

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

Inhalt

5. Qualia.....	2
5.1. Einleitung.....	2
Bedeutung.....	3
5.2. Einschub: Seltsame Vermutungen.....	3
5.3. Warum Qualia in keiner Beschreibung enthalten sind.....	7
Noch einmal: das Problem.....	7
Substanz und Akzidens; das Wesen des Seienden.....	7
Die Antwort.....	12
Satz.....	14
5.4. Die Verwandlung des Seienden vom materiellen Ding zum Quale.....	14
Erste und zweite Substanz.....	14
Der Grund für die Verwandlung.....	15
Satz.....	22
Bemerkungen.....	23
5.5. Kriterium für das Auftreten von Empfindungen.....	26
Ein einfaches ergänzendes Argument.....	28
5.6. Wer oder Was hat Empfindungen?.....	29
5.7. Versuch einer begrifflichen Annäherung an die Verwandlung der Substanz.....	30
5.8. Philosophische Zombies.....	35
5.9. Künstliche Intelligenz.....	36
Verstehen.....	36
Warum Computer keinen Geist hervorbringen; die <i>formale</i> Begründung.....	38
Warum Computer keinen Geist hervorbringen; die <i>metaphysische</i> Begründung.....	46
5.10. Der metaphysische Unterschied zwischen Wirklichkeit und Simulation.....	47
Satz.....	48
Korollar.....	49

5. Qualia

5.1. Einleitung

Nun sind wir endlich bei der Frage angelangt, die das eigentliche Geheimnis des Geistes in sich birgt: der Frage nach dem Grund der Verwandlung eines Netzes von Information in eine Welt der Empfindungen.

Wie wird die Information zur Empfindung? Auf welche Weise verwandelt sich der Automat in ein empfindendes Wesen? Was ist der Unterschied zwischen beiden?

Für alles, was bisher über Geist gesagt wurde, war es ausreichend, die geistige Tätigkeit als Informationsverarbeitung aufzufassen. Die Eigenständigkeit und Dominanz des Geistigen sowie die Existenz der Willensfreiheit ließen sich damit begründen.

Jetzt aber, wo danach gefragt ist, wie eine Abfolge neuronaler Muster zu einem Strom von Erlebnissen wird, kann eine Antwort, die sich nur auf diese Annahme stützt, nicht mehr genügen: Solange man davon ausgeht, dass mentale Prozesse nichts als Informationsverarbeitung sind, bleibt man im Bereich der Informationsverarbeitung gefangen. Daran ändert sich auch nichts, wenn man Repräsentationen vernetzt, oder Repräsentationen von Repräsentationen – also höhere Stufen der Informationsverarbeitung – bildet, oder die Information auf sich selbst zurückwirken lässt: Gleichgültig, welche Funktion man auf Information anwendet – das Ergebnis ist immer bloß Information und sonst nichts. In dieser Betrachtungsweise kann sich keine Verwandlung ereignen. Die Information "rot" wird nicht zur Empfindung *rot*, die Information "Druck" wird nicht zur Empfindung *Schmerz*.

Es gilt also:

Die Annahme, dass Geist Informationsverarbeitung ist, ist für die Ableitung der Willensfreiheit notwendig und hinreichend.¹ Für die Begründung der Qualia ist sie jedoch bloß notwendig und nicht hinreichend.

Das bedeutet zugleich: Die Annahme, Geist sei *nichts als* Informationsverarbeitung, ist falsch.

Zunächst ist zu klären, warum Qualia in keiner Beschreibung enthalten sind. Danach werden wir uns der Frage zuwenden, warum aus einem physikalisch-physiologischen Zustand ein Quale wird.

¹ Das gilt aber nur für Informationsverarbeitung in *biologischen* neuronalen Netzen. Mehr darüber später.

Bedeutung

Geistige Zustände beziehen sich aufeinander. Bei Begriffen ist das selbstverständlich, aber es gilt auch für Wahrnehmungen: selbst wenn sie ein Objekt repräsentieren, ist ihr Inhalt nicht nur durch dieses Objekt bestimmt, sondern auch durch die Vernetzung mit anderen Elementen der geistigen Wirklichkeit.

Repräsentationen, die nicht vernetzt sind, sind auch keine geistigen Zustände. Der Roboter Hans (Abschnitt 2.2) repräsentiert die Eigenschaft eines Objekts, die wir als *rot* wahrnehmen, aber diese Repräsentation ist bei ihm kein geistiger Zustand.

Also ist *Vernetzung* der Repräsentations-Zustände eine notwendige Bedingung dafür, dass sie zu geistigen Zuständen aufsteigen können.

Um geistige Zustände von Repräsentations-Zuständen zu unterscheiden, werde ich den ***Informationsgehalt eines geistigen Zustands***, der eben nicht nur durch das repräsentierte Objekt oder den repräsentierten Sachverhalt definiert ist, sondern auch durch die Vernetzung mit anderen geistigen Zuständen, als ***intrinsische Bedeutung*** oder einfach nur als ***Bedeutung*** bezeichnen.

5.2. Einschub: Seltsame Vermutungen

In letzter Zeit bin ich bei meinen Recherchen so oft auf zwei falsche Hypothesen gestoßen, dass ich kurz darauf eingehen möchte. Ich sehe diesen Abschnitt aber nicht als Teil der ernsthaften Auseinandersetzung mit dem Thema "Qualia", sondern eher als unterhaltsamen Einschub, zu dem ich von einem Artikel verführt wurde, der sich auf diese beiden Hypothesen bezog und den ich so erheiternd fand, dass mich die Lust überkam, mein Vergnügen mit jemandem zu teilen. (Ich komme gleich darauf zurück.)

Die erste Hypothese lautet folgendermaßen:

Wenn bei einem Tier eine neuronale Struktur vorhanden ist, die einer Struktur ähnelt, ohne die Menschen keine Emotionen empfinden können, dann besteht Grund zur Annahme, dass auch dieses Tier zu Emotionen fähig ist.

