem Zustandsraum Ψ eines Quantensystems mit $\sum M_a^* M_a = \mathrm{id}_\Psi$. Wir nennen dann $(M_a)_{a\in A}$ **Messung**.*

Beschreibt $|\psi\rangle \in \Psi$ den Zustand des Systems zum Zeitpunkt der Messung, so gilt:

1. *die Wahrscheinlichkeit $\Pr[a]$ des Eintretens von $a \in A$ bei Durchführung der Messung ist*
$$\Pr[a] = \langle\psi| M_a^* M_a |\psi\rangle\,.$$
2. *Der Zustand des Systems n a c h der Messung, bei der a eingetreten ist, ist*
$$\frac{M_a|\psi\rangle}{\sqrt{\Pr[a]}}.$$

◁

Die Frage, ob Postulat 4 aus den anderen Postulaten abgeleitet werden kann ist unklar, spielt aber für unsere Betrachtungen keine Rolle, solange es ihnen auch nicht widerspricht.

Den Zweck der Forderung $\sum M_a^* M_a = \mathrm{id}_\Psi$ zeigt uns nun folgende Umformung:

$$\sum_{a\in A} \Pr[a] = \sum_{a\in A} \langle\psi| M_a^* M_a |\psi\rangle = \langle\psi| \left(\sum_{a\in A} M_a^* M_a\right) |\psi\rangle = \langle\psi|\psi\rangle \underset{\text{(Postulat1)}}{=} 1.$$

Die Richtung $\sum \Pr[a] = 1 \implies \sum M_a^* M_a = \mathrm{id}_\Psi$ zeigt man ebenso einfach, also ist diese Forderung äquivalent dazu, daß sich die Wahrscheinlichkeiten der Meßergebnisse zu Eins summieren.

Beispiel 2 (Messung eines 1-qBit-Systems). Wir betrachten ein 1-qBit-System mit dem Zustand $|\psi\rangle = a|0\rangle + b|1\rangle$. Ein mögliches Ereignis ist: „das System befindet sich im Zustand $|0\rangle$“ (Ereignis E_1), ein dazu komplementäres: „das System befindet sich im Zustand $|1\rangle$“ (Ereignis E_2). Die Operatoren $A = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$ und $B = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$ bilden eine Messung, denn $A^*A + B^*B = A + B = \mathsf{E}_2$. Zudem erhalten wir:

$$\Pr[E_1] = \langle\psi| A^*A |\psi\rangle = \langle\psi| A |\psi\rangle = (\overline{a}\langle 0| + \overline{b}\langle 1|)A(a|0\rangle + b|1\rangle) = \overline{a}a\langle 0|0\rangle = |a|^2,$$

und analog $\Pr[E_2] = |b|^2$. Falls uns die Messung das Ergebnis E_1 geliefert hat, befindet sich das System danach im Zustand $A|\psi\rangle/|a| = a|0\rangle/|a| = \mathrm{e}^{\mathrm{i}\,\theta}|0\rangle$, wobei $\theta = \arg a \in [\![0, 2\pi[\![$.

Wir können den Wert von θ (die „Phase“ des Zustandes) getrost ignorieren, da offenbar für jeden Operator M gilt: $\langle\psi| M^*M |\psi\rangle = \langle\psi| \mathrm{e}^{-\mathrm{i}\,\theta} M^*M \mathrm{e}^{\mathrm{i}\,\theta} |\psi\rangle$. Aus Sicht des Betrachters befindet sich das System n a c h der Messung also tatsächlich im Zustand $|0\rangle$. ◁

Beispiel 3 (Identifikation von Zuständen). Gegeben einen Zustand $|\psi\rangle \in \{|\psi_1\rangle, ..., |\psi_k\rangle\} \subset \Psi$ mit $|\psi_i\rangle \neq |\psi_j\rangle$ für $i \neq j$, kann man das $j \in \{1, ..., k\}$ mit $|\psi\rangle = |\psi_j\rangle$ bestimmen?

Wir definieren im Falle paarweise orthogonaler $|\psi_i\rangle$ eine Messung wie folgt: [4]

$$M_1 = |\psi_1\rangle\langle\psi_1|, \quad ... \quad, \quad M_k = |\psi_k\rangle\langle\psi_k|, \quad M_0 = \sqrt{\mathrm{id}_\Psi - \sum_{i=1}^{k} M_i}$$

[4] Für eine lineare Abbildung A setzen wir $B = \sqrt{A} :\iff B^2 = BB = A$

M_0 existiert, da $M_1, ..., M_k$ gemeinsam diagonalisierbar sind, z.B. bzgl. einer Basis $(|\psi_1\rangle, ..., |\psi_k\rangle, ...)$, und somit auch ihre Summe und deren Differenz zur identischen Abbildung. In der Basis der Diagonalform ist die Wurzel einer Abbildung aber einfach die komponentenweise Wurzel der Diagonalelemente!

Man rechnet leicht nach, daß die angegebenen Messoperatoren für beliebige orthogonale $|\psi_1\rangle, ..., |\psi_k\rangle$ eine Messung modellieren (M_0 sammelt dabei die „Restwahrscheinlichkeit“). Da, voraussetzungsgemäß, $\Pr[|\psi\rangle = |\psi_j\rangle] = 1$, liefert die Messung mit Sicherheit das Ergebnis $\psi = \psi_j$.

Sind die $|\psi_1\rangle, ..., |\psi_k\rangle$ jedoch nicht orthogonal, so kann man die Zustände nicht eindeutig identifizieren, was man leicht durch Annahme des Gegenteils beweist. Das bedeutet aber, das wir höchstens 2^n Zustände eines n-qBit-Systems voneinander unterscheiden können. ◁

In der Praxis hat sich zudem folgende Beobachtung als nützlich herausgestellt:

Satz 1. *Sind $M = (M_a)_{a\in A}$ und $N = (N_b)_{b\in B}$ Messungen, dann ist*

$$L := M \circ N := (M_a N_b)_{(a,b)\in A\times B}$$

eine Messung und äquivalent dazu, M unmittelbar nach N durchzuführen. •

Das ist nicht wirklich überraschend, denn „anschaulich“ entspricht die Nacheinanderausführung der Messungen ja der Bestimmung der Randverteilung des Ereignisraumes $A \times B$. Von praktischem Nutzen ist dieser Satz vor allem deshalb, weil er uns ermöglicht, Messungen tatsächlich gleichzeitig durchzuführen und somit jeden Einfluß von Postulat 3 auf das System zwischen den beiden Messungen zu eliminieren.

3.1. **Projektive Messung.**

Wir führen nun einen Spezialfall von Messungen ein, der besonders angenehme Eigenschaften hat. Dazu betrachten wir eine hermitesche Abbildung M. M hat (vgl. Abschnitt 6) eine eindeutige Zerlegung

$$M = \sum_{j=1}^{r} \lambda_j P_{\lambda_j},$$

wobei λ_j die Eigenwerte von M (ohne Vielfachheit) sind, und P_{λ_j} jedem Vektor seinen im Eigenraum zu λ_j liegenden Anteil zuordnet (→ Beispiel 4).

Da M hermitesch ist, hat der zugrunde liegende Vektorraum Ψ eine Orthonormalbasis aus Eigenvektoren von M, weshalb die Anteile, in die jeder Vektor bezüglich der Eigenräume von M zerfällt, in der Summe wieder den ursprünglichen Vektor ergeben:

$$v = \sum_{j=1}^{r} P_{\lambda_j} v = \left(\sum_{j=1}^{r} P_{\lambda_j}\right) v = \mathrm{id}_\Psi v.$$

Man mache sich klar, daß aufgrund der Basisunabhängigkeit linearer Abbildungen und der Diagonalisierbarkeit von M auch P_{λ_j} hermitesch sein muß. Zudem gilt offenbar $P_{\lambda_j} P_{\lambda_j} = P_{\lambda_j}$ und somit

$$\mathrm{id}_\Psi = \sum_{j=1}^{r} P_{\lambda_j} = \sum_{j=1}^{r} P_{\lambda_j} P_{\lambda_j} = \sum_{j=1}^{r} P_{\lambda_j}^* P_{\lambda_j},$$

durch M wird also eine Messung beschrieben, das P_{λ_j} entsprechende Ereignis sei im weiteren mit a_j bezeichnet. Eine projektive Messung deren Eigenräume alle die Dimension 1 haben nennen wir, uns an Beispiel 3 erinnernd, ***erschöpfend***.

Eine unmittelbare Konsequenz aus $P_{\lambda_j}^2 = P_{\lambda_j}$ und Satz 1 ist, daß projektive Messungen (abgesehen vom Einfluß von Postulat 3) beliebig oft wiederholbar sind, ohne den Zustand des Systems weiter zu verändern. Einsetzen in Postulat 4 ergibt zudem

$$\Pr[a_j] = \langle\psi| P_{\lambda_j}^* P_{\lambda_j} |\psi\rangle = (P_{\lambda_j} |\psi\rangle)^* (P_{\lambda_j} |\psi\rangle) \underset{(\text{Postulat1})}{=} |P_{\lambda_j} |\psi\rangle|^2,$$

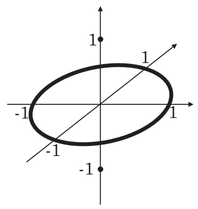
nach Eintreten von a_j erhält man also den Zustand

$$\frac{P_{\lambda_j} |\psi\rangle}{\sqrt{\Pr[\lambda_j]}} = \frac{P_{\lambda_j} |\psi\rangle}{\sqrt{|P_{\lambda_j} |\psi\rangle|^2}} = \frac{P_{\lambda_j} |\psi\rangle}{|P_{\lambda_j} |\psi\rangle|}.$$

Der Zustand nach einer Messung mit Ergebnis a_j ist also ein Eigenvektor von M zum Eigenwert λ_j mit der Länge Eins! Diese Zustände nennt man auch ***Eigenzustände***.

An dieser Stelle sollten wir uns nochmals vor Augen zu halten, in welchem Zusammenhang die Abbildung M zur tatsächlichen Messung steht: lediglich die Eigenräume von M charakterisieren die Projektionsabbildungen P_{λ_j}. Nach der Messung befindet sich das System in einem Zustand, dessen beschreibender Vektor die Länge Eins hat und Teil eines der Eigenräume ist. Die Werte der Eigenwerte von M spielen hier also keine Rolle, es gibt also beliebig viele hermitesche Abbildungen M', die, bis auf die Werte der Eigenwerte, die gleiche Messung charakterisieren.

Tatsächlich wird der Einfluß der Messung also bereits durch die Wahl einer Orthonormalbasis (deren Vektoren dann Eigenvektoren von M sein müssen) und die gewünschte Zuordnung der Basisvektoren zu Eigenräumen bestimmt. Nur, falls der mit einem Meßergebnis assoziierte Eigenraum die Dimension Eins hat, kennen wir nach der Messung den genauen Zustand des Systems (bis auf die „Phase“, was in Beispiel 2 bereits besprochen wurde).



