

# Freiheit und Berechenbarkeit

Von Bernhard LAUTH (München)

In welchem Umfang wird menschliches Verhalten durch empirische Gesetzmäßigkeiten bestimmt? Und in welchem Umfang ist menschliches Verhalten berechenbar bzw. vorhersagbar? Dies sind Fragen von grundlegender Bedeutung für die empirischen Geistes- und Sozialwissenschaften, wenn es darum geht, das Verhalten von einzelnen Personen oder das Verhalten von komplexen sozialen Systemen zu analysieren.

In den Naturwissenschaften kommt man zu Vorhersagen (z. B. über das Verhalten von physikalischen Systemen) auf der Grundlage von mathematischen Modellen, die in Verbindung mit geeigneten Anfangsbedingungen eindeutige Vorhersagen über das Verhalten des Systems ermöglichen. Das Standardschema von Erklärungen und Prognosen in der Physik besteht (im Wesentlichen) aus drei Schritten:

- (1) Man formuliert ein System von Differentialgleichungen (den sogenannten „Bewegungsgleichungen“), die das Verhalten des Systems beschreiben.
- (2) Man bestimmt die Anfangsbedingungen, also den Zustand des Systems zu einem beliebig herausgegriffenen Anfangszeitpunkt  $t_0$ .
- (3) Aus den Bewegungsgleichungen und Anfangsbedingungen ergeben sich mathematisch eindeutige Lösungen, die das Verhalten des Systems zu allen Nachfolgezeitpunkten eindeutig festlegen. Das Schema hat also die Form:<sup>1</sup>

- (1) Bewegungsgleichungen
- (2) Anfangsbedingungen

---

(3) (eindeutige) Lösungen

Die Vorhersagbarkeit bei physikalischen Systemen scheint daher auf zwei wesentlichen Voraussetzungen zu basieren, nämlich (1) Gesetzmäßigkeit und (2) Berechenbarkeit. Gesetzmäßigkeit bedeutet, daß das Verhalten des Systems durch deterministische Modelle beschrieben werden kann, d. h. die Bewegungsgleichungen müssen mathematisch zu beschaffen sein, daß sie zu vorgegebenen Anfangs- und Randbedingungen eindeutige Lösungen besitzen. Berechenbarkeit bedeutet,

---

<sup>1</sup> Ein einfaches Beispiel sind die Hamiltonschen Bewegungsgleichungen für ein System mit  $n$  Freiheitsgraden. Die Bewegungsgleichungen haben die Form  $\partial H/\partial q_i = -dp_i/dt$ ,  $\partial H/\partial p_i = dq_i/dt$  (für  $i = 1, \dots, n$ ). Aus den Anfangsbedingungen  $q_i(0)$ ,  $p_i(0)$  zur Zeit  $t = 0$  ergeben sich dann eindeutige Lösungen  $q_i(t)$ ,  $p_i(t)$ , für alle Zeitpunkte  $t \geq 0$ .

daß die Lösungen selbst berechenbare Funktionen sein müssen, so daß man ein effektives Verfahren (einen Algorithmus) angeben kann, mit dem die Funktionswerte zu allen nachfolgenden Zeitpunkten effektiv ermittelt werden können.

Im Gegensatz zu gewöhnlichen physikalischen Systemen besitzt der Mensch Bewußtsein und einen freien Willen (so meinen wir zumindest), d. h. das Verhalten und die Handlungsweisen einer Person werden maßgeblich durch ihre eigenen Überlegungen und Entscheidungen bestimmt. Willensfreiheit scheint aber unvereinbar mit Berechenbarkeit und Vorhersagbarkeit aufgrund von deterministischen Gesetzmäßigkeiten. Diese intuitive Argumentation beruht auf folgenden Annahmen:

- (1) Berechenbarkeit setzt Determinismus voraus.
  - (2) Determinismus ist unvereinbar mit Willensfreiheit.
- 
- (3) Berechenbarkeit ist unvereinbar mit Willensfreiheit.

Ich werde mich in meinem Aufsatz hauptsächlich mit dem Gesichtspunkt der Berechenbarkeit befassen und nicht so sehr mit der Frage der Gesetzmäßigkeit.<sup>2</sup> Der intuitive Begriff der Berechenbarkeit enthält verschiedene Bedeutungskomponenten, die oft nicht klar genug voneinander getrennt werden. Man kann Berechenbarkeit nämlich sowohl in einem engeren, mathematischen wie auch in einem weitergehenden, epistemischen Sinn verstehen. Die mathematische Bedeutung ist hinreichend klar: Eine Funktion heißt berechenbar, wenn man einen Algorithmus (z. B. ein Computerprogramm oder eine Turing-Maschine) angeben kann, mit dem alle Funktionswerte mechanisch und in endlicher Zeit beliebig genau berechenbar sind.

Berechenbarkeit im epistemischen Sinn bedeutet dagegen eher so etwas wie Vorhersagbarkeit. Vorhersagbarkeit ist aber keine mathematische, sondern eher eine epistemische Kategorie: sie bezieht sich auf unsere epistemischen Möglichkeiten und Ressourcen. Es ist wichtig, sich klar zu machen, daß physikalische Systeme in diesem epistemischen Sinne *oft nicht berechenbar* (d. h. vorhersagbar) sind, auch wenn sie an sich (im mathematischen Sinn) berechenbar sein mögen:

---

<sup>2</sup> Es gibt bekanntlich eine ganze Reihe von Autoren, die die Auffassung vertreten haben, daß Willensfreiheit und Determinismus vereinbar sind (so z. B. G. E. Moore in seiner Ethik oder R. Carnap, *Foundations of Physics* [New York 1966] besonders in Kap. 22). Ich werde diese Frage hier nicht näher diskutieren, weil dies für meine Argumentationsstrategie nicht erforderlich ist, möchte aber darauf hinweisen, daß ich die These für abwegig halte, d. h. ich verwende den Freiheitsbegriff in einem Sinn, der mit deterministischen Verhaltensmodellen inkompatibel ist. Auf den sogenannten „weichen“ Determinismus (der die Vereinbarkeit von Willensfreiheit und Determinismus unterstellt) bezieht sich das alte Diktum von Kant aus der „Kritik der praktischen Vernunft“, wonach die „dogmatischen Lehrer der Metaphysik ... mehr ihre Verschmitztheit als Aufrichtigkeit darin bewiesen (hätten), daß sie diesen schwierigen Punkt soweit wie möglich aus den Augen brachten, in der Hoffnung, daß, wenn sie davon gar nicht sprächen, auch wohl niemand leichtlich an ihn denken würde“ (AAV, 103). Daß die modernen „Lehrer der Metaphysik“ den „schwierigen Punkt“ nicht länger verschweigen, sondern offen thematisieren, wäre insofern als Fortschritt zu bewerten. Ein „locus classicus“ für die Diskussion über die Kompatibilitätsthese ist natürlich van Inwagens „*Essay on Free Will*“ (Oxford 1983).

Die Vorhersagbarkeit kann z. B. daran scheitern, daß das System zu komplex ist, um die Bewegungsgleichungen explizit zu formulieren, oder daß man kein Verfahren zur Integration der Bewegungsgleichungen hat. (Man denke zum Beispiel an das bekannte Mehr-Körper-Problem in der klassischen Physik). Die Vorhersagbarkeit kann aber auch daran scheitern, daß die relevanten Anfangsbedingungen nicht genau genug oder nur unvollständig bekannt sind.

Ich werde im Folgenden von Berechenbarkeit immer nur im engeren, also mathematischen Sinn sprechen, der nicht Vorhersagbarkeit impliziert. Die Frage lautet dann, (1) ob und inwieweit Berechenbarkeit und Freiheit vereinbar sind, (2) in welchem Umfang menschliches Verhalten berechenbar und darüber hinaus auch vorhersagbar ist und (3) ob und in welchem Umfang menschliches Verhalten durch deterministische Gesetzmäßigkeiten gesteuert wird. Ich werde die meisten dieser Probleme hier nicht lösen können, und es wäre beim gegenwärtigen Stand der Diskussion auch ziemlich unrealistisch, eine umfassende Klärung in absehbarer Zukunft zu erwarten. Ich betrachte meine Ausführungen daher eher als Präliminarien zu einer möglichen künftigen Lösung, Präliminarien, die vor allem der begrifflichen Klärung dienen sollen.

Ich werde mich im Wesentlichen auf drei Thesen beschränken, die ich hier schon vorwegnehmend andeuten möchte: Die *erste These* besagt, daß menschliches Verhalten in allen relevanten Aspekten durch berechenbare Funktionen (also durch Algorithmen) beschreibbar ist, die ich als Verhaltensparameter bezeichnen werde. Die Berechenbarkeit dieser Funktionen folgt nicht aus irgendwelchen philosophischen Annahmen oder Voraussetzungen, sondern aus trivialen mathematischen Überlegungen. *Zweitens* werde ich zu zeigen versuchen, daß Berechenbarkeit keinerlei Form von Determinismus impliziert, d. h. man kann von der Berechenbarkeit nicht auf die Existenz von deterministischen Gesetzmäßigkeiten schließen, die das Verhalten steuern. Mit anderen Worten: Ich werde die erste Prämisse in dem oben erwähnten intuitiven Argument in Frage stellen. Berechenbarkeit steht daher *drittens* nicht im Widerspruch zur Freiheit, man kann aber die Berechenbarkeit in dem erwähnten (und noch weiter zu präzisierenden) Sinn als eine Bedingung oder Restriktion auffassen, mit der jede vernünftige philosophische Konzeption der Willensfreiheit vereinbar sein muß.

\*

Ich werde im Folgenden öfter von physikalischen, biologischen oder sozialen *Prozessen* und *Systemen* sprechen, speziell von deterministischen Prozessen, berechenbaren Prozessen usw. Daher wird es sich als zweckmäßig erweisen, eine *Standardbeschreibung* zur Verfügung zu haben, die gleichzeitig präzise und hinreichend allgemein ist, um die verschiedenartigen Anwendungsfälle abzudecken. Diese Standardbeschreibung besteht im Wesentlichen aus drei Komponenten:

(1) einem endlichen Zeitintervall  $T \subset \mathbb{R}$ , (2) einer Menge  $Z$  von „möglichen Zuständen“ des Systems und (3) einer Funktion  $f: T \rightarrow Z$ , die den Zustand des Systems für jeden Zeitpunkt aus  $T$  beschreibt.

$T$  entspricht dem *Beobachtungszeitraum*, in dem der Prozeß stattfindet. Alle Zeitpunkte aus dem Intervall werden wie üblich durch reelle Zahlen dargestellt. Anstelle eines Intervalls werde ich gelegentlich auch diskrete Folgen von aufeinanderfolgenden Zeitintervallen betrachten. In diesem Fall hat  $T$  die Form  $T = \{t_1, \dots, t_n\}$ . Ich bezeichne die Menge  $Z$  aller möglichen Zustände als den „Zustandsraum“ des Systems. Die *Zustandsfunktion*  $f$  gibt also zu jedem Zeitpunkt den „Ort“ des Systems im Zustandsraum an. Ich setze dabei voraus, daß es sich um einen topologischen oder metrischen Raum handelt, so daß man über benachbarte Zustände und konvergente Folgen von Zuständen sprechen kann. Wenn die Zustandsfunktion stetig ist, sprechen wir von einem *stetigen Prozeß*.<sup>3</sup> Wenn dagegen der Definitions- und Wertebereich von  $f$  aus einer diskreten (abzählbaren) Menge von Zeitpunkten und Zuständen besteht, sprechen wir von einem *diskreten Prozeß*.<sup>4</sup> Auf ganz analoge Weise werden wir später von *berechenbaren Prozessen* sprechen, wenn die Zustandsfunktion rekursiv berechenbar ist.

Einige Beispiele sollen zur Erläuterung dieser Definitionen dienen:

*Beispiel 1: Hamilton-Mechanik:* Wir betrachten dazu ein mechanisches System mit  $n$  Freiheitsgraden, z.B. die Planetenbahnen im Sonnensystem oder die Schwingungen eines Pendels im Schwerfeld der Erde. Der Zustand des Systems zu einem bestimmten Zeitpunkt wird durch Angabe von  $n$  Orts- und Impulskoordinaten definiert.<sup>5</sup> Der Zustandsraum ist dann der  $2n$ -dimensionale Phasenraum  $Z = \mathbb{R}^{2n}$ , und die Dynamik des Systems (also die Bewegung des Systems im Phasenraum) wird durch ein System von partiellen Differentialgleichungen erster Ordnung beschrieben, die als Hamiltonsche Bewegungsgleichungen bekannt sind.<sup>6</sup> Es handelt sich dabei um einen stetigen Prozeß, weil alle Komponenten  $q_1$  bis  $p_n$  des Zustandsvektors stetige Funktionen der Zeit sind.

*Beispiel 2: Quantenmechanik:* Ein anderes Beispiel ist der Zustandsbegriff der Quantenmechanik. In diesem Fall ist der Zustandsraum ein (unendlich-dimensionaler, vollständig normierter) Vektorraum (Hilbertraum). In der sogenannten „Ortsdarstellung“ können wir dafür den Raum aller komplexwertigen, quadratintegrierbaren Funktionen  $\psi: \mathbb{R}^n \rightarrow \mathbb{C}$  mit  $\int |\psi|^2 dx_1 \dots dx_n = 1$ , wählen.<sup>7</sup> Ein Zustandsvektor bildet also jedes Koordinaten- $n$ -tupel (sagen wir für ein System mit  $n$  Freiheitsgraden) in die komplexen Zahlen ab.<sup>8</sup> Der quantenphysikalische Zu-

<sup>3</sup> Genauer formuliert: Ein Prozeß  $(T, Z, f)$  heißt stetig, wenn  $T \subset \mathbb{R}$  ein (offenes oder abgeschlossenes) Intervall ist und wenn  $f: T \rightarrow Z$  stetig (in Bezug auf die Topologie des Zustandsraums) ist, d.h. für jede offene Menge  $O \subset Z$  ist das Urbild  $f^{-1}(O)$  offen in der natürlichen Topologie der reellen Zahlen.

<sup>4</sup> Es kann also Prozesse geben, die weder stetig noch diskret im Sinne der angegebenen Definitionen sind, z.B. wenn der Definitionsbereich  $T$  der Zustandsfunktion überabzählbar ist und die Funktion  $f: T \rightarrow Z$  eine nicht-leere Menge von Unstetigkeitsstellen besitzt.

<sup>5</sup> Die Zustandsfunktion hat also die Form  $f(t) = \langle q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t) \rangle$ , wobei  $q_1(t), \dots, q_n(t)$  die Ortskoordinaten des Systems und  $p_1(t), \dots, p_n(t)$  die zugehörigen (kanonischen) Impulskoordinaten sind.

<sup>6</sup> Vgl. Fußnote 1.

<sup>7</sup>  $\mathbb{C}$  ist die Menge der komplexen Zahlen.

<sup>8</sup> Eine alternative Möglichkeit für die Darstellung der Zustände ist der Hilbertsche Folgenraum  $l_2$ . Ein Zustandsvektor ist dann darstellbar als eine abzählbar-unendliche Folge von komplexen Zahlen  $c_n$ ,  $n = 1, 2, \dots$ , wobei die Summe der Betragsquadrate  $|c_n|^2$  gegen einen endlichen Wert konvergieren muß.

stands begriff hat keine unmittelbar anschauliche Bedeutung. Seine Bedeutung besteht darin, daß durch den Zustand des Systems zu einem bestimmten Zeitpunkt eine Wahrscheinlichkeitsverteilung für alle Observablen des Systems festgelegt wird.

Für manche Anwendungen ist es zweckmäßiger, eine *diskrete* Folge  $T = \{t_1, \dots, t_n\}$  von Zeitpunkten oder Zeitintervallen zu betrachten, z. B. eine Reihe von  $n$  aufeinanderfolgenden Kalenderjahren oder Stichtagen. Wenn darüber hinaus auch die Zustandsfunktion nur diskrete Werte annehmen kann, sprechen wir von einem *diskreten* Prozeß.

*Beispiel 3: Makroökonomie.* Ein einfaches Beispiel für einen diskreten Prozeß ist die Entwicklung einer Volkswirtschaft über eine Reihe von aufeinanderfolgenden Kalenderjahren  $t = 1, \dots, n$ . Wir können den Zustand des Systems durch die üblichen Größen der volkswirtschaftlichen Gesamtrechnung beschreiben, z. B. das Bruttosozialprodukt zu Marktpreisen und Faktorkosten, die Abschreibungen, den öffentlichen und privaten Konsum, die Bruttoinvestitionen der Unternehmen und der öffentlichen Haushalte usw. Der Zusammenhang zwischen den verschiedenen Größen wird „ex post“ durch die volkswirtschaftliche Gesamtrechnung hergestellt. Die Zustandsfunktion hat dann die Form  $f(t) = \langle Y_m^b(t), C(t), I^b(t), \dots \rangle$  und der entsprechende Zustandsraum ist  $Z \subseteq \mathbb{N}^n$ , wobei  $n$  der Anzahl der im Modell berücksichtigten Parameter (Rechnungsgrößen) entspricht.<sup>9</sup>

Für das Problem der Entscheidungsfreiheit sind vor allem deterministische Prozesse von Bedeutung. Das traditionelle Paradigma dafür stammt aus der klassischen Mechanik und Elektrodynamik. Das Verhalten von klassischen Teilchen oder Feldern kann durch ein System von (einfachen oder partiellen) Differentialgleichungen beschrieben werden, die zu vorgegebenen Anfangs- und Randbedingungen eindeutige Lösungen besitzen. Das bedeutet, daß durch Angabe des Zustands zu einem willkürlich herausgegriffenen „Anfangszeitpunkt“  $t_0$  alle Nachfolgestände (und natürlich auch alle vorhergehenden Zustände) eindeutig bestimmt sind, d. h. der Zustand des Systems zu irgendeinem Zeitpunkt  $t + x$  ist eindeutig festgelegt durch den Zustand des Systems zur Zeit  $t$ . Im Rahmen der von uns gewählten Standardbeschreibung läßt sich diese Aussage folgendermaßen präzisieren: Wir sprechen von einem *potentiell deterministischen Prozeß*,<sup>10</sup> wenn es eine einparametrische Familie von Operationen  $U_x: Z \rightarrow Z$  auf dem Zustandsraum des Systems gibt, die alle Zustandsübergänge eindeutig festlegt. Das heißt konkret: Wenn sich das System zu einer Zeit  $t$  in einem Zustand  $z$  befindet, dann geht das System zu allen Zeitpunkten  $t + x$  in die durch den Operator bestimmten Nachfolgestände  $U_x(z)$  über, d. h. es gilt:  $f(t+x) = U_x(f(t))$  für alle  $t \in T$

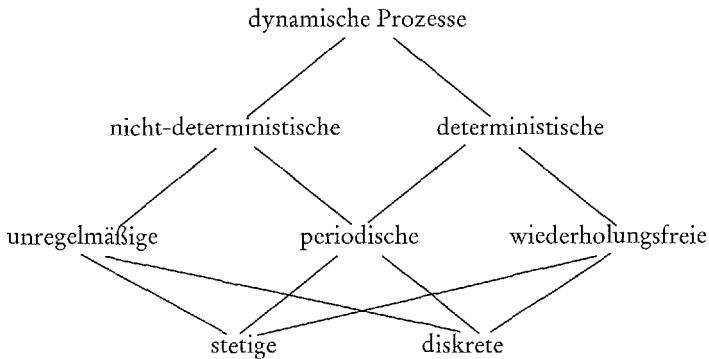
<sup>9</sup>  $Y_m^b$  = das Bruttosozialprodukt zu Marktpreisen,  $C$  = der gesamte (private und öffentliche) Konsum,  $I^b$  = die Bruttoinvestitionen (Unternehmen und öffentliche Haushalte),  $D$  = die Abschreibungen,  $Y_m^n = Y_m^b - D$  = das Nettosozialprodukt zu Marktpreisen,  $X$  = Export,  $M$  = Import, usw. Man kann davon ausgehen, daß alle Rechnungsgrößen nur ganzzahlige Werte annehmen.

<sup>10</sup> Der Grund für die Bezeichnung „potentiell deterministisch“ wird aus den nachfolgenden Erklärungen deutlich werden.

und alle  $x \in \mathbb{R}$  mit  $t+x \in T$ . Der Parameter  $x$  gibt dabei den zeitlichen Abstand zwischen Anfangszustand und Nachfolgezustand an. Wir bezeichnen  $U_x$  auch als *Entwicklungsoperator* des Systems, weil er die zeitliche Entwicklung der Systemzustände festlegt.<sup>11</sup> Die angegebene Gleichung entspricht dann der *Bewegungsgleichung* des Systems.

Ein einfaches Beispiel sind *periodische Systeme*, z. B. ein harmonisches Pendel oder die Planetenbewegungen im Sonnensystem. Periodische Systeme sind dadurch gekennzeichnet, daß das System nach Ablauf einer bestimmten konstanten Periode  $\theta$  wieder in seinen Anfangszustand zurückkehrt, so daß gilt:  $f(t) = f(t + \theta)$  für alle Zeitpunkte  $t$  aus dem Beobachtungszeitraum  $T$ .<sup>12</sup>

Nicht-periodische Systeme sind entweder *unregelmäßig*, d. h. es treten nicht-periodische Zustandswiederholungen  $f(t') = f(t)$  auf, oder kehren nie wieder in ihren Anfangszustand zurück. Im letzteren Fall sprechen wir von wiederholungsfreien Systemen (d. h.  $f(t) \neq f(t')$  für alle  $t \neq t'$ ). Ein physikalisches Beispiel dafür sind Streuungen, z. B. wenn ein negativ geladenes Teilchen durch einen ortsfesten Atomkern aus seiner Bahn abgelenkt wird. Die entstehende Bahn ist eine Hyperbel, so daß das System nie in seinen Ausgangszustand zurückkehrt, wenn es nicht durch äußere Einwirkungen dazu gezwungen wird.<sup>13</sup> Insgesamt erhalten wir das folgende Bild:



Deterministische Systeme können u. U. sehr empfindlich auf kleine Änderungen in den Anfangsbedingungen reagieren. Das bedeutet, daß man bereits bei

<sup>11</sup> Wenn man die ganze reelle Achse als Zeitintervall betrachtet und den Parameter  $x$  über alle reellen Zahlen laufen läßt, dann bilden die Operatoren  $U_x$  eine ein-parametrische Gruppe mit  $U_0$  als neutralem Element,  $U_{-x}$  als inversem Element und der assoziativen Operation  $U_x \circ U_y = U_{x+y}$ .

<sup>12</sup> D. h.:  $U_\theta(z) = z$  für alle  $z \in Z$ . Allerdings sind nicht alle periodischen Systeme deterministisch im Sinne der Definition. (Gegenbeispiel: die Zustandsfolge  $0, 1, 0, 2, 0, 1, 0, 2, 0, \dots$ ).

<sup>13</sup> Rein formal läßt sich für jedes wiederholungsfreie System ein Entwicklungsoperator  $U_x$  angeben: Sei nämlich  $f(t)$  die Zustandsfunktion eines wiederholungsfreien Systems. Wir definieren  $U_x(z) = f(t+x)$ , falls es einen Zeitpunkt  $t$  gibt, in dem  $f(t) = z$  gilt. Nach Voraussetzung kann höchstens ein solcher Zeitpunkt existieren. In allen anderen Fällen sei  $U_x(z) = z$ . Man kann leicht sehen, daß  $U_x$  tatsächlich ein Entwicklungsoperator ist. Daraus folgt: (1) Deterministische Systeme sind entweder periodisch oder wiederholungsfrei. (2) Alle wiederholungsfreien Systeme sind (potentiell) deterministisch. (3) Es gibt periodische Systeme, die nicht deterministisch sind.

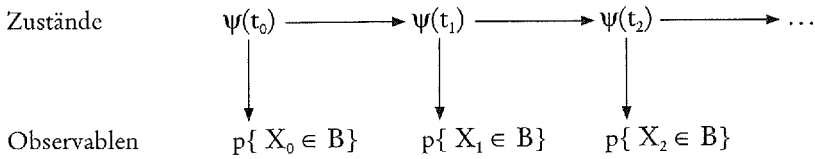
minimalen Veränderungen in den Anfangsbedingungen völlig andere Lösungen für die Bewegungsgleichungen erhält. Wenn die Abweichungen bei den Anfangsbedingungen in der Größenordnung der Meßgenauigkeit liegen, dann lassen sich keine zuverlässigen Vorhersagen über das Verhalten des Systems mehr machen. Man spricht dann gelegentlich von *chaotischen Systemen*. Diese Bezeichnungsweise darf aber nicht darüber hinwegtäuschen, daß es sich dabei trotzdem um deterministische Systeme im klassischen Sinne handeln kann, bei denen alle Zustandsübergänge eindeutig durch die Anfangsbedingungen und Bewegungsgleichungen festgelegt sind.

Im Gegensatz zur klassischen Physik besitzt die Quantenphysik einen inhärent nicht-deterministischen (oder besser: stochastischen) Charakter. Vorhersagen über physikalische Größen wie Orte und Impulse eines atomaren oder subatomaren Systems oder über den Zerfall von radioaktiven Substanzen können nur in einem statistischen Sinn, also mit gewissen Wahrscheinlichkeiten gemacht werden. Die Quantenphysik wird daher oft als eine Widerlegung des klassischen Determinismus interpretiert. Diese Aussage mag zwar im Kern zutreffen, beinhaltet aber dennoch eine gewisse unzulässige Vereinfachung der Situation: auch der Zustand eines quantenmechanischen Systems gehorcht *deterministischen* Gesetzmäßigkeiten, nämlich der bekannten Schrödinger-Gleichung, solange das System nicht durch äußere Einwirkungen gestört wird. Die Schrödinger-Gleichung spielt für die Dynamik eines Quantensystems eine ähnliche Rolle wie die Hamiltonschen oder Lagrange-Gleichungen für die Dynamik klassischer Systeme: zu gegebenen Anfangsbedingungen (Zustand des Systems zur Zeit  $t_0$ ) gibt es eindeutig bestimmte Lösungen. Im Unterschied zur klassischen Physik wird aber der Zustand des Systems nicht durch seinen „Ort“ im Phasenraum, sondern durch einen Zustandsvektor  $\psi$  im Hilbert-Raum beschrieben. Wenn nun der Zustand  $\psi(t_0)$  des Systems zu irgendeinem „Anfangs“-Zeitpunkt  $t_0$  bekannt ist, dann sind alle nachfolgenden Zustände  $\psi(t)$ ,  $t > t_0$ , durch die Schrödinger-Gleichung eindeutig festgelegt.<sup>14</sup> Der indeterministische (bzw. stochastische) Charakter der Theorie zeigt sich erst, wenn das System „gestört“ wird, zum Beispiel durch eine Messung, die am System vorgenommen wird: aus einem gegebenen (Anfangs-) Zustand  $\psi$  kann man nur statistische (Wahrscheinlichkeits-) Vorhersagen über die Meßergebnisse (z.B. bei einer Orts- oder Impulsmessung) ableiten. Eine typische Vorhersage könnte zum Beispiel die folgende Form haben: „Die Wahrscheinlichkeit, daß bei einer Messung einer Orts- (oder Impuls-) Koordinate die Observable einen Wert in einem Intervall  $(x,y)$  annimmt, wenn sich das System im Zustand  $\psi(t)$  befindet, beträgt *soundsoviel* Prozent.“

Ich würde die Situation in der Quantenphysik daher als *stochastischen Determinismus* bezeichnen: Stochastischer Determinismus bedeutet, daß man aus der Kenntnis der Anfangsbedingungen zwar keine sicheren Vorhersagen über die Zukunft ableiten kann, daß aber durch die Anfangsbedingungen eindeutig eine Wahrscheinlichkeitsverteilung für alle Observablenwerte zu allen nachfolgenden

<sup>14</sup> Der entsprechende Entwicklungoperator ist  $U_x = e^{ihx}$ , d. h.  $\psi(t+x) = e^{ihx}\psi(t)$ .

Zeitpunkten festgelegt ist. Diese Situation soll durch das folgende Bild veranschaulicht werden:



Auf der oberen Ebene betrachten wir die physikalischen Zustände des Systems, auf der unteren Ebene die zugehörigen Observablen, also z. B. Orts- und Impulskordinaten des Systems.<sup>15</sup> Die Zustandsübergänge erfolgen in einem klassisch-deterministischen Sinn (gemäß der Schrödinger-Gleichung). Andererseits wird durch den Zustand zu einem Zeitpunkt  $t$  eindeutig eine Wahrscheinlichkeitsverteilung für alle Observablenwerte induziert (wobei man natürlich zu beachten hat, daß es keine gemeinsamen Verteilungen für nicht-kommutierende Observablen gibt). Beide Ebenen zusammengenommen implizieren, daß durch die Anfangsbedingungen (d. h. durch den Zustand des Systems zum Zeitpunkt  $t = 0$ ) eindeutig die Wahrscheinlichkeitsverteilung für jeden Nachfolgezeitpunkt festgelegt wird.

Ich habe oben von *potentiell* deterministischen Systemen gesprochen. Der Grund für diese Einschränkung ist folgender: Der gewöhnliche Sprachgebrauch in Bezug auf deterministische Systeme enthält eine *modale Komponente*, die in der oben angegebenen Definition nicht berücksichtigt wird: Gefordert wird nur, daß sich das *faktische* Verhalten des Systems durch einen geeigneten Entwicklungsoperator  $U$  beschreiben läßt. Ein System ist aber nur dann *im strengen Sinne* deterministisch, wenn das System *zwangsläufig* (im Sinne einer physikalischen Notwendigkeit) aus einem vorgegebenen Anfangszustand  $z_0$  in den entsprechenden Nachfolgezustand  $U_x(z_0)$  übergeht, der durch den Entwicklungsoperator bestimmt ist. In der Sprechweise der „Mögliche-Welten-Semantik“ ließe sich diese Aussage folgendermaßen wiedergeben: In jeder physikalisch möglichen Welt  $w$ , in der das System sich in dem Ausgangszustand  $z_0$  befindet, muß das System (nach Ablauf einer Zeit  $x$ ) in den Nachfolgezustand  $U_x(z_0)$  übergehen.

Statt mit diesen (etwas problematischen) Modalitäten können wir auch mit Wahrscheinlichkeiten arbeiten. Wir betrachten dann deterministische Systeme als Grenzfälle von bestimmten stochastischen (Markoff-) Prozessen.<sup>16</sup> Um über stochastische Prozesse reden zu können, müssen wir unsere Standardbeschreibung geringfügig erweitern und modifizieren: Betrachten wir dazu eine beliebige

<sup>15</sup> Dabei sind  $t_0, t_1, t_2, \dots$  willkürlich herausgegriffene Zeitpunkte aus dem Beobachtungszeitraum, mit  $t_0 < t_1 < t_2 < \dots$ .  $X$  ist eine observable (z. B. eine Orts- oder Impulskordinate, wobei  $X_n =$  der Wert der Observablen zur Zeit  $t_n$  sein soll), und  $B \subset \mathbb{R}$  ist irgendeine (Borel-) Menge von reellen Zahlen, die als Meßwerte auftreten können.

<sup>16</sup> Die Markoff-Eigenschaft entspricht der intuitiven Idee, daß die Wahrscheinlichkeiten nur vom Anfangszustand, aber nicht von der Vorgeschichte abhängen.



Funktion  $f: T \rightarrow Z$ , die alle Zeitpunkte aus dem Beobachtungszeitraum in den Zustandsraum abbildet. Jede solche Funktion beschreibt einen möglichen *Pfad* des Systems durch den Zustandsraum. (In der klassischen Physik spricht man häufig von „Trajektorien“). Jeder möglichen Welt entspricht dann eindeutig ein Pfad  $f \in Z^T$  durch den Zustandsraum. Im stochastischen Fall gibt es zu jedem möglichen Anfangszustand  $z \in Z$  eine Wahrscheinlichkeitsverteilung, die jeder meßbaren Menge von Pfaden eine Wahrscheinlichkeit zuordnet. Wir betrachten nun die Werte der Zustandsfunktion zu einem vorgegebenen Zeitpunkt  $t$  als Zufallsvariable, deren Verteilung durch den Anfangszustand determiniert wird. Im deterministischen Grenzfall gehört zu jedem Anfangszustand ein eindeutig bestimmter Nachfolgezustand, d. h. die Übergangswahrscheinlichkeit kann nur die Werte 0 oder 1 annehmen.

Um diese Idee zu präzisieren, betrachten wir einen stochastischen Prozeß als ein geordnetes Tupel  $(T, Z, (X_t)_{t \in T}, p)$ .  $T$  ist wieder der Beobachtungszeitraum und  $Z$  der Zustandsraum (der wieder als topologischer Raum konstruiert sein soll).  $\Omega := Z^T$  ist die Menge aller möglichen Pfade durch den Zustandsraum und  $\mathcal{O} := \bigcup_{t \in T} B(Z)$  die Gesamtheit aller meßbaren Mengen von Pfaden, wobei  $B(Z)$  die Familie aller Borelmengen im Zustandsraum ist. Jede der Funktionen  $X_t: \Omega \rightarrow Z$  ist dann eine Zufallsvariable (mit  $X_t(f) := f(t)$ ), die den Zustand des Systems zum Zeitpunkt  $t$  angibt.<sup>17</sup> Zu jedem Anfangszustand  $z$  ist  $p^z: \mathcal{O} \rightarrow \mathbb{R}$  eine Wahrscheinlichkeitsverteilung, die jeder meßbaren Menge von Pfaden eine Wahrscheinlichkeit zuordnet und  $P_t(z, B) = p^z \{X_t \in B\}$  ist der zugehörige Markoff-Kern, der die Übergangswahrscheinlichkeiten beschreibt, d. h.:  $P_t(z, B)$  ist die Wahrscheinlichkeit, daß sich das System zur Zeit  $t$  in einem Zustand  $z' \in B$  befindet, wenn  $z$  der Zustand zur Zeit  $t = 0$  war. Im deterministischen Grenzfall gehört zu jedem Anfangszustand  $z$  und zu jedem Zeitpunkt ein eindeutig bestimmter Nachfolgezustand, d. h.  $P_t(z, B)$  kann nur die Werte Null oder Eins annehmen. Man kann leicht sehen, daß es zu jedem deterministischen Prozeß einen entsprechenden Entwicklungsoperator  $U$  gibt:  $U_t(z)$  ist nämlich der eindeutig bestimmte Zustand  $z' \in Z$  mit  $P_t(z, B) = 1$  für alle Borelmengen  $B$  mit  $z' \in B$ .<sup>18</sup>

Am Beispiel der stochastischen Prozesse läßt sich besonders deutlich nachweisen, daß *berechenbare Prozesse nicht deterministisch* sein müssen. Dazu muß allerdings der Begriff der Berechenbarkeit noch genauer präzisiert werden. Ein dynamischer Prozeß ist *berechenbar*, wenn die Zustandsfunktion  $f$  des Systems berechenbar ist, d. h. wenn ein Algorithmus oder eine Turing-Maschine existiert, mit der alle Funktionswerte innerhalb des Beobachtungszeitraums mechanisch berechenbar sind.

Diese Definition hat zunächst nur dann eine klare Bedeutung, wenn alle Zeit-

<sup>17</sup> Der so definierte Prozeß wird auch als „kanonischer Prozeß“ bezeichnet. Für weitere mathematische Einzelheiten vgl. H. Bauer 1978, Kap. XII: Stochastische Prozesse.

<sup>18</sup> Dabei muß vorausgesetzt werden, daß  $Z$  „vernünftige“ topologische Eigenschaften besitzt (z. B. Separabilität oder die Existenz einer abzählbaren Basis). Unter diesen Voraussetzungen folgt aus der Annahme, daß  $P_t(z, B)$  nur die Werte 0 und 1 annehmen kann, daß genau ein Zustand  $z'$  existiert, so daß  $P_t(z, \{z'\}) = 1$  ist.

punkte und Zustände umkehrbar eindeutig durch natürliche Zahlen darstellbar sind (also insbesondere nur für diskrete Prozesse mit abzählbarem Zustandsraum). Man kann den Begriff der Turing-Berechenbarkeit aber auch auf *stetige* Prozesse (mit überabzählbarem Zustandsraum) ausdehnen, wenn gewisse Voraussetzungen erfüllt sind. Genauer gesagt, handelt es sich um zwei Bedingungen, nämlich

(1) Es gibt eine abzählbare Teilmenge  $Z^* \subset Z$  von Zuständen, die dicht im Zustandsraum liegen.<sup>19</sup>

(2) Alle Zustände aus  $Z^*$  sind effektiv und umkehrbar eindeutig durch natürliche Zahlen (Gödelnummern) codierbar, d. h. man kann ein effektives Verfahren angeben, das zu jedem Zustand die entsprechende Gödelnummer berechnet und umgekehrt: ein Verfahren zur Decodierung, das zu jeder Gödelnummer den entsprechenden Zustand angibt.<sup>20</sup>

Eine stetige Zustandsfunktion  $f: T \rightarrow Z$  heißt nun (exakt) *berechenbar*, wenn  $f(t)$  für rationale Argumente  $t \in \mathbb{Q}$  nur Werte in  $Z^*$  annimmt, und wenn darüber hinaus die auf  $\mathbb{Q}$  beschränkte Zustandsfunktion  $f^* := f|_{T \cap \mathbb{Q}}$  Turing-berechenbar (im üblichen Sinne) ist.<sup>21</sup>

Die meisten physikalisch interessanten Zustandsfunktionen sind allerdings nicht exakt, sondern nur *approximativ berechenbar*. Man kann den Begriff der approximativen Berechenbarkeit auf unterschiedliche Weise präzisieren, auf die ich hier nicht im einzelnen eingehen kann. Die einfachste Möglichkeit besteht darin, die Gesamtheit  $Z^T$  aller stetigen Abbildungen vom Beobachtungszeitraum in den Zustandsraum als topologischen Raum zu betrachten. Dann ist eine Funktion  $f \in Z^T$  *approximativ berechenbar* genau dann, wenn es in jeder  $\varepsilon$ -Umgebung von  $f$  eine exakt berechenbare Funktion  $f^*$  gibt, also insbesondere dann, wenn die berechenbaren Funktionen dicht in dem Funktionenraum liegen.<sup>22</sup>

Man könnte nun vermuten, daß deterministische Systeme immer auch Turing-

<sup>19</sup> „Dicht“ bedeutet folgendes: Wir können einen beliebigen „Punkt“  $z$  aus dem Zustandsraum auswählen. Dann gibt es in jeder noch so kleinen Umgebung  $U$  von  $z$  einen Zustand  $z' \in Z^*$  mit  $z' \in U$ . Wir können also jeden Zustand aus  $Z$  beliebig genau durch Zustände aus  $Z^*$  „approximieren“.

<sup>20</sup> In dem oben erwähnten Beispiel des klassischen Phasenraums  $\mathbb{R}^{2n}$  bilden beispielsweise alle Punkte mit rationalen Koordinaten  $Z^* = \mathbb{Q}^{2n}$  eine dichte, abzählbare Teilmenge, die in der beschriebenen Weise durch Gödelnummern codierbar ist. Im Fall der Quantenmechanik handelt es sich bei dem Zustandsraum stets um einen *separablen* Hilbertraum, so daß man immer eine abzählbare Teilmenge von Zustandsvektoren angeben kann, die dicht in  $H$  liegt. In der oben erwähnten „Ortsdarstellung“ handelt es sich dabei aber nicht um Zahlen, sondern um Hilbertvektoren  $\psi: \mathbb{R}^n \rightarrow \mathbb{C}$ . Diese Funktionen müssen also in umkehrbar eindeutiger Weise durch natürliche Zahlen (Gödelnummern) codiert werden, um eine berechenbare Zustandsfunktion zu erhalten. Man kann dies erreichen, indem man eine Basis im Hilbertraum wählt. Jeder Vektor kann dann durch eine abzählbar unendliche Liste von komplexen Koeffizienten dargestellt werden. Wir erzeugen eine abzählbare, dichte Teilmenge, indem wir nur solche Vektoren berücksichtigen, bei denen alle Komponenten bis auf endlich viele Ausnahmen den Wert Null annehmen, wobei die verbleibenden Koeffizienten einen rationalen Real- und Imaginärteil besitzen müssen.

<sup>21</sup> Man beachte, daß wegen der Stetigkeit der Funktion alle Funktionswerte  $f(t)$  eindeutig durch die Werte  $f(t)$  für rationale  $t$  bestimmt sind.

<sup>22</sup> Man beachte, daß dadurch der Begriff der approximativen Berechenbarkeit immer relativ zu einer Topologie definiert wird.

berechenbar (in dem oben beschriebenen Sinn) sein müssen. Aber diese Annahme ist falsch, wie Roger Penrose an einem einfachen Modellbeispiel gezeigt hat.<sup>23</sup> Ich will hier die Einzelheiten des Arguments nicht wiederholen. Im Wesentlichen handelt es sich bei Penroses Beispiel um einen diskreten Prozeß, dessen Zustände umkehrbar eindeutig durch natürliche Zahlen codiert werden können. Die Dynamik des Systems wird dementsprechend nicht durch Differentialgleichungen, sondern durch Differenzgleichungen bestimmt, und zwar in deterministischer Weise, also so, daß es zu gegebenen Anfangsbedingungen stets eindeutig bestimmte Nachfolgezustände gibt. Die entstehende Zustandsfolge ist aber nicht mehr berechenbar, wie sich durch eine einfache Reduktion auf das bekannte Halte-Problem in der Automatentheorie beweisen läßt.

Für unsere Zwecke ist der umgekehrte Zusammenhang wichtiger: Das Beispiel der stochastischen Prozesse zeigt, daß man auch in der umgekehrten Richtung keinen zwingenden Zusammenhang hat, d. h. man kann nicht allgemein von der Berechenbarkeit auf den deterministischen Charakter eines Prozesses schließen. Betrachten wir dazu als einfaches Beispiel einen stochastischen Prozeß, der aus einer endlichen Folge von Würfeln mit einer unverfälschten Münze besteht. Bei jedem Wurf gibt es zwei mögliche Ergebnisse (Kopf oder Wappen), die durch die Zahlen 0 und 1 repräsentiert werden können. Bei  $n$  Würfeln gibt es  $2^n$  mögliche Ergebnisfolgen, z. B.

$$0101101001 \quad (n = 10)$$

Jede Ergebnisfolge besitzt die Wahrscheinlichkeit  $1/2^n$ . Jede Folge von  $n$  Zahlen (0 oder 1) ist eine endliche, primitiv-rekursive Folge von natürlichen Zahlen. Somit sind alle Pfade des Prozesses berechenbare Funktionen, obwohl es sich *per Definition* um einen nicht-deterministischen Prozeß handelt.

\*

Wir betrachten im Folgenden auch den Menschen als ein (sehr komplexes) dynamisches System mit physikalischen, biologischen und mentalen (psychologischen) „System-Komponenten“. Wenn wir unsere Standardbeschreibung dynamischer Prozesse auch auf das menschliche Verhalten anwenden wollen, stehen wir zunächst vor der folgenden Frage: Was verstehen wir genau unter dem „Verhalten“ einer Person?

Menschliches Verhalten ist bekanntlich eine sehr vielschichtige Angelegenheit, und man kann dieses Verhalten auf unterschiedlichen Ebenen beschreiben: das Verhalten im Straßenverkehr, die Eßgewohnheiten, die durchschnittliche Frequenz von Theater- oder Kinobesuchen, das Sexualverhalten und viele andere

<sup>23</sup> Vgl. R. Penrose, *Shadows of the Mind* (Oxford 1994) 30.

Faktoren mehr machen in ihrer Gesamtheit das aus, was wir gewöhnlich abkürzend unter den Begriff „menschliches Verhalten“ subsumieren.

Es liegt auf der Hand, daß wir im Rahmen der Standardbeschreibung immer nur *Teilaspekte* des Gesamtverhaltens einer Person explizit erfassen und beschreiben können. Ich werde in diesem Zusammenhang von verschiedenen „Verhaltensparametern“ sprechen, mit denen bestimmte Aspekte des Verhaltens modelliert werden können. Unter einem *Verhaltensparameter* verstehe ich allgemein eine Zustandsfunktion, die über einem endlichen (stetigen oder diskreten) Beobachtungszeitraum definiert ist, wobei die Funktionswerte maßgeblich durch die Handlungen und Entscheidungen der betreffenden Person bestimmt werden.

Einige einfache Beispiele mögen verdeutlichen, was damit gemeint ist:

*Beispiel 1:* Wir betrachten das Konsumverhalten eines Rauchers über einen längeren Beobachtungszeitraum  $T$ . Wir beschreiben das Verhalten der Person durch eine *diskrete* Funktion  $f(t)$ , die für jeden Tag  $t = 1, 2, \dots$ , (sagen wir über einen Zeitraum von zwölf Monaten), den täglichen Zigarettenkonsum angibt, also beispielsweise,  $f(1) = 10$  Zigaretten,  $f(2) = 11$  Zigaretten,  $f(3) = 5$  Zigaretten, usw. Wir könnten diese Beschreibung noch beliebig verfeinern, indem wir zusätzliche Informationen berücksichtigen, zum Beispiel durch Angabe der Zigarettenmarke oder durch Angaben über den Nikotingehalt usw.<sup>24</sup>

*Beispiel 2:* Auf ähnliche Weise könnten wir das allgemeine Konsumverhalten einer Person  $P$  beschreiben durch eine Funktion  $f(t)$  mit ganzzahligen Argumenten und Funktionswerten, die den Gesamtwert aller (Konsum-) Ausgaben der Person für  $n$  aufeinanderfolgende Rechnungsperioden (Tage, Wochen oder Monate) beschreibt. Wieder können wir die Genauigkeit der Beschreibung verbessern, indem wir zusätzliche Informationen in die Funktion einbauen: beispielsweise könnten wir die Ausgaben aufgliedern nach verschiedenen Konsumgütern und Dienstleistungen oder durch Angaben über die Zahlungsweise (z. B. Barzahlung, Bezahlung mit Scheck oder Kreditkarte, Ratenkauf usw.).

Weitere Beispiele für Verhaltensparameter sind unschwer anzugeben: Die Gesamtdauer (in Stunden), die die Versuchsperson jeden Tag mit Fernsehen verbringt, die Auswahl der Programme, die Anzahl der täglichen Gesprächspartner, die relative Häufigkeit von Nebensätzen und Konjunktiven in mündlichen oder schriftlichen Äußerungen der Versuchsperson usw.

*Beispiel 3:* In den vorhergehenden Beispielen hatten wir es mit *diskreten* Parametern zu tun, d. h. Argumente und Funktionswerte lassen sich umkehrbar eindeutig durch natürliche Zahlen darstellen. Es gibt aber natürlich auch stetige Funktionen, die das Verhalten einer Person beschreiben. Ein besonders wichtiges Beispiel ist die räumliche Bewegung (Ortsveränderung). Wenn es nicht um motorische Einzelheiten, sondern nur um eine grobe Lokalisierung des Aufenthaltsorts geht, könnte man zum Beispiel die Person mit einem Minisender versehen. Die Si-

<sup>24</sup> Wenn es auf dem Markt  $N$  verschiedene Marken gibt, dann könnten wir  $f(t)$  ersetzen durch einen „Zustandsvektor“  $f^k(t) = \langle n_1, \dots, n_N \rangle$ , wobei  $n_k$  = die Anzahl der Zigaretten der Marke  $k$ , die von der Person am Stichtag geraucht worden sind. Man erhält die ursprünglichen Funktionswerte  $f(t)$ , indem man über alle Vektorkomponenten aufsummiert, d. h.  $f(t) = n_1 + \dots + n_N$ .

gnale des Senders könnten über einen entsprechenden Empfänger an einen Rechner weitergeleitet werden, der dann das Bewegungsprofil in einem geeigneten zwei- oder dreidimensionalen Koordinatensystem abträgt. Wir erhalten auf diese Weise eine stetige Zustandsfunktion  $f(t) = \langle X(t), Y(t), Z(t) \rangle$ ,  $f: T \rightarrow \mathbb{R}^3$ , wobei die Funktionswerte zu jedem Zeitpunkt die Koordinaten der Person zu dem betreffenden Zeitpunkt angeben. Auch hier läßt sich die Genauigkeit der Beschreibung – zumindest theoretisch – beliebig verfeinern. Im Idealfall würden wir eine komplexe Zustandsfunktion erhalten, die die Koordinaten aller einzelnen Atome und Moleküle aus dem Körper der Testperson angibt (wobei die Genauigkeit der Angaben natürlich durch die Heisenbergschen Unschärferelationen beschränkt bleibt).

Es ist nun unschwer einzusehen, daß alle diskreten und stetigen Verhaltensparameter der oben beschriebenen Art durch berechenbare Funktionen dargestellt oder approximiert werden können, d. h. die Parameterwerte sind durch Algorithmen beliebig genau berechenbar:

(\*) Alle stetigen und diskreten Verhaltensparameter der Form  $f: T \rightarrow \mathbb{N}$  mit  $T = \{t_1, \dots, t_n\} \subset \mathbb{N}$  oder  $f: T \rightarrow \mathbb{R}$  mit  $T = [a, b] \subset \mathbb{R}$  sind exakt oder approximativ Turing-berechenbar.<sup>25</sup>

Diese Aussage folgt nicht aus irgendwelchen philosophischen oder naturwissenschaftlichen Voraussetzungen, sondern aus trivialen mathematischen Gründen. Zur Begründung der These genügt es, die folgenden beiden Tatsachen zu betrachten:

- (1) Jede diskrete Funktion  $f: T \rightarrow \mathbb{N}$  mit ganzzahligen Argumenten und Werten über einem endlichen Definitionsbereich  $T = \{t_1, \dots, t_n\} \subset \mathbb{N}$  ist primitiv-rekursiv berechenbar.
- (2) Jede stetige, reellwertige Funktion  $f: [a, b] \rightarrow \mathbb{R}$  über einem kompakten Intervall  $a, b$  ist approximativ berechenbar.<sup>26</sup>

<sup>25</sup> Tatsächlich gilt die These allgemein für beliebige stetige Zustandsfunktionen  $f: T \rightarrow Z$ , wo  $Z$  ein topologischer Raum mit den oben beschriebenen Zusatzeigenschaften ist, sofern die berechenbaren Funktionen dicht in  $Z^T$  liegen.

<sup>26</sup> Die Aussage (1) bedarf wohl keiner näheren Erläuterung. Die Aussage (2) läßt sich am einfachsten mit dem Approximationssatz von Weierstraß begründen: Danach gibt es zu jeder stetigen Funktion  $f(x)$  über einem kompakten Intervall  $[a, b]$  ein Polynom  $p(x)$  mit rationalen Koeffizienten, so daß der Abstand der Funktionswerte  $|p(x) - f(x)|$  im ganzen Intervall kleiner als ein beliebig vorgegebenes  $\varepsilon$  ist, das heißt:

(i) Zu jeder stetigen Funktion  $f: [a, b] \rightarrow \mathbb{R}$  gibt es ein Polynom  $p(x)$  mit rationalen Koeffizienten, so daß  $|f(x) - p(x)| < \varepsilon$  für alle  $x$  aus  $[a, b]$  gilt.

Sei nun  $C[a, b]$  die Klasse aller stetigen Funktionen über diesem Intervall, d. h.

(ii)  $C[a, b] := \{f: [a, b] \rightarrow \mathbb{R} / f \text{ ist stetig}\}$ .

Wir definieren den Abstand zwischen zwei Funktionen  $f, g$  als das Supremum der Abstände zwischen den Funktionswerten der beiden Funktionen:

(iii)  $d(f, g) := \sup \{ |f(x) - g(x)| / x \in C[a, b] \}$

Mit dem Weierstraß'schen Approximationssatz folgt dann unmittelbar

(iv) Zu jeder stetigen Funktion  $f$  über  $[a, b]$  und jeder noch so kleinen Zahl  $\varepsilon > 0$  gibt es eine exakt berechenbare Funktion  $f^*$  mit  $d(f, f^*) < \varepsilon$ .

Ich möchte noch einmal betonen, daß es sich bei der oben erwähnten These um eine triviale mathematische Feststellung handelt. Um diesen Punkt besonders deutlich zu machen, möchte ich zum Abschluß einige Anmerkungen machen, die zur Vermeidung von Mißverständnissen dienen sollen:

- (1) Die Berechenbarkeitsthese gilt völlig unabhängig von bestimmten philosophischen Annahmen (oder Vorurteilen) über die Natur der Kausalität oder das Verhältnis von Geist und Körper. Es wird also weder eine behavioristische, noch irgendeine funktionalistische oder gar reduktionistische Auffassung über die Natur mentaler Zustände und deren Zusammenhang mit dem beobachtbaren Verhalten einer Person vorausgesetzt.<sup>27</sup>
- (2) Noch wichtiger ist die folgende Feststellung: Die Berechenbarkeitsthese impliziert keinerlei Form von Determinismus. Wir haben bereits oben gesehen, daß Berechenbarkeit und Determinismus logisch unabhängige Eigenschaften von dynamischen Prozessen sind.
- (3) Berechenbarkeit steht daher auch nicht im Widerspruch zur Willensfreiheit. Vielmehr muß man Berechenbarkeit (in dem beschriebenen Sinn) als eine Bedingung oder Restriktion betrachten, die ein vernünftiger Begriff von Willensfreiheit erfüllen muß. Die zentrale philosophische Konsequenz ist daher:
- (4) Jeder wissenschaftlich haltbare Begriff von Willensfreiheit muß mit der Berechenbarkeit menschlichen Verhaltens (im Sinne von (\*)) verträglich sein.

### *Literatur*

Die umfangreiche Literatur zum Freiheitsbegriff zu bibliografieren möchte ich mir an dieser Stelle ersparen. Das folgende Verzeichnis enthält nur die wenigen im Text erwähnten Werke:

- H. Bauer: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie (Berlin, New York 1978).
- R. Carnap: *Philosophical Foundations of Physics* (New York 1966), deutsche Übersetzung: Einführung in die Philosophie der Naturwissenschaft (München 1976) besonders: Kap. 22: „Determinismus und freier Wille“.
- I. Kant: *Gesammelte Schriften*, hrsg. von der Preußischen Akademie der Wissenschaften (Berlin 1902 ff.).
- G. E. Moore: *Ethics* (Oxford 1966), deutsche Übersetzung: Grundprobleme der Ethik (München 1975).
- R. Penrose: *Shadows of the Mind* (Oxford 1994).
- P. v. Inwagen: *An Essay on Free Will* (Oxford 1983).

<sup>27</sup> Die These wäre also auch mit einer klassisch-dualistischen Auffassung von Körper und Geist verträglich. Nach der klassischen, cartesianischen Konzeption sind Geist und Körper verschiedene Entitäten (*res cogitans* – *res extensa*). Dementsprechend gibt es eine scharfe Trennung zwischen physikalischen und mentalen Systemzuständen und Verhaltensparametern. Aber selbst in diesem Fall würde gelten, daß alle stetigen und diskreten Verhaltensparameter (*sowohl* physikalische *wie* mentale) beliebig genau Turing-berechenbar sind.