

(Dies ist ein Teil des Buchs [Der Begriff der Wirklichkeit.pdf](#))

Inhalt:

2. Geist und Materie: Vorbereitende Bemerkungen.....	2
2.1. Einleitung: Ein fundamentaler Mangel.....	2
2.2. Übersicht über das weitere Vorgehen.....	8
Willensfreiheit und das Vollständigkeitsaxiom der Naturwissenschaft.....	9
Das Problem der Qualia.....	14
3. Willensfreiheit.....	17
3.1. Vorübung: Die Rechtfertigung psychologischer Begriffe.....	17
3.2. Der Weg zur geistigen Wirklichkeit über sieben Stationen.....	20
3.3. Der letzte Schritt: die Begründung der Willensfreiheit.....	33
3.4. Geordnete Zustände in neuronalen Netzen.....	37
3.5. Zusammenfassung.....	44
Bemerkungen.....	47
Schluss.....	49

2. Geist und Materie: Vorbereitende Bemerkungen

2.1. Einleitung: Ein fundamentaler Mangel

Können wir die Welt begreifen?

"Keinesfalls!" entgegnet die gegenwärtige Physik und liefert scheinbar unwiderlegbare Beweise: etwa das Doppelspaltexperiment, das angeblich auf keine denkmögliche Weise beschreibbar ist, oder die Bellsche Ungleichung, die jede lokale Interpretation der Welt ausschließen soll, oder die relativistischen Raum-Zeit-Verhältnisse, von denen behauptet wird, dass sie unserer *a priori* vorgegebenen Vorstellung von Raum und Zeit widersprechen.

Wenn die üblichen Interpretationen dieser Szenarien tatsächlich die einzig möglichen wären, dann würde jeder Versuch, herauszufinden, *was* die Wirklichkeit ist und *woraus* sie besteht, augenblicklich scheitern, ja er wäre geradezu unsinnig, weil unsere Begriffsbildungen dann zum Verständnis der Wirklichkeit offenbar vollständig ungeeignet wären.

Glücklicherweise hat sich herausgestellt, dass dieser unerfreuliche Befund falsch ist. Im ersten Teil ist gezeigt worden, dass sich nicht nur die eben erwähnten, sondern auch viele andere paradigmatische physikalische Szenarien auf einsichtige Weise begrifflich interpretieren und verstehen lassen, und dass sich die einzelnen Interpretationen zu einem alternativen Bild der Wirklichkeit zusammenfügen, aus dem alle Absurditäten verschwunden sind.

Aber selbst wenn nun vorausgesetzt werden könnte, wir verstünden die Natur und die uns bekannten Naturgesetze seien wahr, so wäre doch der dadurch gewonnene Begriff der Wirklichkeit noch immer unvollständig und zutiefst unbefriedigend, weil in ihm jener Bereich der Wirklichkeit, den wir *Geist* nennen, nicht enthalten ist; – und mehr noch: weil es, trotz gegenteiliger Behauptungen von Seiten mancher Hirnforscher, innerhalb des gegenwärtig geltenden Verständnisses der Natur sogar *vollkommen ausgeschlossen* ist, geistige und materielle Wirklichkeit in *einem* Bild zu vereinen.

In der Vergangenheit war diese Unvereinbarkeit rein philosophischer Art. Für Kant war es noch möglich, den Widerspruch zwischen der Gewissheit, dass wir Teil der vollständig naturgesetzlich bestimmten Wirklichkeit sind, und der Überzeugung, dass wir einen freien Willen besitzen, als Folge der Undenkbarkeit des *Dings an sich* aufzufassen.

Diese Möglichkeit besteht für uns nicht mehr. Daran hindert uns die Kenntnis des Zusammenhangs zwischen neuronalen Prozessen und geistigen Phänomenen. An dieser neuronal-geistigen Schnittstelle wird der Zugriff der materiellen Natur und ihrer Gesetzmäßigkeit auf den mentalen Bereich so konkret, dass es nicht mehr genügt, den naturwissenschaftlichen Anspruch – oder sagen wir besser: die naturwissenschaftliche Erklärungsmacht – mit allgemein-philosophischen Einwänden zurückzuweisen. Vielmehr sehen wir uns mit der Herausforderung konfrontiert, zu analysieren, *inwiefern* die bisherige naturwissenschaftliche Sicht zu kurz greift und diese Analyse zur Grundlage der Erklärung der Beziehung zwischen Geist und Materie zu machen.

Als die Naturwissenschaft gerade erst ihr Newtonsches Fundament erhalten hatte, erschien es noch kaum denkbar, dass durch Interaktion materieller Objekte Geist entstehen könnte. Hören wir, was Gottfried Wilhelm Leibniz, der sich wegen ebendieser Schwierigkeit gegen die atomistische Lehre seines Zeitgenossen Newton wandte, in seiner 1714 verfassten *Monadologie* dazu zu sagen hatte:¹

§. 17. Man ist außerdem genötiget zu bekennen daß die perception und dasjenige / was von ihr dependieret / auf mechanische Weise / das ist / durch die Figuren und durch die Bewegungen / nicht könne erkläret werden. Und erdichteten Falls / daß eine Machine wäre / aus deren Structur gewisse Gedanken / Empfindungen / Perceptionen erwachsen; so wird man dieselbe denkende Machine sich concipieren können / als wenn sie ins große nach einerlei darinnen beobachteter Proportion gebracht worden sei / dergestalt daß man in dieselbe / wie in eine Mühle / zugehen vermögend wäre. Wenn man nun dieses setzet / so wird man bei ihrer innerlichen Besichtigung nichts als gewisse Stücke / deren eines an das andere stosset / niemals aber etwas antreffen / woraus man eine Perception oder Empfindung erklären könnte. Daher muß man die Perception in der einfachen Substanz / und keines weges in dem Composito oder in der Machine suchen. [...]

In diesem Paragraphen, der später unter der Bezeichnung "Mühlengleichnis" bekannt wurde, schließt Leibniz also aus, dass Gedanken, Empfindungen oder Wahrnehmungen durch das "Stoßen von Stücken" – was wir heute wohl als "Wechselwirkung von Teilchen" bezeichnen würden – erzeugt werden könnten. Wenn diese Fähigkeiten aber nicht aus dem Zusammenwirken von Objekten hervorgehen, dann, so folgert er, müssen sie den *einfachen Substanzen* zuerkannt werden (die er *Monaden* nennt).

Wie stellt sich das Verhältnis Geist – Materie in der Gegenwart dar? Lässt sich die Kantische Antinomie von Freiheit und Kausalität durch das Wissen über neuronale Netze auflösen? Ist das

¹ In einer zeitgenössischen Übersetzung des französischen Originals von Heinrich Köhler. (Nachzulesen unter <http://gutenberg.spiegel.de/buch/2790/1>)

Leibnizsche Mühlengleichnis noch gültig? Gibt es weitere Einwände gegen die Hypothese, dass Geist von neuronalen Netzen hervorgebracht wird?

Ich nehme hier folgenden Standpunkt ein:²

Zum Problem der Freiheit

Die Kantische Antinomie ist nach wie vor aufrecht. Sie kann nicht aufgelöst werden, ohne das naturwissenschaftliche Paradigma zu korrigieren, demzufolge alles, was sich ereignet, vollständig durch Gesetze bestimmt ist – wobei es gleichgültig ist, ob diese Gesetze deterministisch oder "objektiv" probabilistisch sind. Unter dieser Voraussetzung sind Naturwissenschaft und Freiheit unvereinbar; aus dem auf diese Weise lückenlos geknüpften universellen Netz der Kausalität gibt es kein Entrinnen.

Die Annahme der Freiheit des Geistes ist somit gleichbedeutend mit einer Änderung bzw. Erweiterung der derzeit vorherrschenden naturwissenschaftlichen Sicht der Welt.

In den Kapiteln 3. *Willensfreiheit* und 4. *Das veränderte Bild der Wirklichkeit* werde ich zeigen, worin die Täuschung besteht, der die Naturwissenschaft bisher unterlag; dadurch wird diese Täuschung zugleich aufgehoben. Die Möglichkeit der Willensfreiheit ist eine Folge dieser Korrektur.

Zur Frage der naturwissenschaftlichen Beschreibbarkeit mentaler Prozesse; das Problem der Qualia

Mentale Prozesse haben einen *Informationsgehalt* und einen *Empfindungsgehalt*.

Die gegenwärtige naturwissenschaftliche Sicht des Zusammenhangs zwischen neuronalem Netz und Geist erzwingt nicht nur die Preisgabe der Freiheit, sie scheitert überdies an der Erklärung, warum mentale Zustände *Qualia* sind, d.h. warum sie eine Empfindungsqualität besitzen.

Durch die Korrektur der Sicht der Wirklichkeit, die in den Kapiteln 3 und 4 vorgenommen wird, können geistige Leistungen – Wahrnehmungen, Gedanken, Bewusstsein usw. – und auch die Freiheit des Geistes auf naturwissenschaftliche Weise verstanden werden; Aber auch in dieser erweiterten Sichtweise ist der Empfindungsgehalt mentaler Zustände noch *nicht* mit eingeschlossen.

² Ich werde zunächst meine Sicht der Probleme nur kurz skizzieren, die Lösungen bloß ankündigen und erst später alle Aspekte eingehend erörtern.

Das Problem der Qualia ist noch fundamentaler als das der Freiheit, weil es hier nicht nur um die Gültigkeit der *naturwissenschaftlichen* Beschreibung der Wirklichkeit geht, sondern darum, ob und wie weit die Möglichkeit einer Beschreibung der Wirklichkeit überhaupt gegeben ist: Die Definition des Begriffs "rot" kann den mentalen Zustand *rot* – d.h. die Empfindung *rot* – ebenso wenig vermitteln wie die Videoaufnahme oder die Gleichung des dynamischen neuronalen Musters "rot", das diese Empfindung hervorbringt, und dasselbe gilt für jede Art von Beschreibung oder Abbildung. Die Empfindung *rot* ist in keiner Beschreibung enthalten. Sie stellt also ein *irreduzibles* Seiendes dar, das nicht *beschrieben*, sondern nur *erlebt* werden kann.

Solange nicht geklärt ist, warum die Empfindung *rot* – oder irgendeine andere Empfindung – dennoch als natürliches Seiendes verstanden werden kann, wissen wir nicht, ob geistige Phänomene nicht unserer Vorstellung von der Wirklichkeit widersprechen, und wir können dann nicht einmal sicher sein, ob wir das, was die Wirklichkeit *eigentlich* ist, bei unseren naturalistischen Modellannahmen nicht vollständig verfehlen.

Umgekehrt entbehrt aber der Versuch, den Geist selbst zur Basis der Wirklichkeit zu machen, jeder Grundlage. Alles, was wir über die Natur wissen, verdanken wir der Naturwissenschaft, die der Untersuchung materieller Strukturen und Prozesse gewidmet ist. Die Annahme, nicht Materie, sondern Geist sei das grundlegende Prinzip, hat dagegen außer Spekulationen bisher überhaupt nichts hervorgebracht. Vielmehr sprechen alle Erfahrungen eindeutig dafür, dass Materie – oder wie auch immer man dasjenige Existierende nennen mag, das von der Naturwissenschaft beschrieben werden kann – eine *notwendige Bedingung* für das Auftreten geistiger Phänomene ist.

Im Kapitel 5. *Qualia* wird geklärt, was Beschreibungen leisten können bzw. nicht können und warum das so ist. Der Unterschied zwischen materiellen und geistigen Entitäten wird analysiert. Die Ergebnisse dieser Analyse ermöglichen es, Qualia als Zustände der sich entfaltenden Natur zu verstehen.

Gibt es weitere begründete Einwände gegen die Hypothese, dass Geist von neuronalen Netzen hervorgebracht wird?

Nein. Unser Wissen über Informationsverarbeitung und Repräsentation in neuronalen Netzen ist so weit fortgeschritten, dass für grundsätzliche Zweifel an dieser Hypothese kein Raum mehr bleibt.

(Allerdings muss aufgeklärt werden, in welcher Beziehung Wirklichkeit und Repräsentation stehen. Auch das wird im Kapitel 5. *Qualia* geschehen.)

Falls es also gelingt, vom naturwissenschaftlichen Weltbild ausgehend durch logische Argumentation die Freiheit des Willens und die Existenz der Qualia zu begründen, dann sind alle scheinbar unüberwindlichen Probleme beseitigt, an denen die Integration von Geist und Materie bis jetzt gescheitert ist. Dann können – ohne Zuhilfenahme exotischer Zusatzannahmen – beide Phänomenbereiche in einem Bild zugleich gedacht werden, ohne dass sie durch diese Vereinigung etwas von ihrem Wesen verlieren.

Für den Augenblick genügt dieser kurze Abriss der Problematik des Zusammenhangs zwischen mentalen Zuständen und neuronalen Prozessen.

Das Schlusswort dieser Einleitung überlasse ich Erwin Schrödinger, der wie kaum ein Anderer die beiden Hauptprobleme der naturwissenschaftlichen Weltansicht erkannt hat: die Uninterpretierbarkeit physikalischer – vor allem quantentheoretischer – Sachverhalte und das Scheitern des Versuchs, Geist in das naturwissenschaftliche Weltbild zu integrieren.

Schrödinger schreibt:³

"Wir müssen bekennen, [...] dass es uns jedenfalls bisher nicht gelungen ist, ein einigermaßen verständliches Weltbild aufzubauen, ohne unseren eignen Geist, den Schöpfer des ganzen Weltbildes, daraus zu verbannen, derart, dass darin für ihn kein Platz ist. Der Versuch, ihn hineinzuzwängen, führt notwendig auf Ungereimtheiten.

[,,] im physikalischen Weltbild fehlen alle Sinnesqualitäten, aus denen das Subjekt der Erkenntnis sich eigentlich zusammensetzt. Dem Modell fehlen Farben, Töne, Greifbarkeit. Ebenso und aus dem gleichen Grund mangelt der Welt der Naturwissenschaft alles, was eine Bedeutung in bezug auf das bewusst anschauende, wahrnehmende und fühlende Wesen hat; Von alledem enthält sie nichts. Vor allem denke ich an die sittlichen und ästhetischen Werte, Werte von jeder Art, an alles was auf Sinn und Zweck des ganzen Geschehens Bezug hat. Nicht nur fehlt dieses alles, sondern es kann von einem rein naturwissenschaftlichen Standpunkt aus überhaupt nicht organisch eingebaut werden. Wenn man es einzubauen versucht, wie ein Kind seine schwarzweiße Malvorlage koloriert, so passt es nicht hinein. Denn alles, was man in dieses Weltmodell eingehen lässt, nimmt stets die Form einer naturwissenschaftlichen Aussage an, ob man will oder nicht; als solche aber wird es falsch."

Soweit Schrödingers Darstellung des grundsätzlichen Mangels der Naturwissenschaft bei der Beschreibung geistiger Phänomene. Ich stimme ihm zu. Nun meint Schrödinger allerdings, aus der Tatsache, dass Bewusstsein nicht in das derzeit gültige naturwissenschaftliche Weltbild integriert

³ Erwin Schrödinger: *Geist und Materie*, Zsolnay 1986, S.95f.

werden kann, folgern zu müssen, dass Geist nicht von Materie erzeugt wird und dass er auf eine Weise existiert, über die uns die Naturwissenschaft keine Auskunft gibt: nicht als individueller Geist, wie wir glauben, sondern als *ein* universaler Geist, wie es der fernöstlichen Tradition entspricht.⁴

Das ist auch der Grund, weshalb ich Schrödinger hier so ausführlich zu Wort kommen lasse: nicht etwa deshalb, weil ich diese Auffassung teile (was ich sicher nicht tue), sondern deshalb, weil er wegen des ihm unvermeidlich erscheinenden Scheiterns der Erklärung des Geistes aus seinen materiellen Voraussetzungen den Bereich des vernünftigen Denkens und Urteilens verlassen hat – so wie viele andere große Denker vor und nach ihm –, und weil seine und deren gedankliche Verirrungen Zeugnis davon ablegen, welche Folgen der fortdauernde Mangel an Einsicht in den Zusammenhang von Geist und Materie hat. Solange diese Einsicht fehlt, werden immer wieder gerade diejenigen, die das Problem am deutlichsten erkennen, der Versuchung ausgesetzt sein, ihr Heil in haltlosen Spekulationen zu suchen, nicht anders, als es nun seit mehr als hundert Jahren in jenem erkenntnistheoretischen Freiraum – nein: in jenem Bereich von Narrenfreiheit geschieht, der durch das Versagen der Interpretation des quantentheoretischen Formalismus eröffnet worden ist.

Es gilt ja ganz allgemein, dass jede Unvollständigkeit der wissenschaftlichen Welterklärung eine Ausstiegsstelle aus dem Bereich vernünftigen Denkens bildet. Die beliebtesten Ausstiegsstellen sind aber jedenfalls diese beiden: Quantentheorie und Geist. Immer mehr Naturwissenschaftler, Esoteriker und Anhänger aller Religionen, aber auch "modern denkende" Durchschnittsbürger drängen sich dort zusammen, um vor der vergeblichen Anstrengung des Begriffs ins Nirwana des Unsinnns zu entfliehen und dort ihren wirren Reigen zu tanzen.

Anmerkung

Zwischen der Frage: *Gibt es einen freien Willen* und der Frage: *Wie können Qualia erklärt werden* besteht ein wichtiger Unterschied:

Im Fall der Qualia handelt es sich um einen echten, im Rahmen der bisherigen naturwissenschaftlichen Sicht nicht zu beseitigenden Defekt unseres Verständnisses der Wirklichkeit. Aus der Tatsache, dass Qualia existieren und jedes Quale ein irreduzibles Seiendes ist, folgt zwingend, dass die derzeitige naturwissenschaftliche Sicht der Wirklichkeit falsch oder unvollständig ist.

Anders verhält es sich im Fall der Willensfreiheit: Der freie Wille muss nicht existieren. Die Annahme, dass er nicht existiert, lässt sich widerspruchsfrei durchhalten. Dass durch die gegenwärtige Natur-

⁴ A.a.O., S.79.

wissenschaft die Existenz der Willensfreiheit ausgeschlossen wird, ist also kein *logischer* Beweis für einen Defekt des derzeit gültigen naturwissenschaftlichen Weltbilds.

Es bleibt allerdings ein massives Unbehagen: die Eliminierung der Willensfreiheit würde bedeuten, dass wir nicht mehr Urheber unserer Handlungen sind. Unser freier Wille wäre dann eine Täuschung – tatsächlich würden wir nur Gesetze vollziehen, die wir zwar nicht vollständig kennen, durch die wir aber genauso determiniert wären wie z.B. Planeten durch das Gravitationsgesetz.

Damit verschwände zugleich der Begriff der *Verantwortung*. Ein Wesen, dessen Verhalten – nicht anders als die Bewegung eines Himmelskörpers – bloß Ausdruck des Gesetzes ist, durch das es bestimmt ist, ist für seine Taten nicht verantwortlich.⁵

2.2. Übersicht über das weitere Vorgehen

Im Folgenden gehe ich von der Voraussetzung aus, dass *Geist* ein *natürliches Phänomen* ist, das durch neuronale Netze von hinreichender Komplexität hervorgebracht wird. Ich werde zeigen, dass er trotz der Einbindung in die Naturkausalität frei ist und dass *Qualia* selbstverständliche Elemente einer folgerichtigen Betrachtung der Wirklichkeit sind.

Die Erklärung ist also zweistufig:

1. Die Begründung der *Freiheit* des Geistes, ohne die wir – wie soeben ausgeführt – nicht Urheber unserer Handlungen wären.
2. Die Begründung der Tatsache, dass geistige Zustände immer *Qualia* sind, also einen subjektiven, nur demjenigen, der sie erlebt, zugänglichen Empfindungsgehalt haben.

⁵ Von Seiten der Hirnforschung wird gelegentlich argumentiert, dass Konzepte wie *Verantwortung* oder *Schuld* trotzdem wegen ihrer Bedeutung im sozialen Kontext beibehalten werden müssten. Hier zeigt sich jedoch deutlich die Seltsamkeit von Theorien, die den Geist als naturgesetzlich determiniert betrachten: Wenn *Verantwortung* und *Schuld* keine geistigen Phänomene sind, sondern neuronale, dann sind auch *Argument* und *Wille* neuronale Phänomene. Wenn die kausalen Zusammenhänge in der neuronalen Ebene liegen und nicht in der geistigen, dann kann ein Grund kein Grund sein und ein Wollen kein Wollen. Niemand hat je eine Wahl – es geschieht immer, was aufgrund physikalischer Ursachen geschehen muss. Etwas zu wollen oder für etwas zu argumentieren ist daher unsinnig. Deterministen, die glauben zu argumentieren unterliegen einer Selbsttäuschung.

Diese Einteilung ergibt sich zwingend, weil einerseits die Argumentation für die Existenz der Qualia auf der für die Existenz der Freiheit aufbaut, andererseits aber die beiden Argumentationsszenarien doch so weit selbständig sind, dass jedes für sich allein betrachtet werden muss.

Das Ziel dieses zweistufigen gedanklichen Wegs ist nichts Geringeres als eine neue Sicht der Wirklichkeit. Deshalb will ich es nicht bei dieser kargen Ankündigung bewenden lassen, sondern mit einigen Bemerkungen wenigstens eine Übersicht über die Reihenfolge der argumentativen Schritte vermitteln und die zugehörigen Voraussetzungen ein wenig ausführlicher skizzieren.

Willensfreiheit und das Vollständigkeitsaxiom der Naturwissenschaft

Die derzeitige wissenschaftliche Sicht der Natur lässt sich durch folgende Annahme – nennen wir sie A_{N_0} – charakterisieren:

A_{N_0} : *Alles, was sich ereignet, folgt aus universell gültigen Naturgesetzen und Anfangsbedingungen.*⁶

Was sich in einem Ausschnitt der Wirklichkeit – in einem "System" – ereignet, ist jedoch in vielen Fällen nicht nur durch universelle Naturgesetze festgelegt, sondern auch durch weitere Gesetze, die nur in bestimmten Systemen gelten.⁷

Das lässt sich leicht am Beispiel neuronaler Netze demonstrieren:

Die Beschreibung der Dynamik eines neuronalen Netzes bezieht sich zunächst und vor allem auf die *Struktur* des Netzes und nicht auf die Details physikalischer oder chemischer Vorgänge. Jeder Hirnforscher wird das neuronale Netz auch als *kybernetisches* und nicht nur als physikalisch-chemisches System auffassen. Die Betrachtungsweise des Netzes als kybernetisches System erlaubt es sogar, von der Art der physikalischen Realisierung völlig abzusehen.

Das bedeutet, dass hier – zusätzlich zu den Naturgesetzen – noch eine andere Art von Gesetzen den Verlauf der neuronalen Dynamik bestimmt, die man *Strukturgesetze* oder *Gesetze der Form* nennen könnte.

⁶ Quantenmechanisch: *Die Wahrscheinlichkeit jedes Ereignisses folgt aus ...*

⁷ Dies ist ein Sachverhalt von grundsätzlicher Bedeutung. In den folgenden beiden Kapiteln (3. *Willensfreiheit* und 4. *Das veränderte Bild der Wirklichkeit*) werde ich ausführlich darauf eingehen.

Dasselbe gilt auch für die einzelnen Neurone: Neurone sind Systeme, bei denen zwischen Input und Output eine gesetzmäßige Beziehung besteht. Dieses *neuronale Input-Output-Gesetz* hängt mit der Form und Struktur des Neurons zusammen. Auf diese Weise stehen Neurone also in Wechselwirkung, nicht anders als z.B. elektrisch geladene Teilchen in elektromagnetischer Wechselwirkung stehen.

Das neuronale Input-Output-Gesetz kann daher als *Wechselwirkungsgesetz der Neurone* aufgefasst werden. Es ist ein Gesetz, das nur für Neurone gilt und zusätzlich zu den Naturgesetzen auftritt. Die mathematische Formulierung des Gesetzes ist *von seiner physikalischen Realisierung unabhängig*

Um die Tatsache der Existenz solcher spezifischen, nur für bestimmte Systeme geltenden Gesetze zu berücksichtigen, muss die Annahme A_{N_0} zur Annahme A_N erweitert werden:

A_N : *Alles, was sich ereignet, folgt aus Gesetzen und Anfangsbedingungen.*

Für die Beschreibung von Systemen werden die Gesetze in die Form spezifischer Gleichungen gebracht. Gleichungen sind quantitative Beziehungen zwischen Variablen, Anfangsbedingungen sind die Werte dieser Variablen zu einem bestimmten Zeitpunkt.

Also bedeutet A_N , dass es zu jedem Bereich der Wirklichkeit ein System von Gleichungen gibt, das *alle* Informationen enthält, die für eine vollständige Beschreibung dieses Bereichs erforderlich sind. In diesem Sinn kann also jeder Bereich der Wirklichkeit auf ein solches System abgebildet werden.

Ich werde A_N deshalb das ***Vollständigkeitsaxiom der Naturwissenschaft*** nennen.

A_N beinhaltet zwei Forderungen:

Die erste Forderung betrifft die Gesetze. Der Ausdruck "folgt aus" setzt voraus, dass die den Gesetzen entsprechenden Gleichungen einen *Algorithmus* bereitstellen, d.h. ein Verfahren, das es ermöglicht, aus den gegenwärtigen Bedingungen durch Anwendung definierter Regeln künftige Ereignisse abzuleiten bzw. vorauszuberechnen.

Die zweite Forderung betrifft die Anfangsbedingungen. Wenn das Gleichungssystem *alle* Informationen über die Wirklichkeit enthalten soll, dann muss es *unendlich genau* sein. Selbstverständlich sind aber Anfangsbedingungen niemals "unendlich genau" verfügbar: weder ist es möglich, mit unendlicher Genauigkeit zu messen, noch kann irgendein System ganz von der Umgebung getrennt werden. (Also müsste in jedem Fall das ganze Universum mit einbezogen werden.)

Außerdem wäre allein schon für die unendlich genaue numerische Darstellung des Wertes einer einzigen Variablen in den meisten Fällen unendlich lange Zeit erforderlich.⁸

Also ist A_N keine Aussage über ein wirklich existierendes Gleichungssystem, sondern über ein rein gedankliches Gebilde.

Ist A_N dann überhaupt noch eine sinnvolle Aussage? Ich denke ja. A_N gibt die Überzeugung wieder, dass *die Natur selbst ein algorithmisches System* ist, dass sie also aus den in einem bestimmten Augenblick gegebenen Bedingungen nach feststehenden Regeln die Zukunft produziert, und dass dieser regelhafte Vorgang im Prinzip durch ein Beschreibungssystem reproduziert werden kann – auch wenn diese Reproduktion durch die obigen Einschränkungen nur näherungsweise durchführbar ist. A_N für sinnvoll zu halten, bedarf bloß zweier Annahmen, die eigentlich selbstverständlich erscheinen: die erste Annahme ist, dass die Variablen zu jedem Zeitpunkt exakte Werte *haben*,⁹ und die zweite Annahme ist, dass *die Natur* den die Zukunft erzeugenden Algorithmus mit ebendiesen Werten ausführt. Beide Annahmen werden durch die Tatsache, dass *wir* in endlicher Zeit diese Werte weder numerisch niederschreiben noch den Algorithmus damit durchführen könnten, nicht in Frage gestellt.

Kehren wir zu unserem Gegenstand zurück.

Aus naturwissenschaftlicher Sicht müssen sich also neuronale Netze (im Prinzip) auf ein Gleichungssystem abbilden lassen, derart, dass die zeitliche Entwicklung des neuronalen Netzes der zeitlichen Entwicklung des Gleichungssystems entspricht, d.h. dass zu jedem Zeitpunkt die tatsächlichen Werte aller Variablen gleich den (im Prinzip) errechenbaren Werten wären. Da dieses Gleichungssystem niemals wirklich verfügbar ist, muss man sich in der Praxis mit Näherungen zufrieden geben, also z.B. bei Modellierungen des Gesamtsystems nur die Interaktionen großer Ensembles von Neuronen betrachten, oder den Blick nur auf eine beschränkte Zahl von Neuronen richten. Die Annahme A_N garantiert jedoch, dass diesen Näherungen nichts "Wesentliches" fehlen muss, eben deshalb, weil auch das neuronale Netz ein "Bereich der Wirklichkeit" ist, der durch Gesetze *vollständig* bestimmt wird

⁸ Man könnte auch fragen, ob überhaupt vorausgesetzt werden kann, dass sich die Natur selbst unendlich genau an ihr Gesetz hält. Diese Frage wäre aber nur dann sinnvoll, wenn die Natur ihrem Gesetz *gehören* würde; Gemäß den Ausführungen des ersten Kapitels gehorcht sie aber nicht dem Gesetz, sondern sie *ist* das Gesetz.

⁹ Quantenmechanisch heißt das (wie immer), dass es zu jedem Zeitpunkt eine exakte Wahrscheinlichkeitsverteilung der möglichen Messwerte gibt. In meiner Interpretation, die verborgene Parameter enthält und vollständig deterministisch ist, ist diese Ergänzung natürlich überflüssig. Ich füge sie nur hinzu, um die Argumentation zur Willensfreiheit von meinen eigenen physikalischen Hypothesen so weit wie möglich unabhängig zu halten.

und somit auf ein Gleichungssystem abgebildet werden kann – und "vollständig" bedeutet, dass die Abbildung *alles* kann, was das Netz selbst kann.

Wären tatsächlich alle Zustände eines Netzes auf diese Weise erfassbar, dann würde das zugehörige Differenzialgleichungssystem also auch eine Beschreibung der geistigen Tätigkeit enthalten, die durch das Netz hervorgebracht wird. Von Freiheit des Willens könnte unter dieser Voraussetzung keine Rede sein.

Das bedeutet umgekehrt: Aus der Annahme, dass wir uns frei nach unserem Willen entscheiden können, ergibt sich zwingend die Forderung, dass die mathematische Beschreibung neuronaler Netze *unvollständig* ist – und es ist hinzuzufügen, dass es sich hier um eine Unvollständigkeit handeln muss, die über die oben erwähnte Einschränkung infolge der Unmöglichkeit unendlicher Genauigkeit hinausgeht, um eine Unvollständigkeit also, die darauf beruht, dass die Natur *kein algorithmisches System* ist.

Es ist außerordentlich wichtig, zu sehen, dass die Existenz der Willensfreiheit nicht davon abhängt, was *wir* über ein menschliches neuronales Netz wirklich wissen bzw. wissen können. Die Willensfreiheit verschwindet nicht erst dann, wenn eine gesetzliche Beschreibung des Netzes wirklich zur Verfügung steht, sondern bereits dann, wenn die Annahme getroffen wird, dass die *Existenz* einer vollständigen mathematischen Abbildung (der zuvor definierten Art) möglich ist, oder anders gesagt: wenn angenommen wird, dass die Zukunft aus der Gegenwart auf algorithmische Weise hergestellt wird.

Konkret: Für die Annahme der menschlichen Willensfreiheit ist es nicht hinreichend, dass niemand genau weiß, was ein Mensch tun wird, ja es würde nicht einmal genügen, dass dieses Wissen aus *technischen* Gründen niemals erlangt werden könnte; – Dafür muss die wesentlich stärkere Bedingung erfüllt sein, dass ein solches Wissen *aus logischen Gründen unmöglich* ist.

Vielleicht kann das folgende, ein wenig skurrile Gedankenexperiment zur Klärung des Sachverhalts beitragen:

Es steht ohne Zweifel fest, dass die Bahn des Planeten Merkur durch die Gravitation der Körper im Sonnensystem bestimmt ist. Nehmen wir nun an, der Merkur besäße Geist wie wir und wäre überzeugt, einen freien Willen zu besitzen und seine Bahn aus freien Stücken und mit guten Gründen zu wählen; Dann wäre das offenbar eine Selbsttäuschung des Planeten, die darauf beruht, dass er nicht weiß, dass seine Bahn gesetzlich festgelegt ist. Dessen sind wir uns gewiss, obwohl die *exakte* Gleichung für seine Bahn niemals verfügbar wäre – sie müsste ja die Orte und Geschwindigkeiten *aller* Körper des Sonnensystems berücksichtigen – und obwohl diese Gleichung, falls wir sie doch

hätten, keinesfalls lösbar wäre.¹⁰ Wir sind uns dessen gewiss, weil wir einfach *wissen*, dass die Planetenbahnen *ausschließlich* durch das Gravitationsfeld festgelegt sind. Diese Gewissheit ist selbst nicht naturwissenschaftlich im engeren Sinn, sie ist aber natürlich eine Voraussetzung der Naturwissenschaft und ein unbezweifelbares Element der naturwissenschaftlichen Weltanschauung.

Wäre unsere eigene geistige Aktivität in der gleichen Weise durch Gesetze bestimmt wie die Bahn des Merkurs durch das Gravitationsgesetz, dann wären wir offenbar in genau derselben Lage wie der sich frei wählende Planet. Wir würden bloß glauben, willentlich zu handeln, tatsächlich aber würden wir doch stets nur das vollziehen, was uns per Gesetz vorgeschrieben ist. Wir würden das Gesetz unseres Handelns zwar nicht kennen, aber es wäre stets präsent und würde uns führen. Und unser Geist ließe sich dann – wenn auch nur *im Prinzip* und niemals realiter – auf ein Gleichungssystem abbilden, wäre darin gefangen und hätte seine Freiheit verloren.

Unter dieser Voraussetzung dennoch anzunehmen, wir handelten nach unserem Willen, wäre ebenso unsinnig wie die Behauptung, nicht nur Gravitation würde die Bahn der Himmelskörper lenken, sondern auch ein *vielarmiger kosmischer Affe*.

Wir sind also zu dem Schluss gelangt, dass höchstens *eine* der beiden Behauptungen wahr ist:

1. Willensfreiheit existiert. 2. Das Vollständigkeitsaxiom A_N ist korrekt.

Das ist der gegenwärtige Stand der Dinge: Eine unerfreuliche Alternative. Einerseits ist die Annahme der Willensfreiheit für unser Selbstverständnis und für den Sinn unseres Daseins von essenzieller Bedeutung, andererseits ist aber völlig unklar, auf welche Weise die prinzipielle Möglichkeit der Abbildung der Wirklichkeit auf ein System von Gleichungen und Anfangsbedingungen ausgeschlossen werden könnte.

In den Kapiteln 3. *Willensfreiheit* und 4. *Das veränderte Bild der Wirklichkeit* steht genau diese Frage im Mittelpunkt:

Warum ist das Vollständigkeitsaxiom der Naturwissenschaft ungültig?

Es wird gezeigt, dass eine Darstellung der Dynamik neuronaler Netze durch ein System von Gleichungen und Anfangsbedingungen nicht möglich ist, und zwar, wie gefordert, nicht bloß aus technischen, sondern aus logischen Gründen.

¹⁰ Mit "lösbar" meine ich hier und im Folgenden, dass es ein Verfahren gibt, durch das zukünftige Variablenwerte exakt berechnet werden können.

Der Beweis wird durch die Ableitung des folgenden Satzes erbracht:

Es gibt Zustände neuronaler Netze, deren formale Repräsentationen Aussagen sind, die nicht aus irgendeinem System von Gleichungen und Anfangsbedingungen ableitbar sind und die nicht als Lösung eines solchen Gleichungssystems für irgendeinen bestimmten Zeitpunkt aufgefasst werden können.

Das Vollständigkeitsaxiom der Naturwissenschaft ist daher falsch. Jede formale Beschreibung der Natur durch ein System von Gleichungen und Anfangsbedingungen ist unvollständig.

Das naturwissenschaftliche Paradigma muss korrigiert werden.

Die Argumentation, die zum Beweis dieser Behauptung erforderlich ist, genügt für die Begründung der Eigenständigkeit des Geistes und der Existenz der Willensfreiheit. Allerdings stellt diese Begründung nur eine notwendige und keine hinreichende Voraussetzung für das Verständnis von Geist dar, denn es fehlt noch immer die Erklärung dessen, was geistige Zustände *eigentlich* sind, nämlich *Qualia*.

Das Problem der Qualia

Nichts wäre einfacher, als einen Roboter zu konstruieren, der die Frequenzen und Frequenzmischungen des Lichts, die wir als *rot* empfinden, mit dem Wort "rot" verknüpfen könnte. Man könnte ihn so programmieren, dass er "rot" sagt, wenn sein optischer Chip ihm die Information liefert, dass eine solche Frequenzmischung vorliegt, während er gleichzeitig seinen Arm auf den betreffenden Gegenstand richtet und sich darauf zu bewegt. Für diese Leistung wäre noch nicht einmal künstliche Intelligenz erforderlich.

Es wäre dann vollkommen klar, dass der Roboter – ich will ihn *Hans* nennen, um später auf ihn zurückkommen zu können – zwar die **Information** hat, die für die Definition von "rot" erforderlich ist, dass ihm aber die **Empfindung** *rot* fehlt. *Rot* als Quale – eben genau jene charakteristische Farbe-Empfindung, die *wir* haben, wenn wir *rot* wahrnehmen oder uns vorstellen – kann in seinem simplen Programm unmöglich auftreten.

In diesem Beispiel wird deutlich, dass zwischen *Information* und *Empfindung* unterschieden werden muss. Daraus ergibt sich aber eine Reihe von Fragen, von denen bis heute keine einzige auch nur ansatzweise gelöst ist.

Um nur einige aufzuzählen:

Wenn ein Mensch *rot* empfindet, dann enthält die Beschreibung seiner neuronalen Aktivität die Information, ohne die es diese Empfindung nicht gäbe. Wie ist hier die Beziehung zwischen Information und Empfindung? Wodurch wird die Empfindung ermöglicht?

Worin besteht der Unterschied zwischen einem Systemzustand, der nur Information enthält, und einem Systemzustand, der eine Empfindung hervorbringt?

Können künstliche neuronale Netze Empfindungen haben?

Offenbar hängt die Fähigkeit, etwas zu empfinden, mit der Komplexität der neuronalen Netze zusammen, die sich im Laufe der Evolution entwickelt haben. Damit erhebt sich die Frage:

Auf welcher Sprosse der evolutionären Leiter wird die Information zur Empfindung? Wo wird der Automat zum empfindenden Wesen? Bei den Wirbellosen? Bei den Reptilien? Bei den Fischen? Bei den Vögeln? Oder erst bei den Säugetieren?

Wodurch ereignet sich diese Wesensänderung?

Man kann auch konkreter fragen:

Nehmen Bienen Farben wahr? Ist ihre Welt wirklich *bunt*? Oder sind sie bloß Automaten, die auf Licht-Frequenzen reagieren? Empfinden Krebse Schmerz? Oder reagieren sie nur auf Reize?

"Komplexität" wird häufig als Zauberwort benutzt, so als wäre klar, dass bei hinreichender Komplexität eines neuronalen Netzes "von selbst" Empfindungen auftreten würden. Das ist natürlich unsinnig. Komplexität ist zwar eine notwendige Bedingung, besagt aber nichts weiter.

Dasselbe gilt für die populäre Außen-Innen-Unterscheidung. Es wird behauptet, Information und Empfindung seien *dasselbe*: genau das, was von außen – von einem Beobachter – als neuronale Aktivität gesehen wird, wird von innen – vom "System" selbst – als Empfindung erlebt. Aber auch diese Behauptung ist keine Erklärung, und überdies ist sie sogar falsch, wie sich später herausstellen wird.

Es bleibt also die zentrale Frage: *Was sind Qualia? Auf welche Weise können Qualia als natürliche Zustände verstanden werden?*

Die formal-logische Argumentation, die im Kapitel 3. *Willensfreiheit* der Begründung der Willensfreiheit dient, ist hier deshalb nicht ausreichend, weil die Erklärung der Qualia auf dem nur metaphysisch definierbaren Unterschied zwischen *wirklich existierenden Objekten* und deren *Repräsentationen (Beschreibungen)* basiert.

Von dieser Voraussetzung ausgehend gelangt man mittels einiger einfacher metaphysischer Schlussfolgerungen zu einem erweiterten Verständnis des Seienden, das sowohl die naturwissenschaftlich beschreibbare als auch die geistige Wirklichkeit umfasst. Die drei Welten – die materielle Welt, die Welt der geistigen Prozesse und die Welt der Entdeckungen und Schöpfungen des Geistes – können unter einen Begriff gebracht werden.

3. Willensfreiheit

Gegenstand der folgenden Erörterung ist die *metaphysische* Frage nach der Möglichkeit des freien Willens. Es geht also nicht um psychologische, soziologische oder andere Abhängigkeiten von Willensentscheidungen, sondern ausschließlich um die Frage, ob freier Wille *überhaupt möglich* ist – unter der Voraussetzung, dass wir zu einem Universum gehören, in dem nach gegenwärtiger Überzeugung *alles* gesetzmäßig abläuft.

3.1. Vorübung: Die Rechtfertigung psychologischer Begriffe

Zurzeit ist die Annahme verbreitet, dass Geist auf neuronale Prozesse zurückgeführt werden kann. Bezeichnen wir diese Position als "neuronalen Reduktionismus". Er bedeutet Folgendes:

Geistige Zustände werden von uns zwar als die Agenzien des geistigen Geschehens erlebt, aber wir unterliegen bei dieser Auffassung einer Täuschung – *eigentlich* ist das geistige Geschehen bloß ein reines Epiphänomen des neuronalen Geschehens, durch das es vollkommen determiniert wird. Man könnte auch sagen: Geistige Phänomene sind *eigentlich* neuronale Phänomene.

Unter dieser Voraussetzung mag es zunächst so scheinen, als wäre die Beschreibung geistiger Tätigkeit durch psychologische Begriffe – wie Gefühle, Gedanken, Gründe usw. – nur eine ungenaue und überholte Art der Darstellung, die nur solange angewendet werden sollte, bis eine exaktere Darstellung durch Begriffe und Konzepte der neuronalen Ebene zur Verfügung steht. Erst dann würden wir ja wirklich wissen, was z.B. der psychologische Begriff "Grund" *wirklich* bedeutet; – er könnte etwa ein lokales Energieminimum der Dynamik¹¹ eines Teilbereichs des neuronalen Netzes sein.

Untersuchen wir, in welchem Maß diese Vorstellung einer möglichen Eliminierung psychologischer Begriffe und Konzepte gerechtfertigt ist.

Notwendige Voraussetzung des neuronalen Reduktionismus ist die Annahme der Eigenständigkeit der neuronalen Beschreibung, von der die reduktionistische Argumentation ausgeht. Damit ist gemeint,

¹¹ Der Terminus "Dynamik" bezeichnet die zeitliche Entwicklung des Zustands eines Systems. "Zustand" bezeichnet die Menge der Werte aller Variablen (quantenmechanisch die Werte aller gleichzeitig messbaren Observablen) des Systems zu irgendeinem Zeitpunkt.

dass man von *Neuronen* und *neuronalen Prozessen* reden kann und nicht auf Moleküle, Atome oder Elementarteilchen und deren Wechselwirkungen zurückgreifen muss. Kann man diese Eigenständigkeit voraussetzen? – Das scheint in mancher Hinsicht so selbstverständlich, dass man meinen könnte, sich gar nicht darum kümmern zu müssen. Man *sieht* ja, dass Neurone existieren, dass sie miteinander agieren, dass durch ihre *Form* die Bewegung der Moleküle festgelegt ist, und deshalb erscheint eine Beschreibung des neuronalen Geschehens mittels der Terminologie und der Modelle der Neurowissenschaft und Neuroinformatik angemessen.

Tatsächlich kann aber die Möglichkeit, als Grundelemente der Beschreibung eines Bereichs der Wirklichkeit Strukturen zu benützen, die für sich schon komplexe Aggregate einfacherer Bestandteile sind, vom reduktionistischen Standpunkt aus überhaupt nicht als selbstverständlich vorausgesetzt werden. Der Reduktionist muss in jedem Fall begründen, warum die Elemente seiner Beschreibung (Neurone) selbst schon zusammengesetzt sein dürfen und warum er nicht von *wirklich elementaren* Bestandteilen ausgeht.

Als Grund – und es ist anzunehmen, dass auch ein Reduktionist so argumentieren würde – für die Berechtigung der Beschreibung auf der Basis von Neuronen kann angeführt werden, dass die Neurone *als Neurone* aufeinander bezogen sind, sodass Neuronenverbände als kybernetische Systeme aufgefasst werden können, die durch die funktionellen Zusammenhänge zwischen ihren Elementen vollständig definiert sind, weshalb es bei der Darstellung der Struktur und Dynamik der Zustände von Neuronenverbänden unwichtig ist, dass es sich *eigentlich* (– im reduktionistischen Sinn) um Moleküle (oder Atome, oder Elementarteilchen) handelt; Die Art der physikalischen Realisierung ist gleichgültig (– was auch eine notwendige Voraussetzung für die Möglichkeit von Computersimulationen neuronaler Prozesse ist).

Gehen wir nun zu einer höheren Stufe der Komplexität über: betrachten wir die Ebene der *geistigen Zustände und Prozesse*. "Geistige Zustände" definieren wir wie in der Hirnforschung üblich als globale raum-zeitliche Muster des neuronalen Netzes mit gleichzeitiger Aktivierung vieler Areale dieses modular aufgebauten Netzes, die verschiedene Aspekte eines Sachverhalts oder einer Situation repräsentieren und für die Dauer des Zustands miteinander verbunden sind. Ohne ins Detail gehen zu wollen, welche Areale das sind, lässt sich doch einiges beispielhaft aufzählen: zur Wahrnehmung gehörende neuronale Module, die die (tatsächlichen oder vorgestellten) szenischen Gegebenheiten repräsentieren, diverse Module im vorderen Teil des Gehirns, die der Analyse verschiedener komplexer Zusammenhänge dienen (wie etwa der Einschätzung sozialer Folgen einer Handlung), motorische Module, das Sprachzentrum usw. – in jedem Fall aber auch Teile des Zwischenhirns, die das neuronale System mit chemischen, verhaltensregulierenden Systemen verbinden und damit für die Intentionalität geistiger Zustände unerlässlich sind.

Nun folgt die für die Rechtfertigung psychologischer Begriffe entscheidende Schlussfolgerung:

Alle Argumente dafür, dass es angemessen ist, bei der Beschreibung der Vorgänge im Gehirn von Neuronen und neuronalen Zuständen auszugehen und nicht von Molekülen (oder Atomen, oder Elementarteilchen), gelten identisch oder analog auch dafür, dass es angemessen ist, geistige Zustände und Prozesse durch geistige bzw. psychologische Termini und Konzepte – wie etwa "Wahrnehmung" oder "Gedanke" – zu beschreiben und nicht durch neuronale. Ebenso wie Neurone *als Neurone* interagieren, so interagieren auch geistige Zustände *als geistige Zustände* miteinander; Genauso wie gesagt werden kann: der Output des neuronalen Moduls A verursacht den Zustand des neuronalen Moduls B, so kann auch behauptet werden: der geistige Prozess A hat den geistigen Prozess B zur Folge – eben genau auf die Art, wie das z.B. bei Gedankengängen der Fall ist.

Geistige Zustände und Prozesse sind also ebenso wie Neurone *untereinander vernetzt*, sie bedingen einander, sie bestimmen ihre eigene zeitliche Abfolge, kurz: *sie bilden eine Schicht der Wirklichkeit, die ebenso eigenständig ist wie die neuronale.*

So, wie man bei der Beschreibung der neuronalen Dynamik von der materialen Beschaffenheit der kybernetischen Elemente (der einzelnen Neurone) absehen kann – d.h. von der Tatsache, dass es Aggregate von Molekülen sind –, so kann man auch bei der Beschreibung der geistigen Dynamik von der materialen Beschaffenheit der kybernetischen Elemente (der einzelnen geistigen Zustände) absehen – d.h. von der Tatsache, dass es globale neuronale Muster sind.

Kurz gesagt: *Geistige Prozesse haben genau denselben Anspruch auf eine spezifische, eigenständige Beschreibung wie neuronale Prozesse.*

Aussagen über geistige Zusammenhänge – wie etwa die Behauptung, die den Reduktionisten von Philosophen entgegengehalten wird, dass auf der geistigen Ebene *Gründe* und nicht *Ursachen* gelten, sind also ebenso gerechtfertigt wie Aussagen über neuronale Zusammenhänge – wie etwa die Behauptung, dass eine Vorstellung zur Erinnerung wird, indem die Verbindungen zwischen den gleichzeitig aktiven Neuronen verstärkt werden.

Wir sind also zu folgendem Schluss gelangt:

Wenn die neuronale Schicht der Wirklichkeit als kybernetisches System beschrieben und verstanden werden kann, mit einer eigenen spezifischen Struktur und einer damit zusammenhängenden neuronalen Dynamik, dann gilt das in gleicher Weise für die geistige Schicht der Wirklichkeit: auch sie kann als kybernetisches System aufgefasst werden, das seine eigene spezifische Struktur hat, aus der wiederum eine spezifische *geistige* Dynamik folgt, d.h. eine Abfolge geistiger Zustände, deren

Gesetzmäßigkeit mit der Struktur des Raumes geistiger Zustände zusammenhängt. Wenn es also angemessen ist, neuronale Zustände durch Begriffe und Zusammenhänge auszudrücken, die sich aus der Betrachtung der neuronalen Aktivität als kybernetisches System ergeben, dann ist es auch angemessen, geistige Zustände durch psychologische bzw. geistige Begriffe und Gesetzmäßigkeiten zu beschreiben, die sich aus der Betrachtung der geistigen Aktivität als kybernetisches System ergeben.

Das bedeutet aber zugleich:

Wäre die Annahme korrekt, dass psychologische Beschreibungen durch neuronale ersetzt werden müssten, dann würde sich diese Annahme selbst aufheben, denn dann müsste die neuronale Beschreibung durch eine molekulare ersetzt werden, diese wiederum durch eine atomare usw. Der Reduktionsprozess kann erst dann enden, wenn die elementare Ebene erreicht ist – außer es gibt besondere Gründe für die Eigenständigkeit einer bestimmten Schicht. Solche Gründe gibt es aber für keine Schicht der Wirklichkeit: was für die Eigenständigkeit irgendeiner Schicht spricht, gilt in gleichem Maß für alle anderen.

Zwar ist diese Rechtfertigung einer eigenständigen Beschreibung des Geistes durch mentale Begriffe und Konzepte *konditional* – sie gilt nur für den Fall, dass irgendeine Beschreibung außer der elementaren zulässig ist –, aber das ist keine wirkliche Einschränkung, da es ja schlichtweg absurd wäre, auf alle anderen Beschreibungen zu verzichten.

Also stellt der eben durchgeführte Gedankengang eine vollständige Rechtfertigung der mentalen Beschreibung geistiger Zustände dar. Er enthält aber keine Begründung, wie diese Beschreibung in einen naturwissenschaftlichen Rahmen eingepasst werden kann, und deshalb ist er hier bloß als Einleitung gedacht, als Vorübung, die geeignet ist, aktuelle Denkmuster in Frage zu stellen und auf die nun folgende grundsätzliche Analyse desselben Szenarios vorzubereiten.

3.2. Der Weg zur geistigen Wirklichkeit über sieben Stationen

Um die Entwicklung des folgenden Gedankengangs übersichtlicher zu gestalten, werde ich die verschiedenen jeweils erreichten Stadien als Überschriften hervorheben.

Ausgangspunkt der Analyse ist die Annahme A_N , die in 2.2. als *Vollständigkeitsaxiom* bezeichnet wurde:

A_N : Alles, was sich ereignet, folgt aus Anfangsbedingungen und Gesetzen.

Zunächst eine Bemerkung zu den Voraussetzungen:

Aus physikalischer Sicht besteht die Wirklichkeit aus Objekten, die miteinander wechselwirken. Anfangsbedingungen sind die Werte der Variablen eines Systems – z.B. der Orte und Impulse aller Teilchen – zu irgendeinem Zeitpunkt, Gesetze sind mathematische Beziehungen zwischen diesen Variablen, d.h. Gleichungen.

Hier noch eine ausführlichere Formulierung von A_N , durch die die Annahme der (möglichen) strukturellen Übereinstimmung von Wirklichkeit und Beschreibung hervorgehoben wird:

A_N besagt: Die **Information**, die **in der Natur selbst** vorhanden ist und von der die zeitliche Entwicklung eines Systems abhängt, ist **vollständig in den Anfangsbedingungen und Gleichungen** des Systems enthalten. So, wie in der Wirklichkeit die Zukunft aus der Gegenwart folgt, folgen die künftigen Variablenwerte des Gleichungssystems aus den gegenwärtigen. Das Gleichungssystem ist ein Abbild der Gesetze, durch die festgelegt ist, wie die Zukunft aus der Gegenwart entsteht.

Die Formulierung "Anfangsbedingungen und Gesetz" erweckt zunächst den Eindruck, als wäre "das Gesetz" eine Art Input-Output-Maschine, die, wenn sie als Input die exakten Anfangsbedingungen erhält, als Output die Werte (Wahrscheinlichkeitsverteilungen) aller Variablen zu jedem beliebigen Zeitpunkt liefert. Bezeichnen wir diese erste, vorläufige Vermutung als

Standpunkt 1:

*Die Anfangsbedingungen sind singulär, das Gesetz ist allgemein, d.h. die Anfangsbedingungen werden **einmal** in die Gesetzesmaschine eingegeben (in das Gleichungssystem eingesetzt), und danach enthält der Gesetzesmechanismus das gewünschte Ergebnis **für jeden beliebigen Zeitpunkt**.*

Diese Erwartung bestätigt sich bei einigen einfachen klassischen Beispielen: etwa bei einem schwingenden Pendel, oder bei einem Planeten, der um seine Sonne kreist, oder auch beim allgemeinen Fall zweier gravitierender Körper, die – wenn sie nicht gestört werden – in alle Ewigkeit ihren elliptischen Tanz um den Massenmittelpunkt vollführen (jedenfalls in der Newtonschen Näherung).

Bleiben wir beim Fall gravitierender Massen, deren Dynamik durch das von diesen Massen selbst erzeugte Gravitationsfeld bestimmt wird.

Bei mehr als zwei Körpern ändert sich die Situation auf grundlegende Weise: Die Differenzialgleichungen sind nicht mehr integrierbar, und es ist daher nicht mehr möglich, das Gesetz als

Mechanismus aufzufassen, der aus einem Input den zugehörigen Output produziert.¹² Stattdessen müssen numerische Lösungsverfahren angewendet werden.

Wie ist im Fall von, sagen wir, um irgendeine hinreichend große Zahl zu nennen, 1000 Körpern vorzugehen, die sich relativ zueinander bewegen und durch Gravitation aneinander gebunden sind?

Wenn die Anfangsbedingungen, also die Orte und Impulse aller Körper zur Zeit t_0 bekannt sind und für eine spätere Zeit t_1 berechnet werden sollen, dann muss die Zeitspanne $t_1 - t_0$ in Intervalle unterteilt werden; – in *wie viele*, hängt von der gewünschten Genauigkeit ab: je besser die Näherung sein soll, desto kleiner müssen die Intervalle sein.

Man beginnt mit dem ersten Zeitintervall und berechnet aus den gegebenen Anfangsbedingungen die Orte und Impulse aller Körper am Ende dieses Intervalls. Sie bilden die Anfangsbedingungen für das zweite Intervall. Danach wiederholt man denselben Vorgang für jedes weitere Intervall und erhält auf diese Weise schließlich das gewünschte Resultat mit beliebiger Genauigkeit (wenn man von möglichen Instabilitäten absieht, was aber für den folgenden Gedankengang nicht von Belang ist.)

Was hat sich damit gegenüber der vorherigen Sichtweise geändert?

Die Anfangsbedingungen erscheinen nun nicht mehr nur als anfängliche, einmalige Voraussetzung für den Gesetzesmechanismus, sondern als ständig wiederkehrende Voraussetzung, die den Akt der Informationsgewinnung permanent begleitet. Wenn man die Behauptung aufrecht hält, dass der Beschreibungsformalismus (im Prinzip) *alle* Informationen bereitstellt, die die Natur selbst für ihre zeitliche Entwicklung benötigt, dann werden Anfangsbedingungen und Gesetz sogar völlig gleichwertig, weil für die Gewinnung dieser Information die Dauer der Berechnungsintervalle gegen 0 gehen müsste. In bezug auf die Berechnung der Bahn irgendeines der Körper werden die Anfangsbedingungen zu *Randbedingungen*, die sich mit der Zeit differenziell ändern – genauso wie der Ort und Impuls des Körpers selbst.

Das System wird also durch *Rückkopplung* geregelt: Die lokalen Änderungen (der Orte und Impulse der einzelnen Körper) verändern zugleich die Randbedingungen (die globalen Verhältnisse) und

¹² Bei drei Körpern existieren Lösungen in Form konvergenter Reihen. Das ist aber für die hier durchgeführte Argumentation ohne Bedeutung – hier geht es nicht um Sonderfälle von Lösungen für eine kleine Anzahl miteinander wechselwirkender Körper, sondern um die Frage, was sich beim Übergang auf den allgemeinen Fall *sehr vieler* Körper ändert.

dadurch verändern sich wiederum die Wirkungen, die das Gravitationsfeld auf die Körper ausübt, also deren Beschleunigungen.

Die erste Vermutung muss also korrigiert werden, und wir gelangen zum

Standpunkt 2:

*Gesetz und Anfangsbedingungen sind gleichwertig. Es findet eine permanente **Rückkopplung zwischen globaler Struktur und lokaler Veränderung** statt.*

Schon durch diesen einfachen Schritt vom Standpunkt 1 zum Standpunkt 2 hat sich somit die gängige Vorstellung einer Kausalität "von unten" – also die Vorstellung, dass alles auf der Ebene der "elementaren Teilchen" determiniert ist – als falsch herausgestellt. Sie passt nur für Idealfälle, die niemals wirklich existieren. In realen Fällen folgt aus der Kausalität "von unten" allein überhaupt nichts; um Information über die Zukunft zu produzieren, muss sie mit der Kausalität "von oben" verbunden werden. Die Vorstellung der *Determiniertheit auf der Ebene elementarer Entitäten* muss durch die Vorstellung der *Rückkopplung*, d.h. des *Zusammenwirkens lokaler und globaler Gegebenheiten* ersetzt werden.

Für den nächsten Schritt wechseln wir in ein anderes Szenario. Wir betrachten eine *schwingungsfähige Membran*, die aus einer sehr großen Zahl von Teilchen gebildet wird, die durch elektromagnetische Kräfte aneinander gebunden sind. An ihrem Rand ist die Membran befestigt, so dass sie unter Spannung steht.

Hier gibt es

1. ein Naturgesetz: die Gleichungen des elektromagnetischen Feldes
2. Anfangsbedingungen: die Orte, Impulse und Ladungen der Teilchen zu einem bestimmten Zeitpunkt und
3. Randbedingungen: die *Form des Randes*, durch die die möglichen Schwingungsmuster der Membran festgelegt sind.

Es hat sich aber nun gegenüber dem vorherigen Szenario etwas Entscheidendes geändert: Zuvor waren die Orte und Impulse der Teilchen zu Beginn jedes Zeitintervalls die Anfangsbedingungen für den nächsten Rechenschritt. Wir haben sie als *veränderliche Randbedingungen* bzw. *globale Struktur* bezeichnet.

Die unter Punkt 3 genannten Randbedingungen entsprechen aber keineswegs diesen Randbedingungen des vorigen Szenarios. Vielmehr ist nunmehr die globale Struktur selbst durch weitere Bedingungen – nennen wir sie *Randbedingungen höherer Ordnung* – soweit eingeschränkt, dass sie einem eigenen Gesetz gehorcht: dem *Schwingungsgesetz der Membran*. Die unter Punkt 3 genannten Randbedingungen sind diese Randbedingungen höherer Ordnung.

Das Schwingungsgesetz der Membran ist kein Naturgesetz im üblichen Sinn. Es ist auch nicht aus Naturgesetzen ableitbar, sondern muss zu den Naturgesetzen, denen das Szenario gehorcht, hinzugefügt werden, um eine Beschreibung zu ermöglichen, die alle Informationen enthält. Diese Unabhängigkeit des Schwingungsgesetzes von den Naturgesetzen zeigt sich auch darin, dass man bei der Darstellung des Gesetzes von seiner physikalischen Realisierung absehen kann.

Die Behauptung, dass hier ein *weiteres Gesetz* auftritt, erscheint allerdings zunächst sonderbar. Ist es nicht nach wie vor einfach die elektromagnetische Wechselwirkung, die die Bewegungen der Teilchen und damit zugleich die Bewegungen der Membran bestimmt? Ist das Schwingungsgesetz nicht bloß die bequeme Darstellungsform der Dynamik einer eigentlich elektromagnetisch determinierten Teilchenkonstellation?

Die Antwort ist *nein*. Die elektromagnetische Wechselwirkung ist nur *eine* Voraussetzung. Für die *vollständige Beschreibung* benötigt man jedoch hier – ebenso wie beim vorigen Szenario und wie auch sonst immer – nicht nur das Gesetz, sondern auch die *Anfangsbedingungen* zu irgendeinem Zeitpunkt, also den *globalen Zustand des Systems*. Während es vorher jedoch notwendig war, die Zeitspanne zwischen Anfangsbedingungen und Resultat zu unterteilen und den globalen Zustand für alle Intervalle neu zu ermitteln, ist es nunmehr *wegen des zusätzlichen Gesetzes* möglich, für den Systemzustand eine Gleichung aufzustellen, die analytisch gelöst werden kann.

*Das neue Gesetz ist somit das **dominante** Gesetz: es bestimmt die globale und damit zugleich die lokale Dynamik. Das Naturgesetz ist untergeordnet: die elektromagnetische Wechselwirkung fügt sich dem Gebot des Schwingungsgesetzes.*

Ist aber nicht doch das ganze Szenario auf naturgesetzliche Weise entstanden, so dass auch die Randbedingungen, die hier die globale Dynamik auf gesetzmäßige Weise strukturieren, letztlich *doch* aus Naturgesetzen und Anfangsbedingungen abgeleitet werden können?

Abermals nein. Der Versuch, die Behauptung, dass die globale Dynamik jetzt *neuen* Gesetzmäßigkeiten gehorcht, durch Rückführung auf die Vergangenheit zu widerlegen, scheitert ganz einfach daran, dass man auch bei dieser Rückführung – gleichgültig, wie weit man auch zurückgeht – mit *Gesetz und Anfangsbedingungen* beginnen muss, und dass es dann – gemäß dem zuvor Gesagten –

wegen der *veränderlichen globalen Bedingungen* unmöglich ist, die Zukunft abzuleiten. Genau diese Tatsache verhindert die Rückführung existierender Systeme auf Naturgesetze und Anfangsbedingungen und ermöglicht das Auftreten neuer Gesetze. Die Naturgesetze allein bieten diese Möglichkeit nicht – sie bleiben immer gleich. Die globalen Bedingungen jedoch, die ein zweites, unverzichtbares Element der Organisation der Natur und ihrer Beschreibung darstellen, sind offen für die Entwicklung von Gesetzmäßigkeiten, die zusätzlich zu den Naturgesetzen auftreten.¹³

Dieser zweite Schritt hat uns von der Annahme, dass es nur Naturgesetze gibt, zur Annahme geführt, dass die globale Zustandsdynamik von Systemen nicht nur durch Naturgesetze, sondern auch durch weitere Gesetze bestimmt wird, die man (wie schon in 2.2. festgestellt) *Gesetze der Form* oder *Strukturgesetze* nennen könnte. Ich fasse zusammen:

Standpunkt 3:

Randbedingungen – oder allgemeiner: strukturelle Voraussetzungen – können dazu führen, dass die globalen Systemzustände zusätzlich zu den Naturgesetzen weiteren Gesetzen unterworfen sind. Es sind Strukturgesetze oder Gesetze der Form. Sie sind dominant und von der physikalischen Beschaffenheit des Systems unabhängig.

Zur Ausführung des nächsten Schritts wenden wir uns nun direkt menschlichen neuronalen Netzen zu.

Die einzelnen Elemente, aus denen sie aufgebaut sind – Neurone – sind Systeme, die mit der Membran des vorigen Szenarios auf folgende Weise vergleichbar sind: Der Aufbau bzw. die Struktur eines Neurons bestimmt als *Menge der Randbedingungen* die Dynamik der physikalisch-chemischen Prozesse, die im Neuron stattfinden, also die Dynamik der Systemzustände. Analog zum Schwingungsgesetz der Membran tritt hier ebenfalls ein *neues Gesetz* auf: das bekannte neuronale Input-Output-Gesetz. Auch hier ist es wieder möglich, von der physikalischen Realisierung abzusehen.

Standpunkt 4:

Neurone sind Systeme, deren (interne) Zustandsdynamik durch ein Strukturgesetz – das neuronale Input-Output-Gesetz – geregelt ist. Auch die Dynamik des neuronalen Netzes insgesamt gehorcht diesem Gesetz. Das Netz gleicht in dieser Hinsicht dem vorher beschriebenen Gravitationszenario, dessen Dynamik dem Gravitationsgesetz gehorcht. Das neuronale Input-Output-Gesetz kann als Wechselwirkungsgesetz der Neurone aufgefasst werden.

¹³ Die vollständige Argumentation hierzu folgt im nächsten Kapitel.

Damit sind wir nun also bei jener Ebene der Wirklichkeit angelangt, die Gegenstand der Hirnforschung ist: Das *neuronale Netz* wird als System betrachtet, dessen Elemente Neurone sind und dessen Zustände durch das *Wechselwirkungsgesetz der Neurone* (das neuronale Input-Output-Gesetz) festgelegt sind, ebenso, wie die Zustände von physikalischen Systemen durch Naturgesetze festgelegt sind.

Die groben Umrisse der Zustandsdynamik des Netzes sind direkt beobachtbar. Diverse bildgebende Verfahren zeigen, welche neuronalen Areale aktiv sind und ermöglichen dadurch Einblicke in die funktionelle Topologie des Gehirns. Die Kenntnis der Aktivität und Ausprägung neuronaler Strukturen lässt – in einem gewissen Maß – sogar Schlussfolgerungen über das Ausmaß der zugeordneten geistigen Leistungen zu, insbesondere im Fall pathologischer Einschränkungen.

Die Frage ist:

Entspricht die nun erreichte Ebene der Wirklichkeit – das neuronale Netz als System, das dem neuronalen Input-Output-Gesetz gehorcht – bereits der Ebene der geistigen Aktivität?

Die Antwort ist *nein*, und zwar aus folgendem Grund:

Nehmen wir an, wir hätten die Gleichung bzw. das Gleichungssystem eines menschlichen neuronalen Netzes. (Auch wenn diese Annahme völlig absurd erscheint, so gibt es doch keinen Grund, die *Existenz* eines solchen Gleichungssystems auszuschließen.) Dann sind wir in genau derselben Lage wie beim zuvor beschriebenen Szenario zahlreicher gravitierender Körper. Wir haben *das Gesetz*: das Gleichungssystem des Netzes, und *die Anfangsbedingungen*: die Werte der Variablen, durch die die Zustände der einzelnen Neurone definiert sind – genauso, wie beim Gravitationsszenario das Gesetz und die Anfangsbedingungen gegeben waren.

Wer nun meint, daraus würde bereits die weitere zeitliche Entwicklung des Systems folgen, der würde demselben Irrtum unterliegen, durch den der *vorläufige Standpunkt 1* charakterisiert war – der irrigen Annahme also, dass das Gesetz als Input-Output-Mechanismus funktioniert, in den bloß, als einmaliger Akt, die Anfangsbedingungen als Input eingesetzt werden müssen, und der dann – ohne irgendwelche Zwischenschritte – jede beliebige Information über künftige Systemzustände als Output liefert.

Diese Annahme war beim Gravitationsszenario falsch, und sie ist auch beim neuronalen Netz falsch, weil es in beiden Fällen eine *permanente Rückkopplung* zwischen dem *globalen Systemzustand* und den *lokalen*, durch das Gesetz *und* den globalen Zustand determinierten Veränderungen gibt.

Betrachten wir zum Vergleich nochmals das Gravitationsszenario: Die Änderungen der Variablen der Körper – der Orte und Impulse – ergeben sich aus der Struktur des Gravitationsfeldes, d.h. aus dem Zustand des gesamten Systems: das Gravitationsgesetz kann angewendet werden, um die Bahnen der Körper zu bestimmen – jedoch nur für eine (sehr) kurze Zeitspanne, denn infolge der lokalen Änderungen ändert sich auch das Feld. Diese Änderung muss beim nächsten Schritt berücksichtigt werden. Das veränderte Feld bewirkt nun Änderungen der Orte und Impulse, die wiederum zugleich eine Änderung des Feldes bedeuten usw. Die *singulären Anfangsbedingungen* werden zu *veränderlichen globalen Bedingungen*.

Auf Grund dieser permanenten Rückkopplung ist es, wie anfangs festgestellt, keinesfalls möglich, Information über die zeitliche Entwicklung des Systems zu erhalten, indem man die Anfangsbedingungen in die Gleichungen einsetzt und dann ein Verfahren zum Auffinden von Lösungen anwendet. *Es existiert kein solches Verfahren.*

Beim neuronalen Netz sind die Verhältnisse analog: Alle Elemente des Systems (Neurone) sind – entweder direkt oder über wenige Zwischenschritte – aneinander gekoppelt. Die Änderungen der Variablen der Neurone – Zahl der Synapsen, Verbindungsstärken, Aktivierungsgrad, Frequenz – sind eine Folge des jeweiligen globalen Systemzustands, d.h. das Gesetz – das Gleichungssystem des Netzes – kann angewendet werden. Aber genauso wie zuvor kann auf diese Weise nur Information über eine (sehr) kurze Zeitspanne hinweg gewonnen werden: die lokalen Änderungen ändern zugleich den Systemzustand, der wiederum anders auf die lokalen Variablen wirkt usw. Und – wie beim Gravitationsszenario – folgt aus der Existenz dieser massiven Rückkopplung, dass kein Verfahren zum Auffinden von Lösungen für die Gleichungen des neuronalen Netzes existiert.

Standpunkt 5:

Auf Grund der Rückkopplung zwischen globalen und lokalen Bedingungen ist es nicht möglich, genaue Werte der Variablen für irgendeinen zukünftigen Zeitpunkt zu berechnen.

Es gibt keinen Algorithmus, der von einem bekannten Zustand A des Netzes zu einem künftigen Zustand B führt.

Es müsste somit, wie beim Gravitationsszenario, auf numerische Näherungsverfahren zurückgegriffen werden, *es sei denn*, es gibt weitere Gesetze, die, wie beim Membran-Szenario, die Dynamik des Netzes zusätzlich zum neuronalen Input-Output-Gesetz regeln.

Damit stehen wir also vor der nächsten grundsätzlichen Frage:

Ist das Wechselwirkungsgesetz der Neurone – das neuronale Input-Output-Gesetz – das einzige Gesetz, dem das System gehorcht, oder gibt es weitere Gesetze der Art, die zuvor Strukturgesetze genannt wurden?

Anders gefragt: Ist, wie im Gravitationsszenario, das einzig wirkende Gesetz das Wechselwirkungsgesetz der Elemente des Szenarios, oder treten, wie bei der schwingenden Membran, infolge von *Voraussetzungen struktureller Art* weitere Gesetze auf, denen die globalen Systemzustände unterworfen sind?

Die Antwort ist einfach. Folgendermaßen:

"Geistige Zustände"¹⁴ sind als globale raum-zeitliche Muster des neuronalen Netzes definiert, bei denen etliche Areale des Netzes gleichzeitig aktiviert und miteinander verbunden sind.¹⁵

Damit sind die Verhältnisse denen bei der schwingenden Membran vergleichbar. Die Strukturierung des Netzes durch die darin möglichen geistigen Zustände kann analog zur Strukturierung der Membran durch deren mögliche Schwingungszustände betrachtet werden. So wie der *Rand der Membran* die Bedingung für das Auftreten von geordneten globalen Zuständen der Membran darstellt, so bildet auch *die Form des gesamten Netzes* (genauer: des Phasenraums des Netzes) die strukturelle Voraussetzung für das Auftreten von geordneten globalen Zuständen (Mustern) des Netzes, also von geistigen Zuständen.

Standpunkt 6:

Das neuronale Netz gehorcht zusätzlich zum neuronalen Input-Output-Gesetz einem weiteren Gesetz: seine Dynamik wird durch die globalen Muster strukturiert, die als "geistige Zustände" definiert sind. Ihre strukturierende Funktion kann analog zu der von globalen Schwingungszuständen eines physikalischen Systems verstanden werden: wie diese sind geistige Zustände Attraktoren der globalen Dynamik. (Mehr darüber in 3.4.)

¹⁴ Ich verwende den Terminus "Zustand" hier – analog zum "Schwingungszustand" der Membran – synonym für "raum-zeitliches Muster". Ein geistiger Zustand ist also zeitlich ausgedehnt – im Gegensatz zum physikalischen Begriff "Zustand", der sich nur auf einen bestimmten Zeitpunkt bezieht.

¹⁵ Als weitere Bedingung muss hier angenommen werden, dass dabei auch Teile des Netzes aktiv sind, die keine spezifische Funktionalität haben (sogenannte assoziative Felder). Nur wenn das der Fall ist, können die globalen Zustände auch *geistige Zustände* sein, und nur dann gelten die folgenden Überlegungen. (Ich werde am Ende von 3.5. auf diesen Punkt zurückkommen und im fünften Kapitel nochmals darauf eingehen.)

Der nächste Schritte führt nun über die Membran-Analogie hinaus:

So, wie sicher angenommen werden kann, dass geistige Zustände *geordnete globale Systemzustände* sind, steht auch unbezweifelbar fest, dass sie sich *aufeinander beziehen*, dass sie also *untereinander vernetzt* sind.

Das bedeutet: Während die Übergänge zwischen verschiedenen möglichen Schwingungszuständen eines physikalischen Systems *von außen* verursacht werden, sind die Übergänge zwischen den geordneten globalen Systemzuständen des neuronalen Netzes – den geistigen Zuständen – Teil der *internen Dynamik* des Netzes. Dabei ist jedoch zu beachten, dass die Übergänge zwischen den globalen Zuständen – ebenso wie diese Zustände selbst – nicht aus dem neuronalen Input-Output-Gesetz ableitbar sind: wären die Übergänge ableitbar, dann wäre auch die Folge der globalen Zustände insgesamt ableitbar, und dann müssen auch die Elemente dieser Folge, also die Zustände selbst, aus diesem Gesetz ableitbar sein. Da das aber nicht der Fall ist – sie folgen ja, wie zuvor festgestellt, nicht aus dem neuronalen Input-Output-Gesetz, sondern aus einem *neuen Gesetz der Form* –, muss den geistigen Zuständen also *eine eigene Dynamik* zugeschrieben werden.

Mit anderen Worten: Geistige Zustände beziehen sich *als geistige Zustände* aufeinander. Sie stehen miteinander in spezifischer *geistiger* Wechselwirkung, nicht anders als Teilchen miteinander in spezifischer physikalischer Wechselwirkung stehen.

Damit haben wir nun, ohne dabei jemals den Bereich der naturwissenschaftlichen Methodik zu verlassen, systematisch die Behauptungen begründet, die schon im vorigen Abschnitt 3.1. aufgestellt wurden. Hier die Wiederholung. Sie ist zugleich der

Standpunkt 7:

Geistige Zustände sind – genau wie Neurone oder neuronale Module – untereinander vernetzt, sie bedingen einander, sie bestimmen ihre eigene zeitliche Abfolge, kurz: sie bilden eine eigenständige Schicht der Wirklichkeit.

So, wie man bei der Beschreibung der neuronalen Dynamik von der materialen Beschaffenheit der kybernetischen Elemente (der einzelnen Neurone) absehen kann – d.h. von der Tatsache, dass es Aggregate von Molekülen sind –, so kann man auch bei der Beschreibung der geistigen Dynamik von der materialen Beschaffenheit der kybernetischen Elemente (der einzelnen geistigen Prozesse) absehen – d.h. von der Tatsache, dass es neuronale Muster sind.

ERST JETZT sind wir also auf der Ebene der geistigen Aktivitäten angelangt, d.h. beim Geist selbst. Erst dadurch, dass die globalen Muster des Netzes selbst als Elemente einer *Struktur höherer Ordnung* aufgefasst werden, können geistige Phänomene *als solche* in das naturwissenschaftliche Verständnis der Wirklichkeit integriert werden.

Im Grunde ist der Schritt vom Standpunkt 6 auf Standpunkt 7 nichts Neues – er stellt bloß eine Analogie des Aufsteigens von Molekülen zu Neuronen dar: Er ist ein weiteres Beispiel dafür, dass Aggregate, die aus einfacheren Elementen bestehen, selbst wiederum die Bausteine einer höheren Schicht bilden können, d.h. als Elemente einer dynamischen Struktur höherer Ordnung fungieren. Auf ebendiese Weise sind die globalen Muster des Netzes als Elemente einer höheren Schicht der Wirklichkeit aufzufassen, nämlich der geistigen.

Dem flüchtigen Blick könnte diese Tatsache verborgen bleiben, weil diese Schicht der Wirklichkeit sich niemals *als Ganzes* präsentiert. Das neuronale Netz enthält ja nicht die Aktivitätsmuster selbst, sondern gewissermaßen nur die *Herstellungsanweisungen* für diese Muster. Deshalb ist von allen möglichen geistigen Zuständen jeweils nur *ein einziger* tatsächlich realisiert. Während das neuronale Netz als kybernetisches System immer als Ganzes existiert, sodass seine Struktur (im Prinzip) vollständig sichtbar ist, kann das Netz geistiger Zustände daher nicht direkt beobachtet werden.

Betrachten wir als Beispiel geistiger Aktivitäten einen *Gedankengang*. In unserem Bild ist dies eine Folge geordneter Systemzustände. Jeder Gedanke stellt ein Muster bzw. selbst schon eine Abfolge von Mustern dar. Von diesen Mustern ist jeweils nur eines realisiert. Es ist kurze Zeit aktiv und geht danach in das nachfolgende Muster über. Das nachfolgende Muster kann aber durchaus *in denselben neuronalen Bereichen* realisiert sein wie das vorhergehende. Bei einer Tomographie wäre dieser Übergang dann gar nicht beobachtbar.

Will man also den strukturellen Zusammenhang geistiger Zustände erforschen, dann kann man nicht einfach eine existierende Struktur beobachten; Man muss sich vielmehr einen *virtuellen* Raum vorstellen, der aber (im Prinzip) durch ein Modell veranschaulicht werden könnte. Seine Elemente wären geistige Zustände – Verbindungen von Gedanken, Gefühlen, Vorstellungen usw. – und seine Struktur könnte z.B. durch Pfeile dargestellt werden, die von jedem Zustand zu den möglichen folgenden Zuständen führen, mit Angaben über die Wahrscheinlichkeit des jeweiligen Übergangs.

Die Struktur und Dynamik geistiger Aktivität kann also nicht durch irgendein bildgebendes Verfahren sichtbar gemacht werden. Man muss sich, um einen Überblick zu gewinnen, in den virtuellen Raum geistiger Zustände begeben. Bildgebende Verfahren informieren daher im Grunde nur über jene Aktivitäten und Zustandsfolgen des Netzes, die durch das neuronale Input-Output-Gesetz und durch die funktionelle Struktur des Gehirns determiniert sind und eben *nicht* durch jene Gesetzmäßigkeiten,

durch die die Folge geistiger Zustände, wie etwa die von Gedankengängen, geregelt wird. Es ist also kein Zufall, dass in der Hirnforschung, sofern sie sich auf bildgebende Verfahren stützt, nie von Schlussfolgerungen, Gedankengängen oder Einsichten die Rede ist, die man als das eigentliche Reich des Geistes bezeichnen könnte.

Betrachten wir kurz das bisher Erreichte. Wie weit sind wir gekommen?

Aus unserer Analyse folgt die *Eigenständigkeit des Geistes*. Es ist gezeigt worden, dass geistige Aktivitäten tatsächlich *als solche* beschrieben werden können – und müssen. Dabei sind wir mit der naturwissenschaftlichen Sicht der Welt keineswegs in Konflikt geraten; im Gegenteil: die Argumentation bezog sich ausschließlich auf naturwissenschaftliche Fakten und Methoden.

Die Beschreibung geistiger Phänomene durch die Begriffe, die wir dafür alltäglich gebrauchen, ist damit vollständig gerechtfertigt: Gedanken sind Gedanken, Gründe sind Gründe, Entscheidungen sind Entscheidungen. Dass sie *als sie selbst* auftreten können und nicht *eigentlich etwas anderes* sind, liegt daran, dass sie auf nichts anderes reduziert werden können. Der Grund dafür ist ganz einfach derselbe wie bei den Schwingungszuständen der Membran: wie diese gehorchen auch geistige Zustände einem zusätzlichen, neuen Gesetz und sind daher nicht auf eine einfachere Schicht der Wirklichkeit reduzierbar. Kurz gesagt: die Behauptung, geistige Aktivitäten seien *neuronale Phänomene* ist ebenso falsch wie die Behauptung, die Schwingungszustände einer Membran seien *elektromagnetische Phänomene*.

Geistige Phänomene sind somit *neue, nicht reduzierbare Phänomene*, deren eigenständige Beschreibung nicht bloß gerechtfertigt, sondern notwendig ist.

Besonders hervorzuheben ist, dass durch die bisherigen Schlussfolgerungen aber nicht nur die Eigenständigkeit geistiger Phänomene begründet wird, sondern dass dadurch geistige Zustände auch als *Ursache* physikalischen Geschehens verstanden werden können. Eine Aussage wie: "Diese Person hat so gehandelt, weil sie es für richtig hielt", ist ebenso gerechtfertigt wie die Aussage: "Die rote Kugel wurde versenkt, weil sie von der weißen Kugel im richtigen Winkel angestoßen worden ist". Beide Aussagen sind angemessene kausale Beschreibungen der Vorgänge in der jeweiligen Schicht der Wirklichkeit.

Zur Erinnerung: Die Begründung setzt sich aus zwei Tatsachen zusammen. Die erste ist bereits beim Übergang vom ersten auf den zweiten Standpunkt aufgedeckt worden: der *globale Zustand* eines Systems ist für die künftige Entwicklung des Systems von gleicher Bedeutung wie die lokalen Gegebenheiten (etwa Orte und Impulse von Teilchen).

Dazu kommt als zweite Tatsache, dass es Systeme gibt, deren globale Zustände durch Gesetze bestimmt werden, die zusätzlich zu den Naturgesetzen existieren. Der globale Zustand dominiert dann über die lokalen Verhältnisse. Daher ist es angemessen, ihn als *Ursache* der lokalen Dynamik aufzufassen. In genau diesem Sinn sind geistige Zustände Ursache der Veränderungen neuronaler Variablen und damit zugleich auch Ursache von Handlungen.

Auf diese einfache Weise ist erklärt, wie "Geist" auf "Materie" einwirkt.

Damit sind die Geistigen Zustände und Prozesse selbst Teil der Naturkausalität geworden, aber eben nicht so, wie sie in der klassischen Antinomie von Freiheit und Kausalität auftreten, also nicht als *verursacht*, sondern als *verursachend*. (Ob bzw. in welchem Maß sie selbst verursacht sind, wird im nächsten Abschnitt beantwortet.)

Auch das Auftreten von *Gründen* kann erst dann verstanden werden, wenn man die Tatsache berücksichtigt, dass geistige Zustände untereinander vernetzt sind, dass sie also selbst wiederum Elemente einer weiteren Schicht der Wirklichkeit sind, eines Systems höherer Ordnung, das eine eigene Dynamik hat, d.h. eine Abfolge von Zuständen, die gewissen Regeln gehorcht. *Gründe* sind Beispiele für diese Regeln.

Allerdings fehlen in den bisherigen Überlegungen wichtige Differenzierungen. So ist z.B. klar, dass nicht alle Zustände neuronaler Netze als geistige Zustände aufgefasst werden können. Einfache neuronale Netze funktionieren wie Automaten. Sie sind vollständig durch ihre Schaltungen – also durch das neuronale Input-Output-Gesetz – und durch ihre funktionelle Architektur determiniert. Es wäre also wünschenswert, zu klären, unter welchen Voraussetzungen die verwendeten Analogien und Argumente gelten. (Ich werde das im Abschnitt 3.5, unter "Bemerkungen", nachholen.)

Andererseits geht es hier aber vor allem um die *prinzipielle Möglichkeit*, Geist als natürliches Phänomen zu verstehen und in das naturwissenschaftliche Weltbild zu integrieren, und genau das leisten die bisherigen Schlussfolgerungen: sie zeigen, dass diese Möglichkeit besteht. Geist kann als eigenständige Schicht der Wirklichkeit begriffen werden, die auf keine andere, tiefer liegende Schicht (neuronal, molekular usw.) reduzierbar ist. Psychologische Konzepte und Termini sind daher für die Beschreibung angemessen, und geistige Prozesse sind Ursache für physikalische Veränderungen.

Was allerdings noch fehlt, ist die Begründung der *Freiheit* des Geistes. Es ist zwar gerechtfertigt, Entscheidungsprozesse als geistige Phänomene aufzufassen, aber es ist auf dem gegenwärtigen Stand unserer Überlegungen noch nicht gerechtfertigt, sie *frei* zu nennen.

3.3. Der letzte Schritt: die Begründung der Willensfreiheit

Vor unserem Blick ist nun folgendes Szenario ausgebreitet:

Das neuronale Netz besteht aus mehreren übereinanderliegenden Schichten ansteigender Komplexität.

Die unterste Schicht – bezeichnen wir sie als *Bereich erster Ordnung* – ist die der **Atome und Moleküle**. Sie gehorchen einem Naturgesetz: dem *Gesetz der elektromagnetischen Wechselwirkung*. Ihre Dynamik wird in den Neuronen jedoch von einem *Gesetz zweiter Ordnung* (einem Strukturgesetz) bestimmt: dem *neuronalen Input-Output-Gesetz*.

Die mittlere Schicht – der *Bereich zweiter Ordnung* – ist die der **Neurone**. Sie gehorchen einem Strukturgesetz: dem *neuronalen Input-Output-Gesetz*, das zugleich ihr *Wechselwirkungsgesetz* ist. Ihre Dynamik wird jedoch von einem *Gesetz dritter Ordnung* (einem Strukturgesetz) bestimmt: dem *Gesetz der Abfolge geistiger Zustände*.

Die oberste Schicht – der *Bereich dritter Ordnung* – ist die der **geistigen Zustände**. Sie gehorchen einem Strukturgesetz: dem *Gesetz der Abfolge geistiger Zustände*, das zugleich ihr *Wechselwirkungsgesetz* ist.¹⁶

In diesem Szenario gibt es allerdings einen Punkt, der einer Ergänzung bedarf. Ich komme gleich darauf zurück. Nehmen wir zunächst an, dass alle Aussagen zutreffen.

Nun fügen wir diesen Aussagen eine weitere Aussage hinzu – diejenige, die im Abschnitt 2.2. abgeleitet worden ist. Sie lautet:

Von den beiden Behauptungen

1. *Willensfreiheit existiert* 2. *Das Vollständigkeitsaxiom A_N ist korrekt*

ist höchstens *eine* wahr.

¹⁶ Es überrascht vielleicht, dass hier die Schicht der neuronalen Bereiche unterschiedlicher Funktionalität fehlt, die miteinander ebenfalls auf gesetzmäßige Weise zusammenhängen. Die Berücksichtigung dieser Schicht als weiteren Ordnungsbereich würde jedoch nichts Grundsätzliches ändern, sondern bloß die Argumentation erschweren. Deshalb tritt diese Schicht hier nur in Form der Bedingung auf, dass bei geistigen Zuständen *assoziative Felder* aktiv sein müssen.

Zur Erinnerung nochmals das Vollständigkeitsaxiom:

Alles, was sich ereignet, folgt aus Anfangsbedingungen und Gesetzen.

Überprüfen wir das eben erstellte Szenario. Offensichtlich ist es in der Form, wie es soeben notiert wurde, tatsächlich zur Gänze durch Anfangsbedingungen und Gesetze festgelegt, und daraus würde nun – wie in 2.2. festgestellt – folgen, dass die Annahme, *wir selbst* wären die Urheber unserer Handlungen und würden sie nach unserem freien Willen wählen, genauso absurd wäre wie die Annahme, nicht nur Gravitation würde die Bahn der Himmelskörper bestimmen, sondern auch ein *vielarmiger kosmischer Affe*, der sie nach seinem Gutdünken lenkt.

Untersuchen wir also, ob irgendeine der Behauptungen, die das Szenario enthält, unvollständig oder falsch sein könnte.

Beginnen wir mit den Anfangsbedingungen. Deren Existenz ist selbstverständlich. Am hierarchischen Aufbau ist ebenfalls nicht zu zweifeln. Also muss das Problem in den *Gesetzesannahmen* liegen.

Drei Arten von Gesetzen verschiedener Ordnung treten auf: die *elektromagnetische Wechselwirkung*, das *neuronale Input-Output-Gesetz* und das *Gesetz der Abfolge geistiger Zustände*.

Die ersten beiden Gesetze existieren mit Sicherheit. Hinsichtlich ihrer Definition gibt es kein Problem. Sie gelten stets unverändert.

Durch dieses Ausschlussverfahren haben wir den problematischen Punkt im Szenario ausgemacht – es ist die einzige übrig bleibende Annahme, nämlich die Behauptung:

Es gibt ein Gesetz, das die Abfolge geistiger Zustände regelt.

Was ist daran problematisch? Es kann doch kein Zweifel daran bestehen, dass geistige Prozesse gewissen Regeln folgen?

Untersuchen wir, was geschieht, wenn geistige Prozesse ablaufen. Gemäß unseren Voraussetzungen gehorcht ihr Ablauf einem Gesetz dritter Ordnung.

Jetzt beziehen wir in unsere Überlegung die bekannte Tatsache ein, dass *jede neuronale Aktivität die neuronale Struktur verändert*. Das jeweils aktive Muster wird verstärkt, nicht aktive Muster werden

abgeschwächt, was durch Veränderungen der synaptischen Aktivität, aber auch durch Wachstum oder Abbau von Dendriten oder sogar Bildung oder Abbau von Neuronen bewirkt werden kann.¹⁷

Das bedeutet:

Die geistige Aktivität wirkt auf sich selbst zurück. Sie verändert sich selbst, indem sie ihre neuronale Codierung ändert: durch die Aktivierung eines einzigen geistigen Zustands erfahren alle geistigen Zustände, bei denen sich die neuronalen Bereiche, in denen sie codiert sind, mit denen des aktiven Zustands überschneiden, eine Veränderung. (Da jeder geistige Zustand über weite Bereiche des Netzes verteilt ist, überschneiden sich die Bereiche vieler, wenn nicht sogar aller geistigen Zustände.) Damit ändern sich aber zugleich die Regeln für die Abfolge geistiger Zustände.

Zum Vergleich: Bei Prozessen, die in physikalischen Systemen stattfinden, werden zwar laufend neue Zustände produziert, d.h. die Werte der Variablen ändern sich, die Gesetze bleiben jedoch gleich, und auch die Struktur des Zustandsraumes bleibt gleich, sofern äußere Einflüsse ausgeschlossen werden.

Bei geistigen Prozessen werden hingegen nicht nur neue Zustände produziert, sondern auch *neue Regeln* der Abfolge dieser Zustände, und die Struktur des Zustandsraumes ändert sich permanent. Auch wenn äußere Einflüsse ausgeschlossen werden, verändert das System durch Rückkopplung unaufhörlich seine eigenen Voraussetzungen. Und diese Änderung – auch wenn das Hebb'sche Gesetz ihre notwendige Voraussetzung ist – muss der dominanten Schicht des Szenarios zugeschrieben werden, d.h. der geistigen Schicht. Sie ist ein *geistiges* Phänomen.

Es existieren also keine feststehenden Regeln dritter Ordnung.

Es ist aber noch eine weitere Frage zu klären: Gibt es vielleicht gleichbleibende *Metaregeln*, also Regeln über die Veränderung der Regeln geistiger Aktivität?

Die Existenz *universell gültiger* Metaregeln würde einen universell gültigen Zusammenhang zwischen dem Informationsgehalt geistiger Zustände und dessen neuronaler Codierung voraussetzen. Ein solcher Zusammenhang existiert aber nicht. Falls Metaregeln existieren, sind sie daher systembezogen, und das heißt: auch sie sind den Veränderungen durch Rückkopplung unterworfen. Es gibt somit auch keine gleichbleibenden Metaregeln.

¹⁷ Diese Erkenntnis geht auf Donald Hebb zurück, der 1949 in *The Organization of Behavior* feststellte: *When an axon of cell A is near enough to excite B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.*

Wir haben also den Satz abgeleitet:

Die formalen Repräsentationen geistiger Prozesse sind aus keinem vorgegebenen System von Anfangsbedingungen und Gleichungen ableitbar. Falls Gleichungen bzw. Regeln für diese Repräsentationen und deren Übergänge existieren, werden sie fortwährend durch Rückkopplungsprozesse verändert.

Da die Codierung geistiger Zustände auf der neuronalen Ebene erfolgt, und weil die Elemente dieser Ebene ja wiederum aus physikalischen Elementen aufgebaut sind, folgt aus diesem Satz, wie in 2.2. angekündigt, auch der Satz:

Die physikalischen Zustände und Zustandsfolgen, durch die geistige Prozesse in neuronalen Netzen codiert werden, sind nicht aus irgendeinem vorgegebenen System von Gleichungen und Anfangsbedingungen ableitbar, und sie können nicht als Lösung eines solchen Gleichungssystems für irgendeinen bestimmten Zeitpunkt aufgefasst werden.

Das Vollständigkeitsaxiom der Naturwissenschaft ist daher falsch.

Gerade deshalb, weil geistige Prozesse *natürliche Prozesse* sind, gilt:

Jede formale Beschreibung der Natur durch ein System von Anfangsbedingungen und Gesetzen ist unvollständig.

Wenn das Vollständigkeitsaxiom falsch ist, *kann* Willensfreiheit existieren. Existiert sie?

Die Antwort ist *ja*, und diese Antwort setzt sich aus drei Aussagen zusammen, die zuvor abgeleitet worden sind:

1. Willensentscheidungen sind geistige Prozesse. Als solche sind sie nicht aus physikalischen Gegebenheiten ableitbar, und zwar nicht aus technischen, sondern aus prinzipiellen – oder sagen wir: aus metaphysischen Gründen.
2. Willensentscheidungen wirken *kausal* auf materielle Gegebenheiten ein.
3. Obwohl geistige Prozesse eigenen Regeln unterworfen sind, ist es nicht möglich, daraus eine Willensentscheidung abzuleiten: sie kann in diesen Regeln nicht enthalten sein, weil die Regeln durch den geistigen Prozess, der der Entscheidung vorausgeht, geändert werden können. Während dieser

Prozess stattfindet, können sich die Gesetze, denen er gehorcht, ändern – oder genauer: *er selbst* kann die Gesetze ändern, die vor seinem Beginn galten.

Zusammengefasst: Willensentscheidungen sind Ursachen von Handlungen, und sie sind nicht vorher festgelegt. Sie sind also frei.

3.4. Geordnete Zustände in neuronalen Netzen

Im letzten Abschnitt wurden geistige Zustände als *geordnete globale Zustände* eines neuronalen Netzes bezeichnet, durch die die Dynamik des Netzes gesetzmäßig strukturiert wird. Das kann als *Kausalität von oben* aufgefasst werden. Dieser Sachverhalt soll nun konkretisiert werden.

Ich werde dafür Modelle verwenden, die in einem so hohen Maß idealisiert und vereinfacht sind, dass sie die Bezeichnung "Modell" kaum verdienen. Die Rechtfertigung dieser Methode besteht darin, dass sie ermöglicht, grundlegende Eigenschaften neuronaler Netze auf besonders einfache und abstrakte Art zu erklären und direkt aufzuzeigen, auf welche Weise diese Eigenschaften mit mathematischen Prinzipien verknüpft sind und wie sie zu geistigen Leistungen aufsteigen können.

Stellen wir uns ein natürliches (biologisches) neuronales Netz von, sagen wir: einigen zehntausend Neuronen vor, das keine spezifische Funktionalität hat, in diesem Sinn also unstrukturiert ist. Es sei zunächst isoliert, d.h. ohne Verbindung nach außen. Seine elektrische Aktivität ist schwach, und sie ist *zufällig*, was einfach bedeuten soll, dass die Trajektorie im (physikalischen) Zustandsraum des Netzes keine erkennbaren Muster aufweist.

Jetzt verbinden wir das Netz mit der Außenwelt durch ein Sinnesorgan bzw. Eingabegerät, das optische Signale in neuronale Aktivität umwandelt. Der Zustand des Netzes hängt nun von diesem optischen Input ab.

Präsentieren wir nun dem Eingabegerät über einen gewissen Zeitraum ein *Objekt*. Wir nehmen an, dass sich im Netz auf Grund des spezifischen Inputs ein Muster ausbildet, das so lange fortbesteht bzw. sich wiederholt, wie das Objekt an Ort und Stelle bleibt.

Jetzt folgt der Punkt, der für die Strukturierung der Dynamik des Netzes entscheidend ist:

*Wegen des Hebbschen Gesetzes werden die neuronalen Verbindungen, die das gerade aktive Muster bilden, **verstärkt**.*

Mathematisch bedeutet das: *Das aktive Muster wird zum Attraktor.*

Der Zustandsraum des Netzes hat sich damit verändert: er ist nicht mehr unstrukturiert, sondern besitzt einen Attraktor, der ein Objekt repräsentiert.¹⁸

Der Attraktor hat einen *Einzugsbereich*. Es gibt also jetzt einen Teilbereich des Zustandsraums, innerhalb dessen die Trajektorien nicht mehr zufällig sind, sondern dem Attraktor zustreben.

Aus diesem einfachen Konzept folgt unmittelbar eine Reihe grundsätzlicher Aussagen, die nicht nur für unser simples Modellnetz, sondern ganz allgemein für neuronale Netze gelten, in denen es einen Bereich gibt, dessen Dynamik nicht schon von Beginn an (genetisch) festgelegt ist, sondern sich erst im Lauf der Zeit – wie soeben dargestellt – durch die Ausbildung von Attraktoren entwickelt:

1. Zunächst leistet der Begriff des Attraktors offenbar genau das, was von den geistigen Zuständen gefordert war: dass sie *geordnete Zustände* sind, die den *neuronalen Zustandsraum strukturieren* und die *Dynamik des Netzes bestimmen*.

Wenn geistige Zustände als solche Attraktoren aufgefasst werden, dann sind sie tatsächlich den Schwingungszuständen einer Membran vergleichbar, die ja ebenfalls Attraktoren des Zustandsraumes der Membran darstellen.

Es lassen sich aber noch weitere Schlüsse ziehen:

2. Es gilt: *Wahrnehmen = Wiedererkennen*.

Der Grund dafür ist folgender: Der Einzugsbereich des Attraktors liefert eine Definition für *Ähnlichkeit*. Jeder Input, der einen neuronalen Zustand verursacht, dem ein Punkt im Einzugsbereich des Attraktors entspricht, ist dem Input, der den Attraktor erstmalig erzeugt hat, hinreichend ähnlich, um als *dasselbe Objekt* erkannt zu werden; Dadurch, dass der Punkt im Einzugsbereich des Attraktors liegt, wird die Trajektorie auf den Attraktor zustreben und dann auf diesem verlaufen, so dass das neuronale Muster, das das Objekt repräsentiert, sich wieder ausbildet. Es entsteht also bei einem (hinreichend) ähnlichen Input nicht etwa ein *ähnliches*, sondern *dasselbe Muster*. Da kein Objekt bei seinem Wiedererscheinen einen identischen Input liefert – einige Variable ändern sich mit Sicherheit –

¹⁸ Dass ein neuronales Erregungsmuster, das hinreichend lange aktiv ist, zum Attraktor wird, ist keine Hypothese, sondern eine Tatsache: es ist bloß der mathematische Ausdruck für den erwiesenen neuronalen Sachverhalt, dass ein aktives Muster verstärkt wird.

ist die Annahme einer Konvergenz, wie sie das Attraktor-Konzept beinhaltet, für das Verständnis des Wiedererkennens von Objekten unerlässlich.¹⁹

In künstlichen neuronalen Netzen, die z.B. der Gesichtererkennung dienen, ist eine solche Konvergenz überflüssig: der Output muss keinen bestimmten Wert haben, es genügt, wenn er zu einem Intervall gehört. Diese "Rasterung" des Outputfeldes ist aber von außen – vom Programmierer – verfügt; In biologischen Netzen gibt es keine Rasterung, und daher ist für jede Art des Wiedererkennens die Konvergenz von Zuständen, die von verschiedenen Inputs verursacht werden, auf ein einziges Muster hin notwendig, das dann als Output für weitere Prozesse dienen kann.

Natürlich findet in biologischen neuronalen Netzen auch eine Merkmalsanalyse statt. Aber auch hier gilt wieder: kein Merkmal erzeugt mehrmals einen völlig identischen Input, und deshalb ist es abermals notwendig, Ähnliches als Dasselbe zu erkennen; Und das macht dann die Annahme notwendig, dass nicht nur das Objekt als Ganzes, sondern auch jedes seiner Merkmale durch einen Attraktor repräsentiert wird.

3. Das Attraktor-Konzept wirft außerdem ein Licht auf die philosophische Frage nach dem Status der *Allgemeinbegriffe*. Folgendermaßen:

Wie soeben gezeigt, werden Objekte durch Attraktoren repräsentiert. Wendet man nun dasselbe Prinzip auf die Repräsentationen selbst an – nimmt man also an, dass diese Repräsentationen selbst wiederum *intern* (im neuronalen Netz) durch Attraktoren repräsentiert werden, dann hat man damit die Ebene der *Begriffe* erreicht. Wird ein Begriff als neuronaler Attraktor aufgefasst, dann wird nun abermals *Ähnlichkeit* durch den Einzugsbereich des Attraktors definiert, und alle Objekt-Repräsentationen, die in diesem Sinn einander ähnlich sind, werden dadurch demselben Begriff zugeordnet.

¹⁹ Es sei denn, man nimmt an, das aktuelle Muster werde mit schon vorhandenen Mustern auf irgendeine Weise *verglichen*. Aber wie sollte ein solcher Vergleich stattfinden? Das würde doch zunächst voraussetzen, dass neben dem aktuellen Muster zusätzlich eine Reihe schon vorhandener, bekannter Muster aktiviert wird, damit der Vergleich durchgeführt werden kann. Das ist wenig wahrscheinlich, und überdies ist auch dann noch nicht klar, *wie* dieser Vergleich eigentlich durchgeführt werden könnte. Demgegenüber erscheint die Erklärung des Wieder-Erkennens durch das Attraktor-Konzept einfach und schlüssig: ein Objekt wird wiedererkannt, auch wenn der Input in einigen Variablen verändert ist, weil sich doch wieder derselbe Attraktor, d.h. dasselbe neuronale Muster ausbildet. Es ist dann unmittelbar einsichtig, wie ein Input, der ja auch dann, wenn er als Folge desselben Objekts auftritt, niemals identisch sein kann, stets demselben Objekt zugeordnet wird. (Allerdings muss diese Musterbildung im visuellen Gedächtnis stattfinden und nicht – wie das aus unserem zu stark vereinfachenden Modell geschlossen werden könnte – im Sehfeld selbst.)

Allgemeinbegriffe sind also Konstrukte, die aus den Gesetzmäßigkeiten der neuronalen Repräsentation folgen.

Wenn Objekte und Merkmale von Objekten durch Attraktoren im neuronalen Netz repräsentiert werden, dann sind wir überhaupt nicht imstande, Einzelnes als solches wahrzunehmen. Wir erkennen nicht Einzelnes, sondern nur Allgemeines. Darüber werden wir nur durch die Tatsache getäuscht, dass im Einzugsbereich eines bestimmten Attraktors meist nur *ein* Objekt der alltäglichen Umgebung liegt. Aber am Beispiel von Zwillingen zeigt sich sofort, dass das nicht immer der Fall ist.

Dasselbe gilt für das Denken: ebenso, wie wir nur Allgemeines *wahrnehmen*, können wir auch nur Allgemeines *denken*. Wenn wir etwas Bestimmtes – ein Einzelnes – meinen, dann wird es nur deshalb zum Einzelnen, weil es entweder raum-zeitlich lokalisiert ist, oder weil es das einzige Objekt ist, das allen allgemeinen Bestimmungen genügt, die auf es zutreffen, mit anderen Worten: weil es das einzige Objekt ist, das sowohl als Ganzes wie auch in bezug auf jedes seiner relevanten Merkmale im Einzugsbereich der jeweiligen Attraktoren liegt.

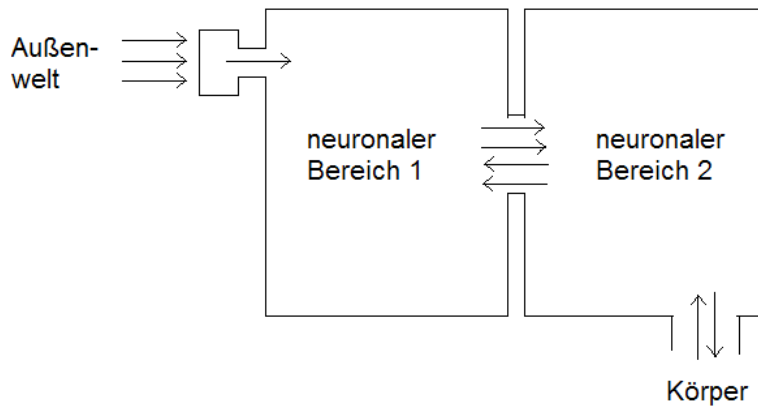
Kehren wir wieder zu unserem anfänglichen neuronalen Netz zurück. Nehmen wir an, der optische Input werde nun nicht mehr durch ein einziges Objekt erzeugt, sondern durch eine Reihe aufeinanderfolgender Objekte.

Jedem dieser Objekte entspricht ein neuronales Muster. Die Reihe von Objekten wird also durch eine Reihe von Attraktoren repräsentiert.

Wenn diese Objekte in derselben Reihenfolge wiederholt als Input für das Netz dienen, dann werden auch die Übergänge – oder genauer: die dabei aktiven neuronalen Verbindungen – zwischen den neuronalen Mustern verstärkt; Mathematisch formuliert: die Bahnen im Zustandsraum, die von einem Attraktor zum jeweils folgenden führen, werden zu *attrahierenden Trajektorien*, in dem Sinn, dass Bahnen, die hinreichend nahe an diesen Trajektorien verlaufen, auf sie zustreben.

Das bedeutet: Der Zustandsraum des Netzes ist jetzt nicht nur durch Attraktoren strukturiert, sondern auch durch Trajektorien, die von einem Attraktor zum nächsten führen – mit anderen Worten: die Repräsentationen der Objekte sind nun *assoziativ verbunden*.

Wir erweitern jetzt das Modell auf folgende Weise: Zusätzlich zum ersten Bereich, der mit der *Außenwelt* verbunden ist, existiere jetzt ein weiterer neuronaler Bereich, der einen Input erhält, den wir als *Körperinformation* auffassen. Die beiden neuronalen Bereiche sind miteinander verbunden, arbeiten aber weitgehend selbständig.



Was hat sich dadurch geändert?

Der Zustand des Netzes hängt jetzt nicht mehr nur von der *äußeren Information* ab, sondern auch von der *Körperinformation*. Er besteht aus zwei Attraktoren, die jedoch über Neuronen miteinander verbunden sind und permanent Information austauschen, so dass die Variablen, durch die ein Attraktor definiert ist, auch vom anderen Attraktor abhängen. Die Innen- und Außeninformationen werden auf diese Weise miteinander verknüpft. Auch der Gesamtzustand des Netzes kann als Attraktor aufgefasst werden.

An dieser Stelle kommt eine wichtige Eigenschaft von Attraktoren ins Spiel: Die Muster, denen sie entsprechen, bilden sich auch dann aus, wenn die Anfangsbedingungen nur teilweise realisiert sind. Eine Teilmenge der Variablenwerte, die zu einem Punkt im Einzugsbereich eines Attraktors gehören, genügt als Bedingung für die Ausbildung des Attraktors.

In unserem Zusammenhang bedeutet das Folgendes: Wenn ein Objekt (oder eine Situation) und ein körperlicher Zustand mehrmals gemeinsam auftreten, dann reicht *später* – infolge der neuronalen Verbindung beider Bereiche – das Muster, das sich in einem der beiden Bereiche ausbildet, für sich allein schon aus, um das zugehörige Muster im anderen Bereich zu verursachen. Der zur Körperinformation gehörende Attraktor stellt also den zur Objektinformation gehörenden Attraktor her und umgekehrt.

Nimmt man an, dass nicht nur der Körper das Netz steuert, sondern auch das Netz den Körper – was durch den nach unten führenden Pfeil in der Skizze angedeutet ist –, dann beeinflusst das Objekt den

Körperzustand: zunächst bildet sich der zum Objekt gehörende Attraktor aus, dieser erzeugt dann den Attraktor im zweiten Bereich des Netzes, und dieser stellt schließlich den zugehörigen Körperzustand her.

Durch diese Beziehungen werden Objekte (und Situationen) mit Körperzuständen verknüpft. Die Körperinformation stellt somit eine *Bewertung* der Objekte dar: die Objekte werden – je nach Art des körperlichen Anregungszustandes – positiv oder negativ bewertet. Überdies hängt das Ausmaß der Verstärkung aktiver Muster vom Grad der körperlichen Erregung ab.

Es ist nun Folgendes zu bedenken: Die neuronalen Muster, die Objekte repräsentieren, sind Attraktoren im Zustandsraum des Netzes, und die assoziativen Verbindungen zwischen den Objekt-Repräsentationen sind attrahierende Trajektorien in diesem Zustandsraum. Es handelt sich also um Strukturen, die *im Netz selbst existieren*, und das bedeutet, dass die Dynamik des Netzes auch dann durch diese Strukturen bestimmt wird, wenn *kein Input* vorhanden ist.

Das Netz stellt also selbständig Folgen von Repräsentationen her, die mit Körperinformationen verknüpft sind.

Dies kann als Grundlage von *Intentionalität* verstanden werden.

Damit sind die Möglichkeiten dieser einfachen Modellierung weitgehend ausgeschöpft. Immerhin waren sie ausreichend, um die Konzepte des vorigen Abschnitts, die dort in einer weitgehend abstrakten Weise dargelegt wurden, zu begründen und an die Realität heranzuführen. Durch seine Fähigkeit, ein bestimmtes Muster auch aus veränderten oder bruchstückhaften Anfangsbedingungen wiederherzustellen, erlaubt es das Attraktor-Konzept, zu verstehen, wie geistige Zustände, etwa Repräsentationen oder Gedanken, in assoziativer Folge aneinander gereiht werden. "Assoziativ verknüpft" bedeutet ja nichts anderes als: "Durch eine Trajektorie im Zustandsraum verbunden, die wiederholt durchlaufen und daher verstärkt wurde".

Ich will diesen kurzen Ausflug in die Gesetzmäßigkeiten neuronaler Netze mit einigen allgemeinen Bemerkungen zum Status naturwissenschaftlicher Beschreibungen neuronaler Leistungen beenden.

Einerseits ist klar, dass es allein schon wegen der Komplexität menschlicher Netze nicht möglich ist, geistige Phänomene mathematisch zu analysieren oder künstlich nachzubilden. Die Komplexität ist derart hoch, dass sie theoretisch und technisch nicht beherrschbar ist.²⁰

²⁰ Im Kapitel 5. *Qualia* wird gezeigt, dass die Simulation von Geist aus *prinzipiellen* Gründen unmöglich ist.

Auf der anderen Seite ist es jedoch einfach und weitgehend unproblematisch, ein grundsätzliches Verständnis davon zu erlangen, wie geistige Leistungen von neuronalen Netzen hervorgebracht werden können.

Ein Beispiel: Wir verstehen – wenn auch nicht in allen Details –, wie Objekte der Wirklichkeit visuell wahrgenommen werden. Wir wissen, wie die Information auf physikalischem Weg zur Retina und von dort über die Sehbahn zur primären Sehrinde gelangt, wie die gesamte Objektinformation in verschiedene Komponenten – Farbe, Textur, Merkmale, Kontrast, Umriss, Bewegung, Orientierung, Tiefe – zerlegt wird, die zunächst voneinander getrennt weiterverarbeitet und dann wieder zusammengeführt werden. Dieses Wissen vermittelt uns ein grundsätzliches Verständnis dafür, wie Information in neuronalen Netzen codiert wird und wie diese Codierung Objektrepräsentation ermöglicht.

Ein weiteres Beispiel: Bewusstsein. Es ist möglich, Bewusstsein auf einfache und schlüssige Weise durch die Annahme zu verstehen, dass es sich dabei um eine *Metarepräsentation* handelt, d.h. um einen mentalen Zustand, der dadurch charakterisiert ist, dass der Informationsgehalt von Zuständen des neuronalen Netzes selbst in einem anderen Bereich des Netzes repräsentiert und weiterverarbeitet wird. Von der Annahme, dass der Informationsgehalt von Wahrnehmungszuständen eine Repräsentation von Ausschnitten der Wirklichkeit darstellt, ist es kein großer Schritt mehr zur Annahme, dass in neuronalen Netzen von hinreichender Komplexität in gleicher Weise auch der Informationsgehalt von *inneren Zuständen* repräsentiert werden kann. Die Möglichkeit einer solchen Metarepräsentation erfordert bloß die Existenz einer weiteren, hierarchisch höheren Verarbeitungsstufe.

Ganz allgemein lässt sich behaupten, dass das Verständnis geistiger Phänomene auf der Basis der Annahme, dass das neuronale Netz ein informationsverarbeitendes System ist, zu keinen grundsätzlichen wissenschaftlichen oder philosophischen Problemen führt – allerdings immer vorausgesetzt, dass dabei nur vom *Informationsgehalt* der mentalen Zustände die Rede ist und nicht von ihrem *Empfindungsgehalt*.

Bemerkung:

Ich betone nochmals, dass die in diesem Abschnitt vorgestellten Modelle unrealistische Vereinfachungen sind. Dennoch erfüllen sie ihren Zweck: zu zeigen, wie die Struktur des Zustandsraums eines neuronalen Netzes entsteht und wie durch diese (zumindest kurzzeitig annähernd konstante) Struktur die neuronalen Abläufe bestimmt werden; – und das ist der Kern der Argumentation dafür, dass Kausalität hier von oben nach unten – von Geist auf Materie – wirkt.

3.5. Zusammenfassung

Da ich im nächsten Kapitel die Schlussfolgerungen, die zur Willensfreiheit geführt haben, noch einmal in allgemeinerer Form darstellen werde, kann ich mich hier kurz fassen.

Was ist die Ursache dafür, dass ein eingestrichenes c erklingt, wenn ich die Teetasse, die vor mir steht, mit dem Löffel anschlage? Was ist die Ursache für den Charakter dieses Tons? Nicht etwa die Tatsache, dass die Tasse aus Teilchen besteht, und auch nicht, dass die Teilchen elektromagnetisch wechselwirken. Nein, die Ursache für die Höhe und Klangfarbe des Tons ist die *Form* der Tasse.

Vom Wechselwirkungsgesetz folgt nur die Geschwindigkeit der Ausbreitung der anfänglichen Störung. Sie bestimmt zusammen mit der Form der Tasse die Tonhöhe. Der Toncharakter – die spezifische Obertonstruktur – hängt ausschließlich von der Form der Tasse ab.

Schon an diesem einfachen Beispiel zeigt sich, dass in Systemen, deren Dynamik durch Randbedingungen bestimmt wird, zusätzlich zu den Naturgesetzen spezifische Gesetze der Form auftreten. Diese Gesetze sind dominant. Der globale Systemzustand ist die *Ursache* für die lokalen Vorgänge; die Kausalität wirkt also nicht von unten nach oben, sondern von oben nach unten.

Auf analoge Weise lässt sich verstehen, dass Geist die Ursache für die Dynamik von menschlichen neuronalen Netzen ist, also auch die Ursache menschlichen Handelns und der zugehörigen physikalischen Veränderungen.

Geistige Zustände sind geordnete globale Zustände (Muster, Attraktoren) menschlicher (und einiger tierischer) neuronaler Netze. Als solche bestimmen sie die Dynamik des Netzes. Während aber die Tasse eine triviale Dynamik hat – sie besitzt nur einen einzigen möglichen geordneten (Schwingungs-) Zustand –, gibt es im neuronalen Netz eine ungeheure Zahl möglicher geistiger Zustände und Übergänge zwischen ihnen.

Diese Zustände sind nicht aus dem neuronalen Input-Output-Gesetz ableitbar (auch der Zustand der Tasse ist nicht aus dem elektromagnetischen Gesetz ableitbar), und dasselbe gilt für die Übergänge. Geistige Zustände müssen daher *als solche* aufeinander bezogen sein.

Der Bereich der Wirklichkeit, den wir "Geist" nennen, ist also eine eigenständige Schicht der Wirklichkeit, der eine eigene Dynamik zugeschrieben werden muss. Die Gesetze dieses Bereichs sind dominant. Was im Netz vor sich geht, ist somit durch *Geist* verursacht, nicht durch Neurone und die Gesetze neuronaler Interaktion oder durch die funktionelle Architektur des Netzes..

Um zu erfahren, was im Netz vor sich geht, muss man sich also auf die jeweilige geistige Aktivität und deren individuelle Regeln beziehen. Zu diesen Regeln gehören z.B. *Gründe*.

Es ist daher ebenso gerechtfertigt, jemanden nach seinen Gründen zu fragen, um herauszufinden, was er tun wird, wie es gerechtfertigt ist, den Punkt festzustellen, in dem eine Billardkugel auf eine andere trifft, um vorauszusagen, wohin sich die getroffene Kugel bewegen wird, oder wie es gerechtfertigt ist, die Form der Tasse zu bestimmen, um ihren Klang zu berechnen.

Allgemein gesprochen: *Für die Kenntnis der Dynamik eines Systems ist die Kenntnis der Gesetze erforderlich, die in diesem System dominant sind. Im Fall menschlicher neuronaler Netze sind das die zum jeweiligen Netz gehörenden (subjektspezifischen, veränderlichen) Gesetze der geistigen Aktivität.*

An diesem Punkt des Gedankengangs ist das Phänomen *Geist* vollständig in seine Rechte gesetzt, mit anderen Worten: Geist ist als das Phänomen wiedererstanden, als das er uns intuitiv gegeben ist.

Dafür war es insbesondere notwendig, dass wir uns von der verbreiteten Verwechslung von *neuronaler Aktivität* und *geistiger Aktivität* befreit haben.

Um nun zur Begründung der Freiheit zu gelangen, ist ein weiterer Schritt erforderlich.

Voraussetzung ist die Aussage: Die Gesetze der geistigen Aktivität bestimmen die Dynamik des Netzes.

Wären diese Gesetze *feststehend* – wie physikalische und neuronale Gesetze –, dann gäbe es keine Freiheit. Das ist aber nicht der Fall, weil die geistige Aktivität *auf sich selbst zurückwirkt*:

Die physiologische Voraussetzung dieser Rückkopplung ist das Hebbsche Gesetz: Benachbarte Neurone, die gleichzeitig aktiv sind, verändern sich physiologisch derart, dass ihre gegenseitige Stimulation verstärkt wird; umgekehrt werden nicht benutzte Verbindungen abgebaut. Die geistige Aktivität ändert also ihre eigenen Voraussetzungen. Sie wirkt auf ihre neuronale Codierung zurück, und damit ändert sie zugleich ihre eigenen Gesetze.

Im System geistiger Zustände werden daher nicht nur – wie in physikalischen Systemen – neue Zustände erzeugt, sondern auch neue Gesetze. Das System gerät in eine Rückkopplungsschleife. Die Regeln, denen die geistige Aktivität unterworfen ist, ändern sich durch genau diese Aktivität.

Es gibt also keine feststehenden Regeln, durch die festgelegt ist, was geschehen wird.

Das bedeutet: Willensentscheidungen folgen nicht aus irgendeinem System von Anfangsbedingungen und Gesetzen.

Wird die Natur als ein solches System aufgefasst, dann sind Willensentscheidungen, die in Zukunft stattfinden – genau wie alle andern geistigen Prozesse – in diesem System *unentscheidbare Aussagen*. Sie sind im System nicht ableitbar, und das gilt für *jedes* solche System.

Damit ist Raum für Freiheit geschaffen.

Um zu einem abschließenden Urteil über Willensfreiheit zu gelangen, definieren wir nun *Willensentscheidung*:

Eine Willensentscheidung ist ein geistiger Prozess, bei dem

1. die Folgen von Handlungsalternativen *bewertet* werden, und
2. die Alternative gewählt wird, die als die bessere beurteilt wird.

Folgende Aussagen wurden abgeleitet

A1: Geistige Zustände sind nicht aus physikalischen oder neuronalen Gesetzen und Anfangsbedingungen ableitbar. Sie sind untereinander vernetzt und bilden eine eigene Schicht der Wirklichkeit, die eigene Gesetze hat, die jedoch – im Gegensatz zu physikalischen Gesetzen – nicht feststehen. Zu diesen Gesetzen gehören z.B. Gründe.

A2: Die Kausalität wirkt von oben nach unten: geistige Prozesse sind Ursache von neuronalen Vorgängen.

A3: Selbst wenn die Information, die in einem beliebigen Zeitschnitt durch ein menschliches neuronales Netz enthalten ist, vollständig in ein System von Anfangsbedingungen und feststehenden Regeln übertragen werden könnte, so wären doch zukünftige geistige Prozesse in diesem System nicht ableitbar. Die Zukunft ist offen.

Die Aussagen A1, A2 und A3 genügen, um die Freiheit der Willensentscheidung im üblichen Sinn zu begründen.

Darüber hinaus wird durch sie aber genau definiert, was Freiheit im metaphysischen Sinn – innerhalb des sich in Übereinstimmung mit den Naturgesetzen entfaltenden Seienden – bedeutet.

Bemerkungen

Der Unterschied zwischen *geistigen* und *neuronalen* Prozessen

Es ist selbstverständlich, dass es in neuronalen Netzen auch Prozesse gibt, die nicht durch die geistige Aktivität, d.h. durch das Netz der Beziehungen zwischen den geistigen Zuständen selbst determiniert sind, sondern durch die funktionelle Architektur des Netzes, durch das neuronale Input-Output-Gesetz und durch äußere Gegebenheiten. In einfachen neuronalen Netzen finden sogar ausschließlich solche Prozesse statt.

Das Auftreten geistiger Zustände – und damit zugleich eine Dominanz des Geistigen über das Neuronale – ist erst dann möglich, wenn das Netz Bereiche enthält, die *funktionell ungebunden* sind.

Nur Bereiche dieser Art, in denen also, wie schon am Beginn dieses Abschnitts für das einfache Modellnetz gefordert, die neuronale Aktivität nicht durch physiologische Funktionen festgelegt, sondern zunächst zufällig ist, ist offen für die Strukturierung durch diejenigen geordneten Zustände (Attraktoren), die äußere (reale) oder innere (körperliche) Gegebenheiten repräsentieren.

Allerdings sind Zustände, die etwas repräsentieren, noch keine geistigen Zustände. Geistige Zustände müssen auch immer *aufeinander bezogen sein*. Erst durch dieses interne Netz von Beziehungen werden sie zu dem, was sie sind. Deshalb können Repräsentationen nur dann zu Geist werden, wenn sie sich untereinander vernetzen. Und dafür ist offensichtlich die Existenz funktionell ungebundener neuronaler Strukturen eine notwendige Bedingung.

Diese Hypothese, dass Geist nur bei Lebewesen zu finden ist, die über hinreichend entwickelte Bereiche dieser Art verfügen, wird auch durch die Erfahrung mit Tieren gestützt. Ihre geistigen Leistungen hängen von der Größe des Großhirns ab, und nur im Großhirn existieren solche Bereiche.²¹

Der Unterschied zwischen geistigen und neuronalen Prozessen lässt sich auch auf folgende Weise charakterisieren:

²¹ Der Begriff "Großhirn" darf jedoch nicht an die physiologische Struktur gebunden werden, die beim Menschen unter diesem Namen auftritt. Bei den Vögeln ist diese Struktur kaum entwickelt, weswegen ihre Intelligenz lange unterschätzt worden ist. Stattdessen haben sich bei ihnen anscheinend die Basalganglien erweitert, und die trotz des geringen Gehirnvolumens beachtliche Intelligenz einiger Vogelarten lässt vermuten (und hoffen), dass dieses neuronale Gewebe bzw. diese Art "Großhirn" sich zum Denken besser eignet als das menschliche.

Neuronale Prozesse können – zumindest weitgehend – beobachtet, gemessen und vorhergesagt werden. Bei geistigen Prozessen ist das nur in einem sehr geringen Maß möglich. Falls es sich um Gedankengänge handelt, dann gibt es nur eine einzige Methode, etwas darüber zu erfahren: Sie besteht darin, den, der denkt, nach seinen Gedanken zu *fragen*.

Das Ausmaß geistiger Aktivität

Geistige Zustände sind in neuronalen Netzen nicht von Anfang an vorhanden. Sie müssen sich erst entwickeln. In einem gewissen Ausmaß folgt diese Entwicklung vermutlich aus den Eigenschaften des neuronalen Systems.

Was jedoch über dieses Minimum hinausgeht, hängt davon ab, ob Erziehung und kulturelle Bedingungen die Entwicklung der geistigen Aktivität begünstigen. Das gilt auch für die Willensfreiheit. Sie existiert erst dann, wenn die Folgen von Handlungsalternativen beurteilt werden können, und diese Fähigkeit setzt jedenfalls Kenntnis und Urteilsvermögen voraus.

Wie auch immer man zur gegenwärtigen Konzentration der Aufmerksamkeit auf die materiellen (genetischen, neuronalen, chemischen usw.) Voraussetzungen der Persönlichkeit und des Denkens steht – sie wird jedenfalls zum Abbau der Selbständigkeit des Geistes beitragen; Geist existiert nur in dem Maß, in dem er als eigenständige Wirklichkeit begriffen und als solche ausgeübt wird.

Mit Phänomenen wie Willensfreiheit und Verantwortung verhält es sich nicht anders als mit beliebigen anderen geistigen Fähigkeiten, wie z.B. der Sprache: Wenn sie nicht erlernt werden – wobei es oft kritische Phasen gibt, nach deren Ende die jeweilige Fähigkeit nur noch unvollständig ausgebildet werden kann – dann existieren sie nicht oder bloß rudimentär.

Schluss

Wie sich herausgestellt hat, waren Philosophen, die behauptet haben,

- dass im Bereich des Geistigen Gründe und nicht Ursachen gelten,
- dass die Naturwissenschaftler, die Geistiges mit Neuronalem gleichsetzen und so groteske Formulierungen produzieren wie: *Mein Gehirn entscheidet, dass...* Kategorienfehler begehen,
- dass Freiheit als subjektive Gewissheit niemals durch naturwissenschaftliche Argumente aufgehoben werden könne,

immer schon im Recht. Allerdings könnten ihre wahren und schönen Behauptungen den rohen Ansturm der Naturwissenschaft ebenso wenig aufhalten wie der Gesang der Druiden den Angriff der römischen Heerscharen. Es ist einfach eine Tatsache, dass bis jetzt überhaupt keine Möglichkeit in Sicht war, die beiden Phänomenbereiche, die uns so selbstverständlich gegeben sind – Gegenständliches und Geistiges – in *einer* Naturvorstellung zu vereinen.

Seit Jahrhunderten scheinen sie einen unüberwindbaren Gegensatz darzustellen.

Die historischen Lösungswege sind nicht mehr gangbar: Der Descart'sche Dualismus wäre – ebenso wie alle anderen metaphysischen oder religiösen Dualismen – einfach absurd, und den Gegensatz auf Kantische Weise als Antinomie hinzunehmen, erscheint angesichts des enorm angewachsenen Wissens über die Natur, das uns zu einem einheitlichen Verständnis drängt, nicht mehr akzeptabel.

Die Folge sind Reduktionismen und Funktionalismen naturwissenschaftlicher Prägung.

Das naturwissenschaftliche Paradigma ist ja nicht nur unerhört erfolgreich, es ist auch von überzeugender Geschlossenheit: nichts kann den Naturgesetzen entgehen. Unter diesem Würgegriff geraten die Ideen des Geistes, des Ichs und der Willensfreiheit in Erstickungsgefahr. Wenn nicht einmal ihre innere Gewissheit und Selbstverständlichkeit diesen Ideen ausreichenden Schutz gegen die naturwissenschaftlichen Übergriffe gewährt, dann kann dies auch keinen philosophischen Argumenten gelingen.

Nein – wenn es für das naturwissenschaftliche Paradigma Grenzen gibt, dann müssen diese Grenzen aus ihm selbst klar werden.

Genau das haben wir hier erreicht. Gegenständliches und Geistiges können in einem auf naturwissenschaftlicher Basis errichteten Modell vereint gedacht werden, und zwar nicht reduktionistisch oder funktionalistisch, sondern der Erfahrung entsprechend, durch die uns beide Phänomenbereiche gegeben sind, mit ihrer Verschiedenheit, mit der Selbständigkeit des Geistigen und mit dem zugleich bestehenden komplexen gegenseitigen Abhängigkeitsverhältnis von Geist und Materie.

Dabei hat die Naturwissenschaft ihren Absolutheitsanspruch verloren. Es wurde gezeigt, dass Geist durch eine Beschreibung aus Anfangsbedingungen und Gesetzen nicht erfasst werden kann.

Weiter zum nächsten Kapitel [4. Das veränderte Bild der Wirklichkeit](#).

Weiter zu Kapitel [5. Qualia](#).