Der Fehler ist offensichtlich: eine notwendige Bedingung wird mit einer hinreichenden Bedingung verwechselt.

Konkret geht es um Zwischenhirnstrukturen, vor allem um die Amygdala, von der es heißt, sie Sorge für die "emotionale Färbung" neuronaler Prozesse – und das ist nicht etwa als Redensart gemeint, sondern als ontologische Behauptung, so als wären Emotionen bunt und die Amygdala ein Farbtopf.

Was ist die Amygdala? Ein Nervenknotten. Würde man ernsthaft fragen, was solche Strukturen wirklich leisten können, dann käme die Unzulänglichkeit dieser Malkastenvorstellung augenblicklich zum Vorschein:

Was können Nervenknotten leisten? Dreierlei. Sie können

1. Areale des Gehirns miteinander oder mit anderen physiologischen Strukturen (z.B. Gliedmaßen) verknüpfen,
2. Information intern verarbeiten und auf diese Weise Input-Output-Relationen herstellen,
3. das neuronale Netz mit einem der chemischen Regulationssysteme des Organismus verbinden.

Die ersten beiden Punkte illustrieren das, was einleitend festgestellt wurde: Sie führen offensichtlich nicht aus dem Kreis der Information hinaus, und das Auftreten von Emotionen bleibt ungeklärt.

Kommen wir also zum dritten Punkt. Was ist mit chemischen Regulationssystemen gemeint? Z.B. das Transmittersystem, oder das endokrine System. Durch die Verbindung des neuronalen Netzes mit diesen Systemen erfahren neuronale Aktivitäten – und damit zugleich die dadurch repräsentierten Objekte oder Situationen – eine *Bewertung*, in dem Sinn, dass dadurch bestimmte Verhaltensweisen ausgelöst oder modifiziert werden, die sich grundsätzlich in zwei Klassen unterteilen lassen: Anstreben oder Vermeiden.

Können neuronale Strukturen, die Information und Verhalten auf diese Art miteinander verbinden, das Auftreten von Emotionen verursachen? Nimmt man das an, wird man sofort zu absurden Konsequenzen geführt. Solche Strukturen finden sich nämlich in *jedem* Organismus, also auch im einfachsten, weil sie für das Überleben notwendig sind. Jedes Tier muss unterscheiden können, was für sein Überleben günstig ist oder ungünstig; es muss Gefahren ausweichen und Nahrung aufsuchen.

Auch ein Fadenwurm hat ein sogenanntes "Belohnungssystem", und gemäß der erwähnten Hypothese müssten wir nun vermuten, er wäre zu Emotionen fähig. Angesichts seiner äußerst geringen Zahl von Neuronen ist das aber wenig plausibel.

Falls aber irgendjemand meint, es wäre doch möglich, könnte man ihm/ihr folgendes Gedankenexperiment präsentieren:

Nehmen wir an, wir konstruieren einen Roboter – nennen wir ihn *Susi* – der sich wie folgt verhält:

Wenn Susi zart berührt wird, dann gibt sie Wohllaute von sich und rollt näher heran. Übt man dagegen einen festeren Druck aus – so stark, dass er einem Menschen schon unangenehm wäre –, dann verharrt sie auf der Stelle und sagt "au". Drückt man schließlich noch fester zu oder schlägt Susi, dann jammert sie und rollt schreiend davon.

Es ist klar, dass Susis Verhalten mit einfachsten Mitteln zu erzeugen ist. Ein optischer Sensor, ein Drucksensor sowie etwas Elektronik und Mechanik wären ausreichend. Susi hat also mit Sicherheit keine Empfindungen. Aber sie hat eine Struktur, die Information mit Verhalten verknüpft, so dass Annäherung oder Flucht ausgelöst wird. Mit anderen Worten: Susi hat ein Belohnungssystem!

Man könnte also die Äußerung von Victoria Braithwaite, die die Frage untersucht, ob Fische Schmerzen fühlen²: "Auch wenn die entsprechenden Strukturen und Funktionen sehr viel einfacher sind als im limbischen System des Menschen, ist die Entdeckung des Fischäquivalents eine äußerst wichtige Erkenntnis" durch den Satz paraphrasieren: "Auch wenn die entsprechenden Strukturen und Funktionen sehr viel einfacher sind als beim Menschen, ist die Entdeckung des Susiäquivalents eine äußerst wichtige Erkenntnis."

Die erste der beiden Hypothesen, die ich hier kritisiere, ist also sicher falsch. Eigentlich war es überflüssig, das so ausführlich zu erläutern, weil ja schon vorher klar war, dass auf Basis der gegenwärtigen naturwissenschaftlichen Interpretation von Geist die Existenz bestimmter neuronaler Strukturen eben nur eine notwendige und keine hinreichende Bedingung für das Auftreten von Empfindungen ist. Dass bei den Versuchen, Qualia zu erklären, dieser wichtige Unterschied oft zu wenig beachtet wird, bezeugt im Grunde nur den Erklärungsnotstand der Naturwissenschaft, der allerdings im derzeitigen naturwissenschaftlichen Weltbild unvermeidlich ist.

Nun zur zweiten Hypothese. Sie lautet:

Wenn bei einem Tier Verhalten zu beobachten ist, das bei einem Menschen auf eine Empfindung hinweist, dann ist das ein Grund für die Annahme, dass auch dieses Tier etwas empfindet.

Um diese Hypothese zu widerlegen, genügt es, an den Roboter Hans zu erinnern, der sich so verhält, *als ob* er die Empfindung *rot* hätte, oder an Susi, die sich so verhält, *als ob* sie sich wohl fühlen oder Schmerz empfinden würde.