Die Skizze illustriert diesen Sachverhalt für eine Messung eines Systems mit drei Dimensionen, wobei M jedoch nur zwei Eigenwerte hat. Der Kreis und die beiden Punkte kennzeichnen die möglichen Zustände nach der Messung (dabei wurden hier nur die möglichen Realteile dargestellt).

Wir wollen nun den Erwartungswert einer projektiven Messung bestimmen. Sei X dabei die Zufallsvariable, die jedem Elementarereignis a_j den ihm entsprechenden Eigenwert λ_j zuordnet. Dann gilt:

$$\mathbb{E}[X] \underset{\text{def}}{=} \sum_{j=1}^{r} X(a_j)\Pr[a_j] = \sum_{j=1}^{r} \lambda_j \Pr[a_j].$$

Zudem rechnet man nach:

$$\Pr[a_j] = \langle\psi| P_{\lambda_j}^* P_{\lambda_j} |\psi\rangle = \langle\psi| P_{\lambda_j} |\psi\rangle,$$

woraus sich insgesamt wegen der Linearität der P_{λ_j}

$$\mathbb{E}[X] = \sum_{j=1}^{r} \lambda_j \langle\psi| P_{\lambda_j} |\psi\rangle = \langle\psi| \left(\sum_{j=1}^{r} \lambda_j P_{\lambda_j} \right) |\psi\rangle = \langle\psi| M |\psi\rangle$$

ergibt. Falls klar ist, über welches Quantensystem geredet wird, findet sich häufig die Abkürzung $\langle M\rangle := \langle\psi| M |\psi\rangle$. Durch ähnliche Betrachtungen erhält man eine nützliche Interpretation des Skalarproduktes: für eine erschöpfende Messung mit Eigenzustand $|\eta\rangle$, dem das Ereignis a zugeordnet ist, gilt:

$$\Pr[a] = |\langle\psi|\eta\rangle|^2.$$

Beispiel 4 (Zerlegung einer hermiteschen Abbildung). Wir betrachten eine hermitesche Abbildung mit der zugehörigen Matrix

$$M = \begin{pmatrix} 0 & 1 & \mathrm{i} \\ 1 & 0 & -\mathrm{i} \\ -\mathrm{i} & \mathrm{i} & 0 \end{pmatrix}.$$

Diese Abbildung hat die Eigenwerte -2 (einfach) und 1 (doppelt). Der Eigenraum zum Eigenwert 1 wird z.B. durch die Vektoren $\begin{pmatrix} \mathrm{i} \\ 0 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ aufgespannt, der Eigenraum zum Eigenwert -2 z.B. durch $\begin{pmatrix} -\mathrm{i} \\ \mathrm{i} \\ 1 \end{pmatrix}$. Wir erhalten nun die Projektoren P_1 und P_{-2} als

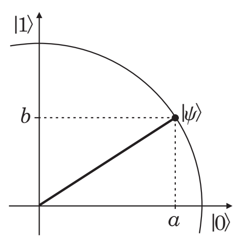
$$P_1 = \begin{pmatrix} -\mathrm{i} & \mathrm{i} & 1 \\ \mathrm{i} & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\mathrm{i} & \mathrm{i} & 1 \\ \mathrm{i} & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} = \frac{1}{3} \begin{pmatrix} 2 & 1 & \mathrm{i} \\ 1 & 2 & -\mathrm{i} \\ -\mathrm{i} & \mathrm{i} & 2 \end{pmatrix}$$

$$P_{-2} = \begin{pmatrix} -\mathrm{i} & \mathrm{i} & 1 \\ \mathrm{i} & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\mathrm{i} & \mathrm{i} & 1 \\ \mathrm{i} & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} = \frac{1}{3} \begin{pmatrix} 1 & -1 & -\mathrm{i} \\ -1 & 1 & \mathrm{i} \\ \mathrm{i} & -\mathrm{i} & 1 \end{pmatrix}.$$

◁

Beispiel 5 (Erzeugung beliebiger Zufallsverteilungen). Wir betrachten ein 1-qBit-System mit dem Zustandsvektor $|\psi\rangle$. Kennen wir $|\psi\rangle$, so fällt es nicht schwer, eine Orthonormalbasis anzugeben, bezüglich der ψ sich als Summe genau zweier Basisvektoren darstellen läßt.

Wählen wir nun eine Abbildung M, die unsere Basisvektoren als Eigenvektoren (zu verschiedenen Eigenwerten) hat, dann erhalten wir wie in Beispiel 2 die Wahrscheinlichkeiten $|a|^2$ bzw. $|b|^2$ für die beiden möglichen Meßergebnisse. Dabei können wir durch Rotatation der Basis den Wert von $|a|^2$ beliebig zwischen 0 und 1 einstellen. Wir können also bei bekanntem $|\psi\rangle$ durch die Wahl unserer Messung eine beliebige Zufallsverteilung des Meßergebnisses erzwingen.



Mit den Resultaten dieses Abschnitts können wir dieses Ergebnis auch etwas leichter gewinnen: Eine hermitesche Abbildung mit den Eigenwerten 0 und 1 hat bei geeigneter Basiswahl die Form $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Falls bezüglich dieser Basis $|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$ ergibt sich als Erwartungswert der Messung $\overline{a}a = |a|^2 = \Pr[E_1]$. Damit könnten Quantensysteme als hervorragende Zufallszahlengeneratoren dienen, eine Eigenschaft, die auch bei Konstruktion der Quanten-TURING-Maschine nicht verloren geht. Allerdings sollte man stets im Hinterkopf behalten, daß diese Erkenntnis auf dem P o s t u l a t Nr. 4 beruht. ◁

4. Die Quanten-Turing-Maschine

Das (im letzten Vortrag vorgestellte) Quantenschaltkreismodell von Richard P. Feynman ist nicht universell in dem Sinne, daß nicht klar wird, wie man damit einen „frei programmierbaren" Quantencomputer herstellen kann. Vielmehr bestand ja seine „Programmierung" darin, ein Quantensystem mit einer Dynamik (d.h., einem Hamilton-Operator) so auszustatten, daß die zeitliche Entwicklung des Systems der gewünschten Berechnung entspricht. Das Programm wurde also als „Hardware" im Hamilton-Operator „fest eingebrannt".

Aus diesem Grund wurde von David Deutsch ein an die Turing-Maschine angelehntes Rechenmodell entwickelt und, darauf aufbauend, eine universelle Quanten-Turing-Maschine (QTM) konstruiert, welche einem solchen „programmierbaren" Quantencomputer entspricht [Deu85]. Zudem unterstellt Deutsch einen fundamentalen Zusammenhang zwischen physikalischen Systemen und dem Berechenbarkeitsbegriff durch eine Neuformulierung der Church'schen These:

These (Deutsch; 1985)**.** *Jedes endliche physikalische System kann durch eine universelle, endliche QTM (beliebig genau) simuliert werden.*

Die universelle QTM von Deutsch hat den Nachteil, daß sie für die Berechnung des Ergebnisses der als Eingabe erhaltenen QTM M (also der Ausführung ihres „Programms") u.U. exponentiell so viel Zeit benötigt wie M, als eigenständige QTM ausgeführt. Bald darauf (1993) wurde jedoch von Ethan Bernstein und Umesh Vazirani nachgewiesen, daß es, wie für klassische Turing-Maschinen, möglich ist, eine universelle QTM zu konstruieren, die die als Eingabe erhaltene QTM stets in polynomieller Zeit simuliert [BV97]. Im selben Jahr hat Andrew C. Yao schließlich die berechnungsmäßige Äquivalenz von Quantenschaltkreisen und Quanten-Turing-Maschinen gezeigt [5] [Yao93].

Ziel dieses Abschnittes ist es, die Konstruktion der universellen QTM von Bernstein und Vazirani vorzustellen. Bevor es aber „richtig losgeht", einigen wir uns noch auf die folgende Definition einer klassischen Turing-Maschine:

Definition (Turing-Maschine)**.** Eine Turing-Maschine wird durch ein 6-Tupel $(\Sigma, \Gamma, Q, q_s, Q_h, \delta)$ mit folgenden Eigenschaften charakterisiert:

Γ	Eine endliche, nichtleere Menge	„Arbeitsalphabet"
Σ	Eine nichtleere Menge, $\Sigma \subsetneqq \Gamma$	„Eingabealphabet"
Q	Eine endliche, nichtleere Menge	„Arbeitszustände"
q_s	$q_s \in Q$	„Startzustand"
Q_h	$Q_h \subseteq Q$	„Haltzustände"
δ	$\delta : Q \times \Gamma \to Q \times \Gamma \times \{R, L\}$	„Zustandsüberführungsfunktion"

Die Turing-Maschine arbeitet auf einem beidseitig unbeschränkten Arbeitsband mit Einträgen aus Γ, von denen nur endlich viele mit Werten aus Σ und der Rest mit einem definierten „Leerzeichen" aus $\Gamma \backslash \Sigma$ initialisiert sind („Eingabe") und startet an der Bandposition 0 im Zustand q_s. Bezeichnet q den aktuellen Arbeitszustand und

[5] Genauer: zu jeder Funktion, die in Polynomialzeit durch eine QTM berechenbar ist, existieren polynomiell große Quantenschaltkreise.

b die Bandinschrift an der Aktuellen Position, so bedeutet $\delta(q, b) = (q', b', X)$, daß b durch b' überschrieben wird, der Schreib-Lese-Kopf der Maschine in Richtung $X \in \{L, R\}$ bewegt wird, und die Maschine danach in den Arbeitszustand q' wechselt. Erreicht die Maschine einen Zustand aus Q_h, so ist die Rechnung beendet, und das Arbeitsband enthält die Ausgabe der Maschine, beginnend an der aktuellen Kopfposition. ◁

4.1. **Das Arbeitsband.**

Das Arbeitsband einer Quanten-TURING-Maschine (QTM) besteht aus einer Folge von qBits, die in der Realität durch Partikel mit Quanteneigenschaften implementiert werden. Diese Partikel müssen bezüglich gewisser meßbarer Attribute die eingangs erwähnten vier Postulate erfüllen.

Beispiel 6. Als Quanten kommen z.B. Elektronen in Frage, wenn man als Meßgröße das Valenzband eines Ions, dem das Elektron gerade angehöre, definiert. Durch Anregung eines Elektrons kann man erreichen, daß es mit gewisser Wahrscheinlichkeit sein Valenzband wechselt. In diesem Falle wäre das Arbeitsband also eine Menge von Ionen und die Zustände würden durch den momentanen Aufenthalt der Ladungsträger kodiert.

Ein weiterer interessanter Ansatz ist die Verwendung von so genannten „Anyonen“ als Quanten. Anyonen sind vorausgesagte Quasiteilchen in zweidimensionalen Medien, die die Eigenschaft hätten, daß sich ihre Bahnen nicht kreuzen würden. Dieser topologischen Eigenschaft entsprechend heißen auf diesem Prinzip aufbauende Quantencomputer „topologische Quantencomputer“. Abgesehen davon, daß der Nachweis von Quasiteilchen mit Anyon-Eigenschaften noch aussteht, könnte dies zu einer deutlichen Verringerung der Fehleranfälligkeit von Quantencomputern führen [FKW00, AG01][Meg02].

Viele weitere Realsierungsmöglichkeiten werden in [NC00] und [Meg02] vorgestellt. ◁

Ein wichtiger Aspekt ist, daß dieses Arbeitsband in unserem Modell für eine gegebene Eingabe zwar endlich aber beliebig lang sein darf. Wir werden uns also eines unendlichdimensionalen Zustandsraumes bedienen müssen. Dieser Raum wird, dadurch charakterisiert sein, daß sich jede gültige Bandinschrift als endliche Linearkombination endlicher klassischer Bandinschriften darstellen läßt. Dieser „Trick“ ermöglicht es, alle endlichen Berechnungen auch in einem endlichdimensionalen Zustandsraum durchführen zu können, bei der Analyse jedoch weniger formalen Beschränkungen unterworfen zu sein.

4.2. **Der Mechanismus.**

Die Arbeitsweise einer QTM während der Rechnung läßt sich in etwa mit einer nichtdeterministischen TURING-Maschine (NTM) vergleichen. Die Zustandsüberführungsfunktion einer NTM ist dabei eine Abbildung

$$\delta : Q \times \Gamma \to \mathfrak{P}(Q \times \Gamma \times \{R, L\}),$$

jede Konfiguration hat also einen ganzen Satz an Folgekonfigurationen, die alle gleichzeitig berechnet werden. Dadurch entsteht ein Baum von Berechnungspfaden von der Wurzel zu den Blättern, die die möglichen Endkonfigurationen repräsentieren.

Am Ende der Berechnung können wir bei einer QTM (im Unterschied zu einer NTM) jedoch höchstens eines dieser Blätter untersuchen. Hier kommt der im nachfolgenden Beispiel demonstrierte Effekt der „Interferenz“ von Berechnungspfaden zum Tragen: eine wichtige Eigenschaft von Quantenalgorithmen besteht nämlich darin, die Anzahl der Blätter während der Berechnung wieder zu reduzieren.

Ebenfalls naheliegend ist der Vergleich von QTM mit probabilistischen Turing-Maschinen (PTM), die aus der Menge möglicher Folgekonfigurationen immer nur eine, entsprechend einer gegebenen Zufallsverteilung $p_{q,\gamma} : Q \times \Gamma \times \{R, L\} \to [\![0,1]\!]$ auswählt:

$$\Pr\left[\delta(q,\gamma) = (q',\gamma',X)\right] = p_{q,\gamma}(q',\gamma',X).$$

Der wesentliche Unterschied zur QTM besteht darin, daß wir mit einer PTM die im folgenden Beispiel angedeuteten „Interferenzeffekte“ zwischen verschiedenen Berechnungspfaden nur mit exponentiellem Aufwand simulieren können.

Beispiel 7. Wir betrachten ein mit $|0\rangle$ initialisiertes qBit und die unitäre Abbildung $C = \sqrt{\frac{1}{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$. Wir erhalten bei Beobachtung von $C\,|0\rangle$ die Ergebnisse $|0\rangle$ und $|1\rangle$ mit gleicher Wahrscheinlichkeit. Wenden wir C ein weiteres mal auf das selbe qBit an, ohne das Ergebnis zu messen, so erhalten wir wegen $C^2 = \mathrm{id}$ jedoch $C^2\,|0\rangle = |0\rangle$. ◁

4.3. Die Formalisierung.

Eine ***Observable*** sei im folgenden immer ein Quantensystem mit einer zugeordneten erschöpfenden Messung.

Definition (Quanten-Turing-Maschine). Eine ***Quanten* - Turing - *Maschine*** wird durch ein Tripel (q, q_h, D) mit folgenden Eigenschaften charakterisiert:

q	Observable von n qBits	„aktueller Zustand“
q_h	Observable eines der qBits von q	„Halt-Indikator“
D	$D : (Q \times \Gamma) \times (Q \times \Gamma \times \{L, R\}) \to \widetilde{\mathbb{C}}$	„Zustandswechselsverteilung“

wobei $Q = \{0, ..., 2^n - 1\}$ die Menge der unterscheidbaren Zustände von q und $\Gamma = \{0, 1\}$ die Menge der unterscheidbaren Zustände eines einzelnen qBits beschreibe. Das Arbeitsband b der QTM besteht aus abzählbar unendlich vielen qBits $(b_i)_{i\in\mathbb{Z}}$, von denen in jeder Berechnung höchstens endlich viele benutzt werden dürfen und die anfänglich alle mit $|0\rangle$ initialisiert sind. Zudem muß es in beschränkter Zeit möglich sein, Informationen zwischen benachbarten Plätze des Bandes auszutauschen. [6] ◁

Die Menge $\widetilde{\mathbb{C}}$ beschreibt die Menge der $z \in \mathbb{C}$, deren Real- und Imaginärteil mit einer Präzision von 2^{-l} deterministisch in in l polynomieller Zeit berechnet werden können, was sich darin begründet, daß die QTM nicht zum Analogrechner verkommen soll und, daß wir sonst gewisse QTM nur mit unvertretbarem Aufwand initialisieren könnten. Beispielsweise könnte man als i-te Nachkommastelle von $D(0, 0, 0, 0, L)$ genau dann eine 1 verlangen, wenn die Turing-Maschine mit der Gödel-Nummer i bei leerem Eingabeband terminiert. Man kann sogar zeigen,

[6] Diese Definition ist etwas restriktiver als bei [BV97], da wir besondere Forderungen an Q und das Alphabet stellen.

daß wir den Wertebereich von D, ohne weitere Einschränkungen zu erfahren, auf $\{0, \pm\frac{3}{5}, \pm\frac{4}{5}, \pm 1\}$ reduzieren können [ADH97][BV97].

Die Elemente von Q werden nachfolgend als ***Arbeitszustände*** bezeichnet, um Verwechslungen mit dem Zustandsvektor des die QTM konstituierenden Quantensystems zu vermeiden. Bei klassischen TURING-Maschinen heißt dagegen der „Zustandsvektor“ ***Konfiguration***, weshalb wir diesen Begriff sinngemäß auch für QTM übernehmen.

Beispiel 8. Zum besseren Verständnis überlegen wir uns, wie die Definition einer klassischen TURING-Maschine in der gleichen Art und Weise ungefähr aussehen müßte:

q	ein n-Bit-Register
q_h	eines der n Bits das, wenn gesetzt, den Haltezustand identifiziert
$\{0,1\}$	Eingabealphabet Σ = Arbeitsalphabet Γ
D	$D : (Q \times \{0,1\}) \times (Q \times \{0,1\} \times \{L, R\}) \to \{0,1\}$

Unsere TURING-Maschine hätte dann 2^n Arbeitszustände, wobei 0 stets der Startzustand wäre. Q wäre auch hier die Menge $\{0, ..., 2^n - 1\}$. Die Hälfte dieser Zustände würde dabei Haltzuständen entsprechen.

Die Abbildung D ist eine implizite Beschreibung der Zustandsübergangsfunktion δ. Der Funktionswert $D(|q\rangle, b, |q'\rangle, v, X) = 1$ bedeutet dann, daß, falls sich die Maschine im Zustand $|q\rangle \in Q$ befindet und auf dem Arbeitsband an der aktuellen Position der Wert $b \in \{0,1\}$ zu finden ist, dieser mit dem Wert $v \in \{0,1\}$ überschrieben wird, und die Maschine danach ihren Schreib-/Lesekopf in Richtung X bewegt und in den Zustand $|q'\rangle$ wechselt.

Es müßte also zusätzlich $D(y, z_1, X_1) = D(y, z_2, X_2) = 1$ für $X_1 \neq X_2$ oder $z_1 \neq z_2$ verboten werden. Zudem müßte gewährleistet sein, daß es zu jedem erreichbaren Zustand einen Folgezustand gibt, also zu jedem während der Rechnung auftretenden Wert von y auch X, z mit $D(y, z, X) = 1$ existieren. Wir erhalten dann mit $D_y(z, X) := D(y, z, X)$ die gewohnte Form der Überführungsfunktion als $\delta(y) = D_y^{-1}(1)$.

Unser Arbeitsband bestünde dann aus abzählbar unendlich vielen Bits $(b_i)_{i \in \mathbb{Z}}$. Wir müßten zudem noch festlegen, wie das Ende der Eingabe markiert wird, da Eingabe- und Arbeitsalphabet identisch sind. Eine Möglichkeit solch einer Festlegung wäre, zu fordern, daß das Eingabeende durch ein gesetztes Bit an gerader Position markiert wird (also nur jedes zweite Bit für die Eingabe benutzbar ist). Alternativ könnte man nur solche Eingaben als gültig definieren, die aus durch einzelne Nullen voneinander getrennten unären Wörtern bestehen. ◁

Sei im weiteren M eine QTM. Um den Zustand von M bei einer Messung vollständig zu beschreiben, benötigen wir eine zusätzliche Observable, p, die uns die aktuelle Bandposition mitteilt. Auch diese wird durch eine beliebig lange Folge von qBits implementiert.

4.4. **Der Zustandsraum.**

Wir werden in diesem Abschnitt einen Zustandsraum $\mathcal{H}$ für M konstruieren. Die wichtigsten Eigenschaften werden am Ende kurz zusammengefaßt, sodaß man die Konstruktion beim ersten Lesen ruhig überspringen kann.

Unser System ist ausschließlich aus qBits zusammengesetzt und Postulat 2 gibt an, wie der zusammengesetzte Zustandsraum unabhängiger Quantensysteme aussieht, womit wir für jede endliche Anzahl k von qBits einen Zustandsraum $\mathcal{H}_k$ angeben können. $\mathcal{H}_k$ ist 2^k-dimensional, also gibt es eine erschöpfende Messung M_k dieses Systems mit 2^k verschiedenen Eigenwerten. Wie im Abschnitt über projektive Messungen bemerkt, können wir diese Eigenwerte ohne Einschränkung als $0, ..., 2^k - 1$ wählen. Wir ordnen jedem Eigenwert λ mit der Binärdarstellung $\lambda = \sum_0^{k-1} \lambda_j 2^j$ wie folgt einen Eigenvektor der Länge 1 zu:

$$|\lambda\rangle := |\lambda_0\rangle \otimes |\lambda_1\rangle \otimes \cdots \otimes |\lambda_{k-1}\rangle =: |\lambda_0 \lambda_1 \cdots \lambda_{k-1}\rangle \,.$$

Beispiel 9. Wir betrachten ein System aus drei qBits. Der Zustandsraum eines jeden qBits wird durch die Vektoren $|0\rangle$ und $|1\rangle$ aufgespannt. Dann bilden (vgl. Abschnitt 6) folgende Vektoren eine Basis des kombinierten Zustandsraumes:

$$\begin{array}{ll} |0\rangle \otimes |0\rangle \otimes |0\rangle =: |000\rangle & |1\rangle \otimes |0\rangle \otimes |0\rangle =: |100\rangle \\ |0\rangle \otimes |0\rangle \otimes |1\rangle =: |001\rangle & |1\rangle \otimes |0\rangle \otimes |1\rangle =: |101\rangle \\ |0\rangle \otimes |1\rangle \otimes |0\rangle =: |010\rangle & |1\rangle \otimes |1\rangle \otimes |0\rangle =: |110\rangle \\ |0\rangle \otimes |1\rangle \otimes |1\rangle =: |011\rangle & |1\rangle \otimes |1\rangle \otimes |1\rangle =: |111\rangle \end{array}$$

Eine erschöpfende Messung des gesamten 3-qBit-Systems wird z.B. durch die Abbildung

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 \end{pmatrix}$$

beschrieben. Wir können die drei qBits z.B. als ein beschränktes Arbeitsband einer QTM der Länge Drei auffassen. Tritt nun bei einer Messung das Ereignis ein, dem wir den Eigenwert $3_{10} = 11_2$ zugeordnet haben, wissen wir, daß sich das System nun im Zustand $|110\rangle$ befindet, den wir mit dem Arbeitsbandinhalt 110 identifizieren können. Befindet sich das System im Zustand $a \cdot |010\rangle + b \cdot |110\rangle + c \cdot |011\rangle$, so hätte dieses Arbeitsband mit Wahrscheinlichkeit $|a|^2$ den Inhalt 010, mit Wahrscheinlichkeit $|b|^2$ den Inhalt 110 und mit Wahrscheinlichkeit $|c|^2$ den Inhalt 011. ◁

Wir führen nun eine nützliche Notation zur Bezeichnung für uns besonders interessanter Eigenvektoren ein. Dazu betrachten wir eine klassische TURING-Maschine T_r, die nur über ein beschränktes Arbeitsband, das nur die Plätze von $-r$ bis r ($r \in \mathbb{N}$) enthält, verfügt. Setzen wir $k = n + 4r + 2$, so können wir jede Konfiguration von T_r mit einer höchstens k-stelligen Binärzahl λ kodieren, in dem wir

$$\lambda = b_r p_{2r} b_{-r} p_{2r-1} \cdots b_2 p_4 b_{-2} p_3 b_1 p_2 b_{-1} p_1 b_0 p_0 q_{n-1} \cdots q_1 q_0$$

setzen und q zur Kodierung des aktuellen Arbeitszustandes, b zur Kodierung des Bandinhaltes und p zur Kodierung der aktuellen Kopfposition verwenden. Da die Verzahnung von b und p dabei unübersichtlich ist vereinbaren wir folgende Schreibweise

$$\begin{aligned} (x; y; z) &:= z_{2r}\ y'_r\ z_{2r-1}\ y'_{-r}\ \cdots\ z_4\ y'_2\ z_3\ y'_{-2}\ z_2\ y'_1\ z_1\ y'_{-1}\ z_0\ y'_0\ x_{n-1} \cdots\ x_1\ x_0 \\ &=: b_r\ p_{2r}\ b_{-r}\ p_{2r-1} \cdots\ b_2\ p_4\ b_{-2}\ p_3\ b_1\ p_2\ b_{-1}\ p_1\ b_0\ p_0\ q_{n-1} \cdots\ q_1\ q_0 \end{aligned}$$

Mit x können wir so den Arbeitszustand identifizieren, mit z eine binäre Bandinschrift (wobei die Kodierung von z einer „Faltung" des Bandes beim Index 0 entspricht) und mit y eine Bandposition, die durch y' geeignet binär kodiert wird,[7] z.B. durch $y'_y = 1$ und $y'_i = 0$ für $i \neq y$.

[7] „geeignet" ist jede mit der später angegebenen Folge von Einbettungen „kompatible" Kodierung

Beispiel. Für $n = 4$, $q = 5$, ein Arbeitsband mit dem Inhalt 11010 auf den Feldern $-2 \cdots 2$ und Kopfposition -2 erhalten wir $z = 01110_2$ und so

$$|5; -2; 14\rangle = |101_2; -2; 1110_2\rangle = |00\ 11\ 10\ 10\ 00\ 0101_2\rangle \mapsto |10100001011100\rangle\,. \qquad \triangleleft$$

Wir können also jeder Konfiguration von T_r eindeutig einen Eigenvektor $|\lambda\rangle \in \mathcal{H}_k$ von M_k zuordnen. Zudem können wir $\mathcal{H}_k$ sehr einfach durch folgende Abbildung in $\mathcal{H}_{k+1}$ einbetten, deren Bild von $|\lambda\rangle$ die selbe Konfiguration wie $|\lambda\rangle$ beschreibt: [8]

$$\mathcal{H}_k \ni |\psi\rangle \mapsto |\psi\rangle \otimes |0\rangle \in \mathcal{H}_{k+1}$$

Den kleinsten HILBERT-Raum, der für jedes k die sukzessive Einbettung

$$\mathcal{H}_k \to \mathcal{H}_{k+1} \to \mathcal{H}_{k+2} \to \cdots$$

enthält, nennen wir $\mathcal{H}$. $\mathcal{H}$ ist offensichtlich unendlichdimensional und die eben eingeführten Zuordnung von Eigenwerten zu Eigenvektoren zählt eine Orthonormalbasis von $\mathcal{H}$ auf, wenn man $|\lambda\rangle \in \mathcal{H}_k$ stets sein durch wiederhohlte Einbettung entstehendes Gegenstück aus $\mathcal{H}$ zuordnet. Dabei wird $\mathcal{H}$ aufgrund seiner Minimalität genau von den endlichen Linearkombinationen dieser Basisvektoren aufgespannt. [9]

Wir haben also als Zustandsraum von M einen unendlichdimensionalen HILBERT-Raum $\mathcal{H}$ konstruiert, in dem jede Konfiguration einer klassischen TURING-Maschine einem Eigenzustand der QTM entspricht. Dabei haben wir für die Konfiguration der klassischen TURING-Maschine die Schreibweise $(q; p; b)$ eingeführt, die bedeutet, daß

- der Inhalt des bei 0 gefalteten und „verzahnten“ Bandes, als Binärzahl aufgefaßt, den Wert b hat
- der Schreib-/Lesekopf sich auf der Bandposition p befindet
- der aktuelle Arbeitszustand der TURING-Maschine die Nummer q hat

und für jede solche Konfiguration den entsprechenden Zustandsvektor $|p; q; b\rangle$ angegeben.

Wenn wir gewährleisten können, daß der Zustand von M sich stets als endliche Linearkombination solcher Zustandsvektoren darstellen läßt, dann können wir, zum einen, sicher sein, daß jede Rechnung nur endlich viele Felder des Arbeitsbandes verwendet, und, zum anderen, jede solche Rechnung auch in einem endlichdimensionalen Unterraum von $\mathcal{H}$ durchführen.

Als gültige Startzustände des Systems lassen wir nun ausschließlich solche zu, die eine Darstellung als

$$|\psi(0)\rangle = \sum_{j=1}^{m} \lambda_j\, |0; 0; t_j\rangle\,, \qquad \lambda \in \mathbb{C}^m, \|\lambda\| = 1, \quad t_j \in \mathbb{N}$$

haben. Dadurch können wir initial nur die Bandinschrift festlegen, diese jedoch als Überlagerung endlich vieler klassischer Bandinschriften.

[8] das Arbeitsband ist zu Beginn überall mit $|0\rangle$ initialisiert

[9] $\mathcal{H}$ entspricht dem HILBERT'sche Folgenraum: Jede endliche Linearkombination von Basisvektoren der Länge 1 hat endliche Länge, ihre Komponenten sind also quadratsummierbar. Der Konvergenzabschluß $\mathcal{H}$ dieser Linearkombinationen ist offenbar isomorph zu ℓ^2.

4.5. Der Hamilton-Operator.

Um zu zeigen, daß M in Einklang mit den Postulaten der Quantenmechanik steht, können wir einen Hamilton-Operator konstruieren, der ihre zeitliche Entwicklung beschreibt. Wir werden feststellen, daß die Auswahl möglicher D dadurch stark eingeschränkt wird.

Die zeitliche Entwicklung von M wird (vgl. Postulat 3) durch den Hamilton-Operator H charakterisiert. Die Schrödinger-Gleichung ist für zeitunabhängige H eine lineare Differentialgleichung mit konstanten Koeffizienten, deren Lösung stets die Form

$$|\psi(t)\rangle = \mathrm{e}^{-\frac{\mathrm{i}\, t H}{\hbar}} \, |\psi(0)\rangle \,,$$

hat. [10] Betrachten wir die Entwicklung von M in diskreten Zeitabschnitten der Länge Δt, so ergibt sich der Zustand des Systems zum Zeitpunkt $k \cdot \Delta t$ als

$$|\psi(k)\rangle := |\psi(k \cdot \Delta t)\rangle = \mathrm{e}^{-\frac{\mathrm{i} \cdot k \cdot \Delta t \cdot H}{\hbar}} \, |\psi(0)\rangle = \left(\mathrm{e}^{-\frac{\mathrm{i} \cdot \Delta t \cdot H}{\hbar}}\right)^k |\psi(0)\rangle =: U^k \, |\psi(0)\rangle \,.$$

Wir betrachten nun für endlichdimensionale Zustandsräume Ψ mit $\dim \Psi = d$ die Gleichung

$$U = \mathrm{e}^{-\frac{\mathrm{i} \cdot \Delta t \cdot H}{\hbar}}$$

H ist hermitesch, hat also bei geeigneter Basiswahl eine Darstellung in Diagonalform $H = (\delta_{ij} h_{ij})$, wir erhalten also mit $c := -\,\mathrm{i} \cdot \Delta t / \hbar$:

$$U = \begin{pmatrix} \mathrm{e}^{c \cdot h_{11}} & & & 0 \\ & \mathrm{e}^{c \cdot h_{22}} & & \\ & & \ddots & \\ 0 & & & \mathrm{e}^{c \cdot h_{dd}} \end{pmatrix}.$$

Da H hermitesch ist, müssen die h_{jj} reell sein, mithin ist $c \cdot h_{jj}$ imaginär ($\hbar$ ist reell). Also folgt wegen $|\,\mathrm{e}^{\mathrm{i}\,\theta}\,| = 1$ für alle $\theta \in \mathbb{R}$ und $z\overline{z} = |z|^2$:

$$UU^* = \mathrm{id}_\Psi .$$

Damit ist U stets unitär. Umgekehrt sind unitäre Abbildungen stets diagonalisierbar, wir können also zu jeder unitären Abbildung in Diagonalgestalt, leicht einen Hamilton-Operator H mit $U = \mathrm{e}^{c \cdot H}$ durch geeignete Wahl der h_{jj} (z.B. als $h_{jj} = (\arg u_{jj})/c$) bestimmen.

Wir werden nun eine unitäre Abbildung U angeben, die als „Interpreter" geeignet gewählter D fungiert. Um mit einer geläufigen Schreibweise in Einklang zu bleiben, beobachten wir zunächst, daß für die darstellende Matrix $A = (a_{ij})$ einer linearen Abbildung bezüglich einer Basis $\mathcal{B} = (|b_1\rangle, ..., |b_d\rangle)$ und Basisvektoren $|b_i\rangle, |b_j\rangle \in \mathcal{B}$ gilt:

$$\langle b_i| \, A \, |b_j\rangle = |b_i\rangle^* \, A \, |b_j\rangle = |b_i\rangle^\mathsf{T} A \, |b_j\rangle = a_{ij}.$$

Dies läßt sich zwar n i c h t auf beliebige unendlichdimensionale Vektorräume übertragen, wir können aber ausnutzen, daß wir zu jedem Zeitpunkt der Arbeit von M nur endlich viele Felder des Arbeitsbandes benutzt haben und wir deshalb statt auf

[10] Das Exponential einer linearen Abbildung A ist als $\mathrm{e}^A := \sum_{j=0}^{\infty} \frac{A^j}{j!}$ definiert, man zeigt leicht, daß dies für A in Diagonalform der Matrix der exponentiierten Diagonalelemente $(\delta_{ij}\, \mathrm{e}^{a_{ij}})$ entspricht.

ganz $\mathcal{H}$ immer die Einschränkung von U auf einen der eingebetteten Vektorräume $\mathcal{H}_k$ betrachten können.

Wir setzen nun

$$\langle q'; p+1; b' | U | q; p; b \rangle := D\left((q, b_p), (q', b'_p), R\right) \prod_{q \neq p} \delta_{b_q, b'_q}$$
$$\langle q'; p-1; b' | U | q; p; b \rangle := D\left((q, b_p), (q', b'_p), L\right) \prod_{q \neq p} \delta_{b_q, b'_q},$$

wobei b_k hier den durch b beschriebenden Bandinhalt an der Bandposition $k \in \mathbb{Z}$ bezeichne. Die betrachteten Einträge von U werden also durch Konfigurationen klassischer TURING-Maschinen indiziert. Jeder von 0 verschiedene Eintrag $u_{ij} = \langle i | U | j \rangle$ von U beschreibt dabei, mit welcher Wahrscheinlichkeit die QTM von der klassischen Konfiguration i in die Konfiguration j wechselt. Dabei wird sichergestellt, daß sich die Kopfposition höchstens um den Wert 1 ändert und zudem nur genau das Feld des Bandes an der Position p modifiziert werden kann (denn sonst würden die Produkte auf der rechten Seite verschwinden). Einzig der Folgezustand q' ist frei wählbar.

Dieses Verhalten überträgt sich auf QTM, die sich in einer endlichen Linearkombination klassischer Konfigurationen befinden. Es ist klar, daß sich die QTM dann auch nach der Anwendung von U in einer endlichen Linearkombination klassischer Konfigurationen befindet.

Eine QTM, für die D unter obenstehender Interpretation eine unitäre Abbildung beschreibt nennen wir, [BV97] folgend, ***wohlgeformt***.

Beispiel 10. Sei T die TURING-Maschine aus Beispiel 8 und $\delta : Q \times \{0,1\} \to Q \times \{0,1\} \times \{L, R\}$ deren Zustandsüberführungsfunktion. Dann wird T durch $M = (q, q_h, D)$ mit

$$D(q, b, q', b', X) = \begin{cases} 1, & \text{falls } \delta(q, b) = (q', b', X) \\ 0, & \text{sonst} \end{cases}$$

simuliert. Insbesondere ist U unitär, wenn δ bijektiv, T also eine reversible TURING-Maschine ist. Jede TURING-Maschine kann man in polynomieller Zeit durch eine reversible TURING-Maschine simulieren [Ben73], mithin auch jede universelle TURING-Maschine. [11] $\triangleleft$

Abschließend wollen wir uns kurz der Frage widmen, wie man herausbekommt, wann die QTM terminiert. Diese Frage ist nicht trivial, da wir q_h nicht beobachten dürfen, wenn es sich in überlagertem Zustand befindet.

Für eine spezielle Klasse von QTM können wir diese Frage jedoch beantworten. Dazu betrachten wir zunächst eine klassische TURING-Maschinen T, deren Ausgabe ausschließlich von der Länge der Eingabe x abhängt. Zu jeder solchen TM T kann man eine klassische reversible TM T' konstruieren, die T effizient simuliert und deren Laufzeit lediglich von der Länge der Eingabe abhängt [BV97].

Gelingt es nun, diese Eigenschaft für alle Berechnungspfade von M nachzuweisen, können wir sicher sein, daß wir q_h stets beobachten dürfen, da alle Berechnungspfade simultan in den Haltezustand wechseln, q_h durch Beobachtung also nie verändert wird. Solche QTM heißen ***anständig*** („well-behaved“). Falls M anständig ist und

[11] Eine universelle TURING-Maschine $\mathcal{T}$ erhält die Kodierung einer TURING-Maschine T als Eingabe, gefolgt von der Eingabe von T und berechnet die Ausgabe von T. Dabei kann man $\mathcal{T}$ so konstruieren, daß $\mathcal{T}$ die Ausgabe von T in höchstens polynomiell so vielen Schritten wie T berechnet.

der Schreib-/Lese-Kopf sich am Ende aller Berechnungspfade wieder in der Ausgangsposition befindet, nennen wir M ***stationär***. Die im nächsten Abschnitt vorgestellte universelle QTM erhält insbesondere diese beiden Eigenschaften, simuliert anständige QTM also anständig und stationäre QTM stationär.

4.6. **Eine effiziente universelle QTM.**

Wie bei einer universellen klassischen TURING-Maschine, ist es Aufgabe einer universellen QTM $\mathcal{Q}$, eine beliebige QTM M zu simulieren. Dabei erhält $\mathcal{Q}$ als Eingabe eine geeignete Kodierung von M, gefolgt von der Eingabe x von M auf dem Eingabeband und berechnet $M(x)$. Von besonderem Interesse ist dabei natürlich das Verhältnis

$$s_{\mathcal{Q}} := s_{\mathcal{Q}}(M,x) := \frac{\mathrm{time}_{\mathcal{Q}}\,(M,x)}{\mathrm{time}_{M}\,(x)}$$

der Anzahl der von $\mathcal{Q}$ benötigten Rechenschritte zur Anzahl der von M benötigten Rechenschritte zur Berechnung von $M(x)$ („slowdown" von $\mathcal{Q}$).

Nun haben wir bereits gesehen, daß es eine QTM gibt, die eine klassische universelle TURING-Maschine $\mathcal{T}$ simuliert ($\rightarrow$ Beispiel 10). Die Idee zur Herstellung einer universellen QTM ist deshalb folgende: wir beschränken uns zunächst darauf, zu überlegen, wie man $\mathcal{T}$ erweitern muß, um einen „Quantenschritt" durchzuführen, d.h., eine Konfiguration zu erreichen, die einer gewünschten Überlagerung mehrerer klassischer Folgezustände entspricht.

Da wir danach davon ausgehen können, daß die QTM ausgehend von jedem dieser Folgezustände „parallel" weiterarbeitet, haben wir so mehrere „virtuelle Instanzen" von $\mathcal{T}$ erzeugt, die ihrerseits wieder beliebige „Quantenschritte" ausführen können, womit wir fertig sind. Insbesondere besteht nicht die Gefahr, daß die Aktionen des Simulators das Gesamtsystem beeinflussen, da dieser bis auf die durch die Eingabe M ausgelösten „Quantenschritte" vollständig deterministisch arbeitet.

Die nun folgende Darstellung ist relativ kompakt und sieht „großzügig" über eine ganze Reihe von „Details" der Konstruktion in [BV97] hinweg (wie z.B. die Frage, ob zwei beliebige QTM zu einer einzigen QTM zusammengesetzt werden können, den Entwurf von Schleifen- und Verzweigungsprimitiven, Fehlerabschätzungen etc.). Nichtsdestotrotz werden wir einige plausible, jedoch nicht triviale Feststellungen aus [BV97] benötigen, die wir die an der jeweiligen Stelle ohne Beweis angegeben werden.

Der Übersichtlichkeit halber statten wir $\mathcal{Q}$ zunächst mit mehreren Arbeitsbändern aus, was jedoch keinen Effizienzverlust bedeutet:

Lemma. *Jede k-Band QTM kann in polynomieller Zeit durch eine* 1*-Band QTM simuliert werden* •

M muß nun auf einem Eingabeband von $\mathcal{Q}$ in qBits kodiert werden. Dazu genügt es, die Anzahl n der qBits, die zur Kodierung des momentanen Arbeitszustandes benötigt werden, den Index h des „Halt-Indikators" unter den n qBits und die Abbildung D anzugeben, um M vollständig zu beschreiben. Da wir bereits jede klassische TURING-Maschine effizient simulieren können, werden wir D als deterministischen Algorithmus $D'(q_1,\gamma_1,q_2,\gamma_2,X,l)$ kodieren, der $D(q_1,\gamma_1,q_2,\gamma_2,X)$ mit der Genauigkeit 2^{-l} in Polynomialzeit (bezüglich l) berechnet. Schließlich benötigen wir noch eine Angabe, mit welcher Genauigkeit ε wir M simulieren sollen, um z.B. l in Abhängigkeit davon geeignet zu bestimmen.

Wie auch bei der Konstruktion einer klassischen universellen TURING-Maschine werden wir M durch Konstruktion der aufeinanderfolgenden Konfigurationen von M simulieren. Diese kodieren wir dabei wie folgt: ein Arbeitsband von $\mathcal{Q}$ wird in Blöcke aus $n+2$ qBits aufgeteilt. Jeder dieser Blöcke ist in der Lage, den aktuellen Arbeitszustand sowie den Bandinhalt von M an der dem Blockindex entsprechenden Position zu kodieren. Dabei wird jedoch nur der Block den aktuellen Arbeitszustand angeben, dessen Index der aktuellen Kopfposition der simulierten QTM entspricht, alle anderen Blöcke enthalten stets den Wert $|0^n\rangle$ auf den ersten n qBits. Das $(n+1)$-te qBit kodiert den Inhalt des Arbeitsbandes von M an dieser Stelle und das letzte qBit dient schließlich als Positionsindikator für den „virtuellen“ Schreib-/Lesekopf von M.

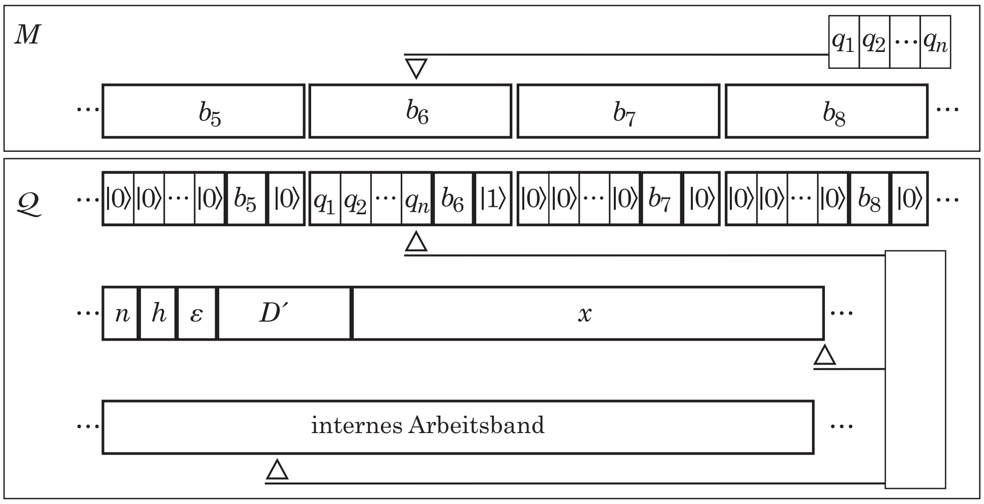


Abb: $\mathcal{Q}$, M in klassischer Konfiguration simulierend

Wir sind nun so weit, uns dem „Quantenschritt“ widmen zu können. Wir betrachten dafür zunächst den Fall, daß M „unidirektional“ ist: eine ***unidirektionale QTM*** ist eine QTM, bei der die Kopfbewegung X sich stets aus dem Folgezustand ableiten läßt, jeder Zustand also immer nur „aus einer der Richtungen“ L oder R erreicht wird.

Da die Kopfbewegung bei einer unidirektionalen QTM keine Rolle mehr für die Invertierbarkeit spielt, beschreibt D eine $\#Q \cdot \#\Gamma = 2^{n+1}$-dimensionale unitäre Abbildung

$$S \simeq (s_{ij}) = (s_{2q_1+\gamma_1, 2q_2+\gamma_2}) = (D(q_1, \gamma_1, q_2, \gamma_2, X(q_2))).$$

Betrachten wir nun eine klassische Konfiguration bei der $|b\rangle$ das Feld an der aktuellen Kopfposition ist, dann berechnet sich der neue Feldinhalt, der neue Arbeitszustand und die neue Kopfposition als

$$|q'b'\rangle = S\,|qb\rangle\,, \qquad |p'\rangle = X\left(|q'\rangle\right),$$

wobei $X(|q'\rangle)$ die komponentenweise Anwendung von X auf $|q'\rangle$ bezeichnet. Dieses Szenario überträgt sich so auch auf $|qb\rangle$ aus nichtklassichen Konfigurationen.

Tatsächlich ist der Fall unidirektionaler QTM aber schon allgemein, denn:

Lemma. *Jede QTM kann in linearer Zeit durch eine unidirektionale QTM simuliert werden.* •

Folgendes Lemma gibt schließlich an, wie schnell sich unitäre Transformationen auf k qBits des Arbeitsbandes ausüben lassen.

Lemma. *Sei $|\psi\rangle := |b_1 b_2 \cdots b_k\rangle$ ein Bandausschnitt und U eine unitäre Abbildung auf dem Zustandsraum von $|\psi\rangle$, die auf einem zweiten Arbeitsband mit ausreichender Präzision kodiert vorliegt. Dann kann U mit Präzision $\varepsilon > 0$ in in 2^k und $1/\varepsilon$ polynomieller Zeit von einer QTM auf $|\psi\rangle$ ausgeübt werden.* •

Diese Formulierung ist sehr lax, für unsere Zwecke soll es aber reichen, zu wissen, daß sich jede unitäre Transformation eines k-qBit-Systems hinreichend genau mit einer QTM in Polynomialzeit, bezogen auf die Präzision ε und die Dimension des Zustandsraumes (also exponentiell in k), durchführen läßt.

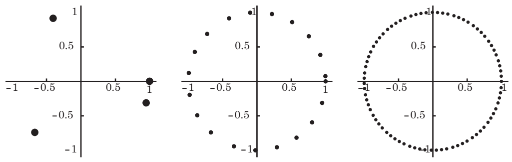


Abb: 3, 19 und 78 sukzessive Rotationen um $2c\pi$

Das dazu in [BV97] angegebene Verfahren approximiert U durch ein Produkt von in 2^k polynomiell vielen Transformationen $U \approx U_1 U_2 \cdots U_n$ in zweidimensionalen Unterräumen von $\mathcal{H}$. Jedes der U_j wird dann durch mehrfache Anwendung einer fixierten Rotationsabbildung beliebig gut approximiert. Dabei wird ausgenutzt, daß durch hinreichend häufige Rotation um $2c\pi$ mit $c \notin \mathbb{Q}$ jeder Winkel θ beliebig gut angenähert werden kann [12] und für geeignet gewähltes c (in etwa einer Intervallschachtelung entsprechend ($c = \sum_{j>0} 2^{-2^j}$)) sogar ausreichend schnell. Dies ist grundsätzlich auch die Idee in [Deu85], nur daß das dort angegeben Verfahren nicht effizient ist.

Insbesondere benötigen wir also für die Ausführung von S auf den ersten $n+1$ qBits eines Blocks stets eine durch 2^{n+1} und $1/\varepsilon$ polynomiell beschränkte Anzahl von Schritten, für fixierte n, ε können wir so die Ausführung von t Schritten von M sicherlich in in t polynomieller Zeit simulieren, es gibt also ein Polynom p mit

$$s_{\mathcal{Q}} = \mathcal{O}(p(\mathrm{time}_M(x))).$$

Die Konstruktion von $\mathcal{Q}$ deutet zudem eine Möglichkeit an, einen universellen Quantencomputer mit relativ einfachen Mitteln herzustellen: wie im Schaltkreismodell von Feynman können wir auch hier alle benötigten Transformationen durch die lokale Anwendung einfacher Gatter durchführen. H wird so als Summe lokaler, lokal konstanter, zeitabhängiger Hamilton-Operatoren realisiert.

5. Komplexität

Was haben wir nun eigentlich gewonnen, außer einer Auffrischung unserer Algebra-Kenntnisse? Um diese Frage zu beantworten, geben wir einen kurzen Überblick über Zusammenhänge zwischen den klassischen Komplexitätsklassen $\mathcal{P}$, $\mathcal{BPP}$ und ihren QTM-Entsprechungen $\mathcal{EQP}$ und $\mathcal{BQP}$.

[12] Genauer: die Gruppe der so erzeugten Rotationen ist dicht in der Menge aller Rotationen, jede Rotation ist also Grenzwert einer Folge von Mehrfach-Rotationen um $2c\pi$. Analog findet man auch solche endlich erzeugten Gruppen, die dicht in der Menge aller unitären Transformationen sind.

$\mathcal{EQP}$ bezeichnet die Menge aller Sprachen, die durch stationäre QTM exakt in Polynomialzeit akzeptiert werden und $\mathcal{BQP}$ die Menge der Sprachen L, für die es eine stationäre, polynomiell laufzeitbeschränkte QTM gibt, die jedes $w \in L$ mit Wahrscheinlichkeit $\frac{2}{3}$ akzeptiert und jedes $w \notin L$ mit Wahrscheinlichkeit $\frac{2}{3}$ verwirft.

Korollar. $\mathcal{P} \subseteq \mathcal{EQP}$ •

Satz (Bernstein,Vazirani; 1993)**.** $\mathcal{P} \subseteq \mathcal{BPP} \subseteq \mathcal{BQP} \subseteq \mathcal{PSPACE}$.

Beweis. [BV97] □

Zudem wurde von Lov K. Grover gezeigt, daß sich jedes $\mathcal{NP}$-vollständige Problem in $\mathcal{O}(\sqrt{2^n})$ mit Hilfe einer QTM lösen läßt [Gro96][BBBV97].

Der „Klassiker“ unter den Quantenalgorihmen ist schließlich das von Peter W. Shor 1994 entwickelte Verfahren zur Faktorisiserung ganzer Zahlen in Polynomialzeit [Sho94]. Bis heute ist kein deterministisches Verfahren bekannt, das diese Aufgabe in Polynomialzeit löst. Dieser und andere Quantenalgorihmen werden detailliert in [NC00] vorgestellt.

6. Mathematische Grundlagen

Diese Ausführungen sind vielleicht hilfreich, um einige Leser nicht durch ein einfaches „$\otimes$“ zu verprellen; übersichtliche, breitere Darstellungen findet man auch in [Fis97], [Kra02] und [Kön97, Kön99].

Sei also im folgenden $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ fixiert und seien U, V, W $\mathbb{K}$-Vektorräume.

6.1. Unitäre Vektorräume und unitäre Abbildungen.

Definition (Eigenwert)**.** Sei f eine lineare Abbildung. $\lambda \in \mathbb{K}$ heißt ***Eigenwert*** von f, falls es ein $v \in V$ mit $v \neq 0$ und $f(v) = \lambda v$ gibt. v heißt dann ***Eigenvektor*** von f zum Eigenwert λ. ◁

Korollar. *Sei f eine lineare Abbildung mit Eigenwert λ. Dann bildet die Menge aller Eigenvektoren zum Eigenwert λ zusammen mit 0 einen Untervektorraum E_λ von V, den* ***Eigenraum*** *von λ.*

Beweis. E_λ ist der Kern der linearen Abbildung $g : v \mapsto f(v) - \lambda v$ und damit ein Untervektorraum von V. □

Definition (Vielfachheit eines Eigenwertes)**.** Sei f eine lineare Abbildung, und $\lambda \in \mathbb{K}$ Eigenwert von f. Die Dimension β_λ des Eigenraumes E_λ von λ heißt ***geometrische Vielfachheit*** von λ. ◁

Definition (diagonalisierbare Abbildung)**.** Eine lineare Abbildung f heißt ***diagonalisierbar***, falls es eine Basis gibt, bezüglich der f eine darstellende Matrix in Diagonalgestalt hat. ◁

Bemerkung. Die Diagonalelemente einer solchen Matrix könne offensichtlich nur Eigenwerte von f sein und jeder Eigenwert taucht mit seiner geometrischen Vielfachheit in der Diagonalen auf. Falls f r paarweise verschiedene Eigenwerte $\lambda_1, ..., \lambda_r$ hat, gilt zudem $V = E_{\lambda_1} \times \cdots \times E_{\lambda_r}$.

Definition (hermitesche M a t r i x). Für eine Matrix $A = (a_{ij})$ setzen wir $A^* := \overline{A^\intercal} = (\overline{a_{ji}})$. Die M a t r i x A heißt ***hermitesch***, falls $A = A^*$. Falls $\mathbb{K} = \mathbb{R}$ nennt man A auch ***symmetrisch***. ◁

Satz. *Alle Eigenwerte einer hermiteschen Matrix sind reell.* •

Korollar. *Für jede hermiteschen Matrix mit den Eigenwerten $\lambda_1, ..., \lambda_r$ gibt es eine Basis bezüglich der A die folgende Form hat:* [13]

$$\begin{pmatrix} \boxed{\lambda_1 \mathsf{E}_{\beta_1}} & & & 0 \\ & \boxed{\lambda_2 \mathsf{E}_{\beta_2}} & & \\ & & \ddots & \\ 0 & & & \boxed{\lambda_r \mathsf{E}_{\beta_r}} \end{pmatrix}.$$

•

Definition (semilineare Abbildung). Eine Abbildung $f : V \to W$ heißt ***semilinear*** (halblinear), falls $f(av + w) = \overline{a} f(v) + f(w)$ für alle $a \in \mathbb{K}$ und $v, w \in V$. [14] ◁

Definition (sesquilineare Abbildung). Eine Abbildung $f : V \times V \to K$ heißt ***sesquilinear*** ($1\frac{1}{2}$-fach linear), falls f im ersten Argument linear und im zweiten Argument semilinear ist, also für alle $v_1, v_2 \in V$, $w_1, w_2 \in W$, $a, b \in \mathbb{K}$ gilt:

$$f(a \cdot v_1 + v_2, b \cdot w_1 + w_2) = a\overline{b} \cdot f(v_1, w_1) + a \cdot f(v_1, w_2) + \overline{b} \cdot f(v_2, w_1) + f(v_2, w_2).$$ ◁

Definition (hermitesche Abbildung). Eine s e s q u i l i n e a r e A b b i l d u n g $f : V \times V \to \mathbb{K}$ heißt ***hermitesch***, falls $f(v, w) = \overline{f(w, v)}$. ◁

Satz. *Sei A hermitesch. Dann ist die Abbildung $f : V \times V \to \mathbb{K}$, $f : (v, w) \mapsto v^\intercal A \overline{w}$ hermitesch.*

Beweis. $\overline{f(v,w)} = \overline{v^\intercal A \overline{w}} = \overline{(v^\intercal A \overline{w})^\intercal} = \overline{\overline{w^\intercal} A^\intercal v} = w^\intercal A^* \overline{v} = w^\intercal A \overline{v} = f(w, v)$; $f(av_1 + v_2, bw_1 + w_2) = (av_1 + v_2)^\intercal A \overline{(bw_1 + w_2)} = (av_1 + v_2)^\intercal A(\overline{b}\overline{w_1} + \overline{w_2}) = a\overline{b} v_1{}^\intercal A \overline{w_1} + a v_1{}^\intercal A \overline{w_2} + \overline{b} v_2{}^\intercal A \overline{w_1} + v_2{}^\intercal A \overline{w_2} = a\overline{b} f(v_1, w_1) + a f(v_1, w_2) + \overline{b} f(v_2, w_1) + f(v_2, w_2)$ □

Definition (Skalarprodukt). Eine hermitesche Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ heißt ***Skalarprodukt***, falls $\langle v, v \rangle > 0$ für alle $v \in V$ mit $v \neq 0$, insgesamt also für alle $v, w \in V$, $a, b \in \mathbb{K}$ gilt:

(1) $\langle v, w \rangle = \overline{w, v}$
(2) $\langle av_1 + v_2, bw_1 + w_2 \rangle = a\overline{b} \langle v_1, v_2 \rangle + a \langle v_1, w_2 \rangle + \overline{b} \langle v_2, w_1 \rangle + \langle v_2, w_2 \rangle$
(3) $v \neq 0 \implies \langle v, v \rangle > 0$. ◁

Definition (unitärer $\mathbb{K}$-Vektorraum). Falls auf V ein Skalarprodukt existiert, so heißt V ***unitär***. Falls $\mathbb{K} = \mathbb{R}$ nennt man V auch ***euklidisch***. ◁

Definition (Norm). Eine Abbildung $\| \cdot \| : V \mapsto \mathbb{R}$ heißt ***Norm*** auf V, falls für alle $x, y \in V$ und $c \in \mathbb{K}$ gilt:

(1) $\|x\| = 0 \iff x = 0$
(2) $\|c \cdot x\| = |c| \cdot \|x\|$
(3) $\|x + y\| \leq \|x\| + \|y\|$ ◁

[13] E_k bezeichnet die $k \times k$-Einheitsmatrix

[14] Offenbar sind für $\mathbb{K} = \mathbb{R}$ alle semilinearen Abbildungen linear.

Korollar. *Für jede Norm $\|\cdot\|$ auf V gilt $\|x\| > 0$ für alle $x \in V$ mit $x \neq 0$.* •

Korollar. *Jedes Skalarprodukt auf V induziert eine Norm auf V, nämlich $\|x\| \mapsto \sqrt{\langle x, x\rangle}$.* •

Definition (normierter $\mathbb{K}$-Vektorraum)**.** Falls auf V eine Norm existiert, so heißt V ***normiert***. ◁

Definition (unitäre Abbildung)**.** Eine lineare Abbildung $f : V \to V$ heißt ***unitär***, falls $\langle f(x), f(y)\rangle = \langle x, y\rangle$ für alle $x, y \in V$. Falls $\mathbb{K} = \mathbb{R}$ nennt man f auch ***orthogonal***. ◁

Satz. *Jede unitäre Abbildung f ist bijektiv, winkeltreu und längenerhaltend. f ist genau dann unitär, wenn $f^{-1} = f^*$, insbesondere f o l g t daraus $|\det f| = 1$.* •

Korollar. *Jede orthonormale Basistransformation ist unitär.* •

6.2. **Tensorprodukt.**

Strukturell kann man sich einen Tensor wie eine als Vektor aufgefaßte Matrix mit bestimmten Eigenschaften vorstellen. Tensoren werden normalerweise n i c h t zur Beschreibung von Abbildungen verwendet, sondern dazu, V und W in besonders angenehmer Art und Weise zu einem einzigen Vektorraum $V \otimes W$ zu kombinieren. Für die Anschauung genügt es häufig, sich $v \otimes w$ als die $(\dim V \times \dim W)$-Matrix vw^T vorzustellen, wobei $\otimes$ offenbar eine bilineare Abbildung ist. Zusammen mit Bemerkung 1 und Satz 2 sollte man für das eigentliche Anliegen des Vortrags gewappnet sein.

Definition (Tensorprodukt, Tensor)**.** Seien $(v_i)_{i\in I}$, $(w_j)_{j\in J}$ Basen von V bzw. W. Wir wählen zudem eine Basis $(b_{ij})_{i\in I, j\in J}$ des $\mathbb{K}^{(I\times J)} =: V \otimes W$. Dann gibt es (vgl. z.B. 6.2.3 in [Fis97]) genau eine bilineare Abbildung $t : V \times W \to V \otimes W$ mit $t(v_i, w_j) = b_{ij}$ für alle $i \in I$, $j \in J$.

Für $t(v, w)$ schreiben wir $v \otimes w$. Wir nennen $V \otimes W$ ***Tensorprodukt*** von V und W. Die Elemente von $V \otimes W$ heißen ***Tensoren***. ◁

Bemerkung 1. In vorstehender Definition kann man $\mathbb{K}$ durch einen beliebigen Körper K ersetzen und $V \otimes W$ hängt nicht von der Wahl der Basen in V und W ab.

Sind $(v_i)_{i\in I}$, $(w_j)_{j\in J}$ Basen von V bzw. W, dann ist $(v_i \otimes w_j)_{i\in I, j\in J}$ eine Basis von $V \otimes W$ und für $x \in V \otimes W$ gibt es $\lambda_{ij} \in \mathbb{K}$ so, daß

$$
\begin{aligned}
x &= \sum_{i,j} \lambda_{ij}(v_i \otimes w_j) = \sum_i \sum_j \lambda_{ij}(v_i \otimes w_j) = \sum_i \sum_j v_i \otimes (\lambda_{ij} w_j) \\
&= \sum_i v_i \otimes \left(\sum_j \lambda_{ij} w_j \right) = \sum_i v_i \otimes w'_i
\end{aligned}
$$

Jedes x läßt sich also als Summe von $\#I$ Tensoren der Form $v_i \otimes w'_i$ mit $w'_i \in W$ bzw. als Summe von $\#J$ Tensoren der Form $v'_j \otimes w_j$ mit $v'_j \in V$ eindeutig darstellen.

Damit ist t in der Regel offenbar nicht surjektiv. Zudem ist t in der Regel nicht injektiv, denn $v \otimes w = (\lambda v) \otimes \left(\frac{1}{\lambda} w\right)$ für $\lambda \in \mathbb{K}$. ◁

Beispiel. Wir betrachten $\mathbb{K}^m \otimes \mathbb{K}^n$. Wir wählen die Standardbasen in $\mathbb{K}^m$, $\mathbb{K}^n$, $\mathbb{K}^{m \cdot n}$ und können so leicht die bilineare Abbildung t konstruieren, denn wir fordern $t(e_i, e_j) = e_{(i-1)\cdot n+j}$. Seien $v = \sum_{i=1}^m \lambda_i e_i$, $w = \sum_{j=1}^n \mu_j e_j$ dann erhält man

$$v \otimes w := t(v,w) = \sum_{i,j} \lambda_i \mu_j t(e_i, e_j) = \sum_{i,j} \lambda_i \mu_j e_{(j-1)\cdot n+i} \simeq (\lambda_i \mu_j)_{i,j} = vw^\intercal.$$

Das t bilinear ist, ist offensichtlich, denn $(\lambda v) \otimes w$ entspricht der Multiplikation aller Spalten der $v \otimes w$ entsprechenden Matrix mit λ und $v \otimes (\lambda w)$ der Multiplikation der Zeilen von $v \otimes w$ mit λ. ◁

Beispiel. Eine Anwendung ist die so genannte K o m p l e x i f i z i e r u n g eines $\mathbb{R}$-Vektorraumes V. Dazu beachten wir $\mathbb{C} \simeq \mathbb{R}^2$. Dann erhalten wir mit $z \in \mathbb{C}$, $v \in V$:

$$v \otimes z \simeq vz^\intercal = \begin{pmatrix} \operatorname{Re} z \cdot v & \operatorname{Im} z \cdot v \end{pmatrix}$$

Falls $(v_j)_{j\in J}$ eine Basis von V ist, so ist offensichtlich $(v_j \otimes 1)_{j\in J} \cup (v_j \otimes \mathrm{i})_{j\in I}$ eine Basis von $V \otimes \mathbb{C}$. Damit gilt für $x \in V \otimes \mathbb{C}$:

$$x = \sum_{j\in J} \lambda_j (v_j \otimes 1) + \mu_j (v_j \otimes \mathrm{i}) = \sum_{j\in J} (v_j \otimes \lambda_j) + (v_j \otimes \mathrm{i}\,\mu_j) = \sum_{j\in J} v_j \otimes (\lambda_j + \mathrm{i}\,\mu_j),$$

Wir können also, wie schon weiter oben bemerkt, jeden Tensor $x \in V \otimes \mathbb{C}$ eindeutig in der Form $x = \sum v_j \otimes z_j$ mit $z_j \in \mathbb{C}$ schreiben. Dies ermöglicht es uns, eine Multiplikation $\mathbb{C} \times (V \otimes \mathbb{C}) \to V \otimes \mathbb{C}$ wie folgt zu definieren:

$$\begin{aligned} (a + \mathrm{i}\,b)x &:= (a + \mathrm{i}\,b) \sum_{j\in J} v_j \otimes z_j := \sum_{j\in J} (a + \mathrm{i}\,b)(v_j \otimes z_j) \\ &:= \sum_{j\in J} a(v_j \otimes z_j) + \mathrm{i}\,b(v_j \otimes z_j) := \sum_{j\in J} v_j \otimes (az_j) + v_j \otimes (\mathrm{i}\,bz_j) \\ &= \sum_{j\in J} v_j \otimes ((a + \mathrm{i}\,b)z_j) \end{aligned}$$

Damit können wir $V \otimes \mathbb{C}$ offenbar als $\mathbb{C}$-Vektorraum mit der Basis $(v_j \otimes 1)_{j\in J}$ auffassen. ◁

Satz 2. *Seien $f : V \to V'$ und $g : W \to W'$ lineare Abbildungen. Dann wird durch $(f \otimes g)(v \otimes w) := (f(v)) \otimes (g(w))$ eine lineare Abbildung $f \otimes g : V \otimes W \to V' \otimes W'$ eindeutig festgelegt, und es gilt für darstellende Matrizen A, B von f bzw. g:*

$$(A \otimes B) \simeq (A_i \otimes B_j)_{(i,j)} \simeq (a_{ij} B)_{ij},$$

wobei A_i die Spalten der Matrix von A, B_i die Spalten der Matrix von B, $(A_i \otimes B_j)_{(i,j)}$ die S p a l t e n von $A \otimes B$ und $(a_{ij}B)_{ij}$ eine Matrix mit Matrixeinträgen beschreibt.

Folgende Skizze illustriert den endlichdimensionalen Fall:

$$\begin{aligned} A \otimes B &\simeq \begin{pmatrix} A_1 \otimes B_1 & A_1 \otimes B_2 & \cdots & A_1 \otimes B_m & A_2 \otimes B_1 & A_2 \otimes B_2 & \cdots & A_2 \otimes B_m & \cdots \end{pmatrix} \\ &\simeq \begin{pmatrix} \boxed{a_{11}B} & \boxed{a_{12}B} & \cdots \\ \boxed{a_{21}B} & \boxed{a_{22}B} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \end{aligned}$$

Beweis. Durch die Bilder einer Basis ist jede lineare Abbildung eindeutig festgelegt. Sind nun V, V', W, W' endlichdimensional und A, B darstellende Matrizen von f bzw. g, Dann sind die $A_i \otimes B_j$ genau die Bilder der Basis von $V \otimes W$. $\mathcal{L}(X, Y)$ bezeichne nun die Menge der Linearen Abbildungen von X nach Y. Wir betrachten nun den eingangs bereits erwähnten linearen Isomorphismus $\varphi : V \otimes W \to \mathcal{L}(V, W)$, $\varphi : v \otimes w \mapsto vw^{\intercal}$. Die erste Spalte von $(a_{ij}B)_{ij}$ entspricht gerade dem zeilenweise von links nach rechts und von oben nach unten abgelesenen Produkt $A_1 {B_1}^{\intercal}$, also $\varphi(A_1 \otimes B_1)$, für die restlichen Spalten gilt dies analog. □

6.3. Hilbert-Räume.

Definition (Cauchy-Folge). Eine Folge $(a_i)_{i \in \mathbb{N}} \subset V$ heißt ***Cauchy-Folge***, falls es zu jedem positiven $\varepsilon \in \mathbb{R}$ einen Index $n_0 \in \mathbb{N}$ so gibt, daß $\|a_i - a_j\| \leq \varepsilon$ für alle $i, j \geq n_0$. ◁

Definition (Grenzwert einer Folge). Sei $a \in V$ und $(a_i)_{i \in \mathbb{N}} \subset V$ eine Folge. a heißt ***Grenzwert*** von $(a_i)_{i \in \mathbb{N}}$ und man sagt $(a_i)_{i \in \mathbb{N}}$ ***konvergiert*** gegen a, falls es zu jedem positiven $\varepsilon \in \mathbb{R}$ einen Index $n_0 \in \mathbb{N}$ so gibt, daß $\|a_i - a_{n_0}\| \leq \varepsilon$ für alle $i \geq n_0$. ◁

Definition (vollständig). V heißt ***vollständig***, falls jede Cauchy-Folge in V einen Grenzwert in V hat. ◁

Bemerkung. Die Begriffe Cauchy-Folge, Grenzwert und Vollständigkeit wurden hier nur für normierte $\mathbb{K}$-Vektorräume erklärt und sind Spezialfälle der selben, auf den allgemeineren metrischen Räumen erklärten Begriffe. ◁

Definition (Hilbert-Raum). Ein mit der durch das Skalarprodukt induzierten Norm vollständiger unitärer $\mathbb{K}$-Vektorraum heißt Hilbert-***Raum***. ◁

Korollar. *Jeder endlichdimensionale $\mathbb{K}$-Vektorraum ist ein* Hilbert-*Raum.*

Beweis. $\mathbb{K}^n$ mit der Standardnorm ist vollständig und alle Normen auf $\mathbb{K}^n$ sind äquivalent (zwei Normen $\|\cdot\|_a, \|\cdot\|_b : V \to \mathbb{R}$ heißen äquivalent, falls es $c_2, c_2 > 0$ mit $c_1\|x\|_a \leq \|x\|_b \leq c_2\|x\|_a$ für alle $x \in V$ gibt). □

Korollar. *Für $m, n \in \mathbb{N}$ ist $\mathbb{K}^m \otimes \mathbb{K}^n \simeq \mathbb{K}^{m \times n}$ mit jeder Norm ein* Hilbert-*Raum.*

Beispiel 11. Der Raum $V = \{x \in \mathbb{C}^{\mathbb{N}} | \max\{i | x_i \neq 0\} < \infty\}$ aller endlichen komplexen Folgen mit dem Standardskalarprodukt $\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$ ist kein Hilbert-Raum, denn er enthält Cauchy-Folgen, deren Grenzwert keine endliche komplexe Folge ist, z.B. $\big((\frac{1}{2} + \frac{1}{2}, 0, 0, ...), (\frac{1}{2}, \frac{1}{4} + \frac{1}{4}, 0, 0, ...), ...(\frac{1}{2}, \frac{1}{4}, ..., 2^{-(n-1)}, 2^{-n} + 2^{-n}, 0, 0, ...), ...\big)$.

Hingegen ist der Raum $\ell^2 = \{x \in \mathbb{C}^{\mathbb{N}} | \sum_{i=1}^{\infty} |x_i|^2 < \infty\} \supset V$ der quadratsummierbaren komplexen Folgen mit dem Skalarprodukt $\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$ durchaus ein Hilbert-Raum, und heißt Hilbert'scher Folgenraum. ◁

7. Notation

i	$\sqrt{-1}$ (man beachte den sich von i unterscheidenden Satz)
$\overline{a + \mathrm{i}\, b}$	$a - \mathrm{i}\, b$ (für $a, b \in \mathbb{R}$), komplexe Konjugation
$A^{\intercal}$	die Transponierte von A
A^*	$\overline{A^{\intercal}}$
v^*	$\overline{v^{\intercal}}$ (nicht: das Dual zu v)

$v \otimes w$	das Tensorprodukt von v und w
$\lvert\psi\rangle$	der Zustandsvektor der Observablen ψ
$\lvert\psi\rangle \lvert\varphi\rangle$	$\lvert\psi\rangle \otimes \lvert\varphi\rangle$
$\lvert\psi\varphi\rangle$	$\lvert\psi\rangle \otimes \lvert\varphi\rangle$
$\langle\psi\rvert$	$\lvert\psi\rangle^*$
$\langle\psi\varphi\rvert$	$\lvert\psi\varphi\rangle^*$
$\langle\psi\vert\varphi\rangle$	$\langle\psi\rvert \lvert\varphi\rangle$, Skalarprodukt von $\lvert\psi\rangle$ und $\lvert\varphi\rangle$
$\langle M\rangle$	der Erwartungswert der durch M definierten projektiven Messung
$\Pr[a]$	Wahrscheinlichkeit des Ereignisses a
$\mathbb{E}[X]$	Erwartungswert der Zufallsvariablen X
$\mathfrak{P}(M)$	Die Menge aller Teilmengen von M (Potenzmenge)

Literatur

[ADH97] Leonard M. Adleman, Jonathan DeMarrais, und Ming-Deh A. Huang, *Quantum computability*, SIAM Journal on Computing **26** (1997), Nr. 5, 1524–1540.

[AG01] Dmitri V. Averin und Vladimir J. Goldman, *Quantum computation with quasiparticles of the fractional quantum hall effect*, `www.arXiv.org` **cond-mat/0110193** (2001).

[BBBV97] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, und Umesh Vazirani, *Strengths and weaknesses of quantum computing*, `www.arXiv.org` **quant-ph/9701001** (1997).

[Ben73] Charles H. Bennett, *Logical reversibility of computation*, IBM Journal of Research and Development **17** (1973), 525–532.

[BV97] Ethan Bernstein und Umesh Vazirani, *Quantum complexity theory*, SIAM Journal on Computing **26** (1997), Nr. 5, 1411–1473.

[DEL99] David Deutsch, Artur Ekert, und Rosella Lupacchini, *Machines, logic and quantum physics*, `www.arXiv.org` **math.HO/9911150** (1999).

[Deu85] David Deutsch, *Quantum theory, the Church-Turing principle and the universal quantum computer*, Proceedings of the Royal Society of London Ser. A **A400** (1985), 97–117.

[Fis97] Gerd Fischer, *Lineare Algebra*, Vieweg, 1997.

[FKW00] Michael H. Freedman, Alexei Kitaev, und Zhenghan Wang, *Simulation of topological field theories by quantum computers*, `www.arXiv.org` **quant-ph/0001071** (2000).

[Gro96] Lov K. Grover, *A fast quantum mechanical algorithm for database search*, Proceedings of the 28th Annual ACM Symposium on Theory of Computing (1996), 212–219.

[Kön97] Konrad Königsberger, *Analysis 2*, Springer, 1997.

[Kön99] Konrad Königsberger, *Analysis 1*, Springer, 1999.

[Kra02] Jürg Kramer, *Notizen zur Vorlesung der Linearen Algebra und Analytischen Geometrie I,II*, `http://www-didaktik.mathematik.hu-berlin.de/org/kramer/laag-scr.pdf`, 2002.

[Meg02] Zdzisław Meglicki, *Introduction to quantum computing (Vorlesungsskript)*, `http://beige.ucs.indiana.edu/M743/M743.pdf`, 2002.

[NC00] Michael L. Nielsen und Isaac L. Chuang, *Quantum computing and quantum information*, Cambridge University Press, 2000.

[Sho94] Peter W. Shor, *Algorithms for quantum computation: Discrete logarithms and factoring*, Proceedings of the 35th Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, 1994, SS. 124–134.

[WC98] Colin P. Williams und Scott H. Clearwater, *Explorations in quantum computing*, Springer, 1998.

[Yao93] Andrew C. Yao, *Quantum circuit complexity*, Proceedings of the 34th Annual Symposium on Foundations of Computer Science (Los Alamitos, CA), IEEE Computer Society Press, 1993, SS. 352–361.

E-mail address: `karge@informatik.hu-berlin.de`