² Zitiert nach Spiegel online, <http://www.spiegel.de/spiegel/0,1518,749108,00.html>

"Der Tierverhaltensforscher Robert Elwood von der Queen's University in Belfast beträufelte die empfindlichen Fühler von Steingarnelen mit Essigsäurelösung. Bis zu fünf Minuten lang rieben die Krustentiere die misshandelten Körperteile – Elwood zufolge eine Reaktion, die an das Schmerzverhalten von Säugetieren erinnere."³

Ja, das ist wahr, genauso wie das Verhalten von Hans und Susi an das Verhalten von Kindern erinnert. Falls Elwood aber meint, seine Beobachtung sei ein Argument dafür, dass Krustentiere etwas empfinden – und das meint er wohl, denn wozu hätte er sich sonst in dieser Weise geäußert – dann muss man ihn darüber aufklären, dass aus seiner Beobachtung ausschließlich hervorgeht, dass sich Steingarnelen ihre Fühler reiben, wenn diese mit Essigsäure benetzt worden sind, und sonst nichts.

Es ist schon erstaunlich, dass solche Äußerungen überhaupt Beachtung finden. Aus dem Reiben der Fühler lässt sich doch ganz offensichtlich nichts anderes schließen, als dass ein bestimmter Reiz ein zugehöriges Verhalten auslöst, das auch länger andauern kann, als der Reiz besteht. Alles weitere ist keine Schlussfolgerung, sondern der Ausdruck subjektiver Neigungen.

Nur um nicht missverstanden zu werden: Wenn eine Person – nennen wir sie Robert – ihre Garnele lieber umarmt statt sie zu verspeisen, oder wenn eine andere Person – nennen wir sie Victoria – ihr Glück mit einem Fadenwurm teilen möchte, dann wäre ich der Letzte, der diese Beiden daran hindern wollte; – und damit ist es mir wirklich ernst, denn ich halte die Buntheit und Vielfalt im Garten des Menschlichen für einen hohen Wert, und ich fände es bedauerlich, wenn sein Boden durch unseren Rationalisierungs- und Optimierungswahn so sehr vergiftet würde, dass darauf nur noch die Gier nach Geld, Macht und Lust gedeihen könnte.

Aber ich würde Robert und Victoria vorschlagen, doch einfach ihren Neigungen zu folgen, und nicht zu versuchen, sie auf scheinbar rationale Weise zu begründen.

Zuletzt, als heiterer Abschluss, noch eine Äußerung, die vom Autor des zitierten Artikels selbst stammt:

"Das Fehlen der Großhirnrinde allein scheint vielen Experten [...] inzwischen nicht mehr ausreichend zu sein, um bewusste Empfindungen ausschließen zu können. Zweifel an der alten Lehrmeinung nähren nicht zuletzt verblüffende medizinische Fallgeschichten: Gelegentlich berichten Neurologen von Menschen mit nur halbem Großhirn. Wo bei anderen graue Zellen schwatzen, schwappt bei ihnen nur Nervenwasser – und dennoch sind sie nicht selten hochintelligent und sozial unauffällig."

³ Aus demselben Artikel.

Ja, und wenn man einen dreibeinigen Hund laufen sieht, dann fragt man sich doch unwillkürlich, ob die alte Lehrmeinung, dass Hunde zum Laufen Beine brauchen, nicht längst revidiert werden müsste!

5.3. Warum Qualia in keiner Beschreibung enthalten sind

Noch einmal: das Problem

Das Versagen der gegenwärtigen naturwissenschaftlichen Sicht des Zusammenhangs zwischen neuronalem Netz und Geist offenbart sich in der Gegenüberstellung der folgenden beiden Sachverhalte. Beide sind so weit abgesichert, dass sie den Status von Fakten beanspruchen können.

1. Gleichgültig, auf welche Weise man das Gehirn oder einen Bereich des Gehirns beschreibt oder analysiert, man wird stets nur Informations- oder Repräsentationszustände finden und niemals ein *Quale*.
2. Das neuronale System bringt *Qualia* hervor. Seine Zustände *sind* Qualia.

Das *Ding*, das beobachtet und beschrieben werden kann, ist ein *dynamisches neuronales Erregungsmuster*. Es kann eingesehen werden, dass dieses Muster etwas *repräsentiert*. Aber das, was dieses neuronale Muster *eigentlich ist* – das *Quale* – scheint die Dinglichkeit des Beschriebenen auf unerklärliche Weise zu transzendieren.

Woher stammt diese unaufhebbare Differenz zwischen dem, was es *ist*, und dem, als was wir es *beschreiben*?

Ich denke, durch die letzte Formulierung ist klar geworden, wo die Erklärung zu finden ist: In der Analyse des Verhältnisses von *wirklich existierenden Dingen* und *Dingen in einer Beschreibung*, die im Abschnitt 1.3 des zweiten Teils des Buchs [*Der Begriff der Wirklichkeit*](#) durchgeführt wurde. Der nun folgende Unterabschnitt enthält eine kurze Wiederholung mit einigen Ergänzungen.

Substanz und Akzidens; das Wesen des Seienden

Die Wirklichkeit besteht *für uns* aus *Dingen*, die *Eigenschaften* haben. Was bleibt übrig, wenn man sich von einem Ding alle Eigenschaften entfernt denkt? Offensichtlich *nichts*, da ein Ding ohne Eigenschaften nicht existiert.

Nichts kann aber keine Eigenschaften haben. Was müssen wir aus diesem Widerspruch schließen? Dass das wirklich existierende Ding, das *Ding an sich*, nicht dem Ding in der Beschreibung, dem *Ding für uns* entspricht.

Worin besteht der Unterschied?

Ein wirklich existierendes Objekt kann nicht verschwinden, wenn ihm alle Eigenschaften genommen werden – eben deshalb, weil *nichts* keine Eigenschaften haben kann –, während ein Objekt, das nur Element einer Beschreibung ist, tatsächlich vollständig verschwindet, wenn alle Eigenschaften entfernt werden.

Dieser Unterschied lässt sich folgendermaßen ausdrücken:

Wirklich existierende Objekte bestehen aus *Substanz und Akzidenzien*, Objekte in einer Beschreibung bestehen *nur aus Akzidenzien*.

Jede Beschreibung handelt ausschließlich von Akzidenzien. Eine Beschreibung ist immer ein Netz von Beziehungen zwischen Objekten, die durch diese Beziehungen definiert sind. Die Objekte bestehen ausschließlich aus dieser Definition, sie *sind nichts als diese Definition*, und deshalb verschwinden sie vollständig, wenn die Eigenschaften entfernt werden, die ja die Voraussetzungen der Beziehungen sind.

Die Substanz kann also in der Beschreibung nicht auftreten. Sie wird durch die Annahme der Existenz des Beschriebenen bloß vorausgesetzt. Dieser Mangel der Beschreibungen ist unaufhebbar. Er bildet eine unüberwindliche Grenze unseres Denkens.

Der Unterschied zwischen Wirklichkeit und Beschreibung kann aber noch auf eine andere Art bestimmt werden:

Es gibt kein Objekt, das bloß existiert; Existieren bedeutet wechselwirken, und wechselwirken bedeutet aktiv sein. Ein Objekt, das nicht *aktiv* ist, existiert nicht.

In der Beschreibung kann zwar die *Art* der Aktivität eines Objekts durch dessen Eigenschaften wiedergegeben werden, aber *die Aktivität selbst* fehlt; Objekte einer Beschreibung sind stets *passiv*.

Das, was Beschreibungen fehlt, lässt sich also durch zwei Begriffe fassen: *Substanz* und *Aktivität*.

Diese beiden Begriffe sind untrennbar miteinander verbunden. *Vor* allem Seienden stellt ihre Verbindung genau das dar, was *für uns* der *Ursprung des Seienden* ist: die erste Substanz. Ich nenne sie AGENS. *Für uns* wird ihre metaphysische Qualität *Aktivität* zum ersten Akzidens "Aktivität". Beides zusammen ist Aktivität von AGENS.

Am Seienden selbst ist die Substanz nicht nur die Bedingung seiner Existenz, sondern zugleich dasjenige, wovon die metaphysische Qualität *Aktivität* des Seienden ausgeht, die die jeweiligen Akzidenzien antreibt – aber nur in der Wirklichkeit; in der Beschreibung verlieren die Akzidenzien diese Qualität.

Ich habe schon einige Male vom *Wesen* eines Seienden gesprochen. Ich hole nun die Definition dieses Begriffs nach:

Definition:

Das Wesen eines Seienden ist dieses Seiende als Einheit von Substanz und Akzidenzien. Es ist das "Ansich-Sein" des Seienden, soweit dieses zum Begriff werden kann.

Das Wesen des Seienden kann nicht gedacht werden. Es bleibt aber auch nicht verborgen. Es ist möglich, sich ihm gedanklich zu nähern. Ich will das am Beispiel des Wesens des physikalischen Seienden demonstrieren.

Die erste Substanz ist AGENS. Auch wenn AGENS nicht gedacht werden kann, wissen wir doch etwas darüber: wir erkennen, dass demjenigen, was nach dem Entfernen aller Eigenschaften von einem Ding übrig bleibt, zwar nicht mehr *Existenz* zuerkannt werden kann, dass es aber dennoch nicht einfach dem rein begrifflichen Nichts entspricht, weil das begriffliche Nichts keine Eigenschaften haben könnte. Deshalb muss AGENS vom begrifflichen Nichts unterschieden werden.

Diese Unterscheidung macht es nun wiederum möglich, den *ontologischen Status* der ersten Substanz zu bestimmen:

Die erste Substanz kann nicht *Nichts* sein. Da ihr alle Akzidenzien fehlen, kann sie aber auch nicht *Etwas* sein. Somit ist ihr ontologischer Status weder der des *Seins* noch der des *Nicht-Seins*, sondern der der *Notwendigkeit* – einfach deshalb, weil das Weder-Etwas-noch-Nichts-Sein keine Alternative hat. Würden wir die erste Substanz als nicht vorhanden denken, dann hätten wir als ihren ontologischen Status *Nichts* gewählt, d.h. eine der beiden Alternativen, die wir zuvor ausgeschlossen haben.

Schon bei diesen ersten Schritten wird ein wenig klarer, was mit "gedanklicher Annäherung" an das eigentlich Undenkbare gemeint ist.

Wiederholen wir noch einige weitere Schritte der Ableitung, die in den Anfangskapiteln des zweiten und dritten Teils durchgeführt wurde:

Aktivität bedeutet Veränderung. Um jedoch eine Aussage zu erhalten, aus der die Beschreibung der Wirklichkeit folgen kann, werden zwei verschiedene Änderungen benötigt. Da der Raum eine notwendige Bedingung der Wirklichkeit ist, und weil es ohne Bewegung keine Änderung gibt, muss die eine Änderung zur Änderung des Raums werden und die andere zur Änderung der Bewegung.

Da es keine Größe gibt, können die Änderungen nur relativ sein und sich auf kein absolutes Maß beziehen. Da es kein Gedächtnis gibt, muss sich die zeitliche Änderung der Bewegung auf den differenziell benachbarten Augenblick in der Vergangenheit beziehen, d.h. sie muss als Differenzialquotient dv/dt ausgedrückt werden.

Auf diese Weise gelangt man zur Vorstellung eines metrischen Kontinuums, das aus beschleunigten Flüssen besteht, wobei die Beschleunigung von der Änderung der relativen räumlichen Dichte (der Länge oder des Winkels) abhängt.

Das bedeutet, dass alles Seiende ein Muster aus Veränderungen der Bewegung des Kontinuums ist, und da alles Seiende eine gewisse Zeitspanne hindurch seine Form annähernd beibehält, bedeutet es zugleich, dass alles Seiende entweder ein Attraktor der lokalen Kontinuumsdynamik ist oder ein Teil eines Attraktors der globalen Kontinuumsdynamik.

Damit ist eine weitere Annäherung an das Wesen des physikalischen Seienden erfolgt. Die erste Substanz AGENS ist durch die soeben durchgeführten Schlussfolgerungen zu einem Raum-Zeit-Kontinuum geworden, und das erste Prädikat *Veränderung* hat die Gestalt von Gleichung (1) angenommen, die das Gesetz dieses Kontinuums darstellt und die das ausdrückt, was die physikalische Wirklichkeit *für uns* ist: ein Gewebe aus differenziellen raumzeitlichen Änderungen.

Verweilen wir ein noch ein wenig beim Begriff des Kontinuums, denn an ihm lässt sich besonders klar demonstrieren, in welchem Maß und bis zu welcher Grenze der direkt nicht denkbare Begriff der *Substanz* sich unserem Denken indirekt erschließt.

Was ist ein Kontinuum? Betrachten wir der Einfachheit halber ein eindimensionales Kontinuum. In der Standardanalysis stehen die Punkte dieses Kontinuums in einer umkehrbar eindeutigen Beziehung zu den reellen Zahlen. Damit wird suggeriert, dass Punkte *existieren*. Punkte haben aber keine

Existenz! Wenn wir also sagen: im Kontinuum gibt es einen beschleunigten Fluss, dann ist diese Aussage nur mathematisch sinnvoll; Wird sie aber ontologisch aufgefasst, dann ist man doch augenblicklich wieder mit der Tatsache konfrontiert, dass sich eigentlich *nichts* bewegt.

Eine andere Frage ist die, welche Zahlen nun tatsächlich den Punkten des Kontinuums zugeordnet werden können. Es mag zunächst überraschend erscheinen, dass es nicht die rationalen Zahlen sind, da diese doch schon *dicht* liegen, d.h. dass zwischen zwei rationalen Zahlen, gleichgültig, wie nah sie aneinander sind, immer unendlich viele andere rationale Zahlen liegen. Dennoch gibt es, wie sich leicht nachweisen lässt, immer noch Platz für überabzählbar viele weitere, nämlich die irrationalen Zahlen.

Und dann?

Lange Zeit meinten die Mathematiker, dass dann die Zahlen vollständig wären, und dass, wenn sie als Punkte einer Geraden gedacht werden, alle möglichen Punkte der Geraden erfasst seien. In der Nicht-Standard-Analysis wird aber gezeigt, dass sich weitere Zahlen definieren lassen, die wiederum *zwischen* den bisherigen liegen. Dieser Prozess lässt sich *ad infinitum* fortsetzen. Es ist allein unsere Entscheidung, wann wir abbrechen.

Das bedeutet, dass das Kontinuum ein Grenzbegriff ist. Letztlich können wir uns keine Vorstellung davon machen: wenn wir es unterteilen, dann verlieren wir es – aber wir *müssen* es unterteilen, um es zu beschreiben.

Trotzdem ist nicht zu leugnen, dass die mathematische Formulierung des *beschleunigten Kontinuumsflusses* mehr an Information enthält als der sprachliche Ausdruck *Veränderung von AGENS*. Und genau diese Annäherung, diese Vermehrung an Wissen, die mit dem Nachdenken über das Wesen des Kontinuums verbunden ist, ermöglicht dann weitere Schlussfolgerungen über die daraus entstehende Wirklichkeit.

Wenn wir das Kontinuum erforschen, dann erfahren wir mehr über die Operationen, die in ihm möglich sind. Falls es unendlich viele sind, werden wir sie zwar niemals alle kennen, aber es gibt auch keinen Grund für die Annahme, dass uns eines seiner Akzidenzien aus prinzipiellen Gründen verborgen bleibt. Wir können mit einiger Sicherheit ausschließen, dass die erste Substanz irgendein Geheimnis birgt, das sich unserem Verständnis für immer entzieht.

Damit ist das Verhältnis zwischen unserem Denken und dem Begriff des Wesens des Seienden bestimmt:

Das Wesen des Seienden kann *als solches* nicht gedacht werden, da sich weder die Substanz selbst noch die Einheit von Substanz und Akzidens – das Kontinuum, das nur als sich Veränderndes existiert – denken lässt.

Andererseits sind aber die Akzidenzien als *Äußerungen* der Substanz aufzufassen, und es gilt: Was auch immer das Wesen des Seienden an *Aktivität* enthält, das *muss* sich als Akzidens äußern, denn andernfalls *wäre* es nicht. Deshalb kann behauptet werden, dass sich uns das Wesen des Seienden, auch wenn es nicht gedacht werden kann, doch durch seine Akzidenzien erschließt.

Das Wesen des Seienden enthält also keine Geheimnisse. Alles, was an ihm ist, offenbart sich uns durch seine Akzidenzien. Es ist nicht denkbar, aber auch nicht verborgen.

Da alles Seiende ein Muster aus Veränderungen der Raumzeit ist, dem *Ursprung des Seienden*, erhebt sich die Frage, ob es auch dessen ontologischen Status *Notwendigkeit* und dessen metaphysische Qualität *Aktivität* enthält.

Notwendigkeit wird offensichtlich nicht vom *Ursprung des Seienden* auf das Seiende übertragen: als Muster kann es sich auflösen. Alles Seiende steht in der Alternative *Sein oder Nicht-Sein*. Notwendig ist nur, *dass* etwas existiert. *Was* existiert, ist jedoch nicht notwendig.

Aktivität jedoch, die metaphysische Qualität, überträgt sich vom *Ursprung des Seienden* auf alles Seiende. Sie äußert sich in Gestalt der jeweiligen Akzidenzien; Sie ist das Treibende hinter den Akzidenzien der wirklich existierenden Dinge, das, was die Akzidenzien der wirklich existierenden Dinge von den Akzidenzien der Dinge in einer Beschreibung unterscheidet.

Die Antwort

Wir nehmen den Tisch wahr, und wir meinen, wir wüssten, was er ist. Wenn wir jedoch versuchen, die Frage, was er ist, bis an ihren Grund zu verfolgen, dann verschwindet das, was er ist. Genau dieses *Verschwinden der Substanz* hat sich in der Physik des zwanzigsten Jahrhunderts ereignet, und es hat sie in eine Krise gestürzt, von der sie zwar kaum noch ein Bewusstsein hat, von der sie sich aber bis heute nicht erholt hat, weil es ihr nicht gelungen ist, sich davon einen Begriff zu machen.

Wir nehmen die Wellen wahr, die sich zum Ufer hin bewegen, und wir meinen, wir wüssten, warum sie sich bewegen. Doch der Begriff der Energie ist für eine Erklärung ungeeignet, weil er bloß mathematisch definiert ist und keine Unterscheidung zwischen Wirklichkeit und Beschreibung

ermöglicht. Auf die Frage, warum sich überhaupt irgendetwas bewegt, gibt es erst dann eine Antwort, wenn das Wesen der Wirklichkeit als *Aktivität* bestimmt worden und zugleich klar geworden ist, dass dieses Wesen beim Beschreiben der Wirklichkeit verloren geht.

Das bedeutet: Seiendes ist immer *mehr* als sein Begriff, mehr als das, als was es in unserer Wahrnehmung oder in unseren Beschreibungen erscheint. Es ist immer *Substanz und Akzidens*, und in der Wahrnehmung und in Beschreibungen treten *nur die Akzidenzien* auf.

Das, was Seiendes eigentlich ist, sein *Wesen*, die Einheit von Substanz und Akzidens, ist für uns undenkbar. Wir wissen also bei *keinem* Seienden, was es ist. Nur die Verwechslung von Wirklichkeit und Beschreibung, die uns bei materiellen Dingen so sehr zur Gewohnheit geworden ist, dass wir sie gar nicht mehr bemerken, vermittelt uns den Eindruck, wir wüssten es.

Der Tisch als wirklich existierender ist *mehr* als der Tisch, den wir wahrnehmen oder denken. Die Welle als wirklich existierende ist *mehr* als die Welle, die wir beschreiben.

Und damit sind wir nun endlich bei der Antwort auf die Frage, warum ein *existierendes* neuronales Erregungsmuster *mehr* ist als ein *gedachtes* neuronales Erregungsmuster.

Es ist nach dem soeben Gesagten weder erstaunlich noch unverständlich, dass Zustände des neuronalen Netzes nicht bloß das sind, als was sie in unserer Beschreibung erscheinen, sondern selbstverständlich und notwendig. Sie *müssen* mehr als das sein, weil ihre Beschreibung nur den Teil des Zustands enthalten kann, der durch eine Beschreibung erfassbar ist, also das *Akzidens*.

Was ist in diesem Fall das Akzidens? Das, was begrifflich erfasst werden kann, also die *Definition* des betreffenden Objekts, und die besteht im Fall eines neuronalen Musters aus der Information, die darin enthalten ist, also aus seiner *intrinsic* Bedeutung.

Damit sind wir zu folgender Erkenntnis gelangt:

Geistige Zustände haben einen Informationsgehalt und einen Empfindungsgehalt. Der Informationsgehalt, die *Bedeutung*, ist ihr *Akzidens*.

Wie alles, was existiert, besteht aber auch ein geistiger Zustand nicht nur aus *Akzidenzien*, sondern auch aus *Substanz*. Er ist also *mehr* als nur Bedeutung.

Was fehlt in der Beschreibung von diesem "Mehr-Sein", was fehlt vom ganzen Sein, vom *Wesen* des geistigen Zustands? EMPFINDUNG. Also ist EMPFINDUNG die *Substanz* geistiger Zustände.

Satz

Geistige Zustände sind Qualia. Bedeutung ist ihr Akzidens, EMPFINDUNG ist ihre Substanz. Als solche ist sie in keiner Beschreibung enthalten.

Das Wesen des Quale ist die Einheit von EMPFINDUNG und Bedeutung.

Damit haben wir den ersten Teil des Problems der Qualia gelöst: die Frage, warum neuronale Muster als *existierende* Entitäten **mehr** sind als die Objekte, als die sie in unseren Beschreibungen auftreten, und warum dieses "Mehr-Sein", das *Quale*, in keiner Beschreibung enthalten ist.

Wenden wir uns nun dem zweiten Teil des Problems zu: der Frage, warum die Substanz im Fall der Qualia nicht mehr dieselbe ist wie bei anderen Arten des Seienden.

5.4. Die Verwandlung des Seienden vom materiellen Ding zum Quale

Erste und zweite Substanz

Es ist nun klar geworden, warum wir RAUMZEIT als *erste* Substanz bezeichnet haben, und Veränderung als *erstes* Akzidens: Offenbar ist *Geist* ein Bereich der Wirklichkeit, in dem das Wesen des Seienden sich gewandelt hat: Seiendes besteht hier nicht nur aus anderen Akzidenzien, sondern auch aus einer anderen Substanz.

Ich bezeichne EMPFINDUNG, die Substanz der geistigen Zustände als ***zweite Substanz***, und ***intrinsische Bedeutung***, ihr Akzidens, als ***zweites Akzidens***.

Die zweite Substanz EMPFINDUNG ist aber nicht etwa als unabhängig von der ersten Substanz RAUMZEIT oder dieser gegenübergestellt zu denken: die zweite Substanz geht aus der ersten Substanz hervor. Wie dieses Hervorgehen des Geistes aus dem neuronalen Netz, diese Verwandlung des physikalischen Seienden zum Quale sich ereignet, ist Gegenstand der nun folgenden Untersuchung.

Klären wir zunächst die Frage, was von einer solchen Untersuchung zu erwarten ist.

Es ist zu fordern, dass erklärt wird, *warum* eine Verwandlung der Substanz stattfindet. Es kann aber nicht gefordert werden, dass die Erklärung eine Ableitung des Wesens der geistigen Zustände enthält;

Das ist jedoch eine selbstverständliche Einschränkung, weil etwas, was in keiner Beschreibung enthalten ist, auch nicht abgeleitet werden kann.

Allerdings stellt diese Einschränkung keine absolute Grenze dar, weil es – wie beim physikalischen Seienden – möglich ist, sich dem Wesen der geistigen Zustände gedanklich anzunähern, auch wenn es nicht gedacht werden kann; – Oder sagen wir besser: nicht *beschrieben* werden kann, denn wir *wissen* ja, was Qualia sind. Genau genommen sind Qualia sogar das einzige Seiende, von dem wir wissen, was es ist, weil unser Bewusstsein ein unaufhörlicher Strom von Qualia ist. Nicht nur, was wir fühlen, auch was wir denken und wahrnehmen ist ein Quale. Auch die abstrakteste geistige Tätigkeit wird von einem Interesse getragen und durch ein Motiv geleitet, und sowohl Interesse als auch Motiv sind Abkömmlinge von Empfindungen, von denen sie nicht getrennt werden können.

Wie also geht die zweite Substanz aus der ersten hervor? Warum verwandelt sich das Wesen des Seienden?

Der Grund für die Verwandlung

Wir können von zwei Voraussetzungen ausgehen:

(1) Die Verwandlung des Wesens des Seienden vom materiellen Objekt zum Quale ereignet sich bei der Entfaltung der Natur zu Schichten ansteigender Komplexität, deren Gesetzmäßigkeiten im vorigen Kapitel beschrieben wurden.

(2) Der uns begrifflich zugängliche Teil dieses Aufstiegs ist jener Teil, der sich auf der Seite der Akzidenzien ereignet. Hier muss also die Argumentation stattfinden.

Zunächst eine Bezeichnung: Das Wesen des Seienden in demjenigen Bereich der Wirklichkeit, in dem es keinen Geist gibt, werde ich mit dem Begriff *Materie* bezeichnen. (Dieser Begriff von Materie enthält somit nicht – wie in der Physik üblich – ausschließlich die Akzidenzien, sondern auch die Substanz.)

Der Punkt, auf den wir unsere Aufmerksamkeit richten müssen, ist die Frage:

Warum verwandelt sich das Wesen des Seienden erst mit der Herausbildung neuronaler Netze hoher Komplexität, die Geist hervorbringen, während es bis dahin durchgehend einheitlich zu sein scheint –

jedenfalls so weit, dass wir uns erst an diesem Punkt zur Einführung einer zweiten Substanz genötigt sehen?

Es gelingt uns mit einigem Erfolg, die Phänomene, die wir in der materiellen Natur vorfinden, zu beschreiben. Dort, wo unser Wissen unvollständig ist, wie etwa bei der Entstehung von Leben, kann diese Lücke zumindest durch wissenschaftliche Hypothesen ausgefüllt werden. Die bei der Beschreibung der Natur auftretenden Probleme erscheinen meist als *technische* Schwierigkeiten, und niemals treffen wir auf ein Phänomen, das sich unserem Verständnis *prinzipiell* zu entziehen scheint – aber das gilt eben nur bis zu dem Punkt der Entfaltung der Natur, an dem neuronale Netze hoher Komplexität erscheinen.

Mit deren Erscheinen tritt ein Phänomen auf, das sich nicht nur der wissenschaftlichen, sondern überhaupt *jeder* Beschreibung entzieht: *Empfindung*.

Beginnen wir nun mit dem Gedankengang, der uns schließlich die *Notwendigkeit der Verwandlung* der ersten in die zweite Substanz einsehen lässt.

Was geschieht, wenn sich bei der Entfaltung der Natur zu Entitäten ansteigender Komplexität die Substanz verwandelt?

Die Antwort folgt aus der Definition der Substanz: Substanz ist ein Teil des Wesens des Seienden, und sie ist dasjenige, was aus Beschreibungen ausgeschlossen bleibt.

Wenn sich also die Substanz verwandelt, dann tritt uns Seiendes entgegen, dessen Wesen auf eine *neue* Art unverständlich ist, eine Art, die uns anders erscheint als die des Wesens des Seienden, dessen Substanz gleich bleibt.

Das klingt allerdings eigenartig und unbestimmt: Wie kann sich die Verwandlung von etwas, was in Beschreibungen nicht auftritt, äußern? Ist es sinnvoll, von verschiedenen Arten des Unbeschreibbaren zu reden?

Das, was geschieht, wenn sich das Wesen des Seienden ändert, und wie sich diese Änderung bemerkbar macht, lässt sich an Hand jenes Falls verdeutlichen, wo sie sich tatsächlich vollzieht: beim Übergang von Materie zu Geist.

Wie zuvor ausgeführt, können wir weder die Substanz des materiellen Seienden – das Raum-Zeit-Kontinuum – denken, noch sein Wesen, d.h. Materie als in der Zeit sich erhaltendes Muster aus raumzeitlichen Änderungen.

Aber unser Nicht-Wissen ist ein *bestimmtes* Nicht-Wissen. Wir haben das, was sich unserem Denken entzieht, bis an seine Grenze verfolgt und uns dadurch das Kontinuum so weit wie möglich erschlossen: Durch die Begriffe *Grenzübergang* und *Grenzwert* haben wir operational verfügbar gemacht, was wir nicht denken können.

Die methodischen Annäherungen an das eigentlich Unverständliche, bei denen wir so erfolgreich waren, dass wir schon beinahe vergessen konnten, dass es letztlich doch *da* ist, versagen jedoch vollkommen, wenn wir versuchen, sie auf das anzuwenden, was an demjenigen Seiendem unerklärbar ist, das unter den Begriff *Geist* fällt. Wir kommen weder der Substanz geistiger Entitäten – EMPFINDUNG –, noch ihrem Wesen – der Einheit von EMPFINDUNG und Bedeutung, dem Quale – irgendwie näher, wenn wir sie als Muster aus raum-zeitlichen Änderungen auffassen.

Diese Aussage muss allerdings präzisiert werden: Selbstverständlich *sind* geistige Entitäten Muster aus raum-zeitlichen Änderungen – *alles* Seiende ist ja ein solches Muster – aber der Ausdruck "sind" ist in dieser Aussage – anders als im Fall von Materie – nicht mehr im Sinn einer Wesensbestimmung zu verstehen: Das Wesen materieller Dinge ist dadurch bestimmt, dass sie raum-zeitliche Muster sind, aber das Wesen geistiger Dinge besteht darin, dass sie Qualia sind, und nicht darin, dass sie raum-zeitliche Muster sind.

Es kann also behauptet werden:

Das Auftreten einer zweiten Substanz zeigt sich uns durch Phänomene, deren Wesen für uns auf eine neue Art unerklärbar ist, so dass die bewährten gedanklichen Mittel der Annäherung an das Unerklärbare versagen.

Die "Mittel der Annäherung" sind aber nichts anderes als die mit der Substanz verbundenen Akzidenzien; Die Substanz äußert sich in ihren Akzidenzien, und sie erschließt sich uns durch deren Erforschung.

Damit haben wir *eine* Richtung des Zusammenhangs zwischen der Verwandlung der Substanz und der Änderung der Akzidenzien bestimmt:

Wenn eine neue Substanz auftritt, dann ändern sich auch die Akzidenzien.

Hier noch eine Kurzform der soeben durchgeführten Schlussfolgerung:

Die Akzidenzien des Seienden sind Äußerungen seiner Substanz. Als solche ermöglichen sie uns eine gedankliche Annäherung an die eigentlich undenkbare Substanz. Wenn nun ein neues Seiendes mit

anderer Substanz auftritt, dann müssen die gedanklichen Annäherungen einem neuen Ziel zustreben, und das bedeutet, dass die Gedankengänge, die zur Erschließung der neuen Substanz benötigt werden, andere sein müssen als zuvor. Mit anderen Worten: die Akzidenzien müssen sich ändern.

Was ist mit der Umkehrung? Tritt eine neue Substanz auf, wenn sich die Akzidenzien ändern?

Es ist sofort zu sehen, dass das im Allgemeinen nicht gilt: Beim Aufstieg des Seienden treten in jeder neu entstehenden Schicht auch neue Akzidenzien auf, ohne dass sich die Substanz verwandelt. Eine Verwandlung der Substanz erfolgt erst bei der letzten, komplexesten Schicht der Wirklichkeit, bei Wesen, die über ein neuronales Netz verfügen, das hinreichend komplex ist, um Geist hervorzubringen.

Damit sind wir zur Bestimmung des logischen Ortes gelangt, an dem die Aufklärung der Frage nach der Verwandlung von Materie in Geist zu finden sein muss:

Die Antwort auf die Frage, warum sich das Wesen des Seienden ändert – derart, dass aus Materie Geist hervorgeht –, muss sich aus der Analyse des Unterschieds ergeben, der zwischen der Art der Änderung der Akzidenzien des Seienden infolge der Herausbildung komplexer neuronaler Netze und der Art der Änderung der Akzidenzien infolge der Entstehung anderer, einfacherer Schichten des Seienden besteht.

Worin besteht dieser Unterschied? Betrachten wir zunächst einige der Akzidenzien, die bei der Entstehung neuer Schichten des Seienden auftreten.

Beginnen wir mit einem Akzidens, das schon am einfachsten Seienden auftritt: *Gravitation*. Wenn *Masse* als dasjenige aufgefasst wird, *wovon* Gravitation ausgeht, dann ist Masse ein Teil der Substanz aller materiellen Objekte.

Im zweiten Teil ist gezeigt worden, dass die Gravitation eines Objekts direkt aus dem Gesetz des Kontinuums folgt, wenn der metrischen Dichte des Raums eine zusätzliche Bedingung auferlegt wird.

Das bedeutet aber nicht, dass irgendein Zustand eines realen physikalischen Systems der Art "materielle Körper mit ihrem Gravitationsfeld" aus einem Zustand des Universums *vor* der Entstehung von Teilchen, die Masse haben und von einem Gravitationsfeld umgeben sind, vollständig abgeleitet werden könnte.

Es gilt nur die wesentlich schwächere Behauptung, dass das Akzidens *Gravitation*, wenn es in idealisierter Form betrachtet wird – d.h. ohne jede Störung von außen und im vollständigen

dynamischen Gleichgewicht – als geordneter Zustand des longitudinalen Flusses erscheint und dass alle Auswirkungen, die dieser Fluss-Zustand auf andere Objekte hat, auf das Gesetz, aus dem er hervorgeht, zurückgeführt werden können.

In diesem Sinn ist also das Akzidens *Gravitation*, das in der ersten und einfachsten Schicht des Seienden auftritt, auf die darunter liegende Schicht – das Kontinuum mit seinem Gesetz – *reduzierbar*.

Auf der Seite der Substanz hat diese Reduzierbarkeit des Akzidens zur Folge, dass der Begriff "Masse" verlustfrei aus der Naturbeschreibung eliminiert werden kann. Es ist also nicht notwendig, Masse als neue Substanz, d.h. als neue, nicht denkbare metaphysische Wesenheit aufzufassen.

Betrachten wir nun ein Akzidens, das zu einer wesentlich komplexeren Schicht des Seienden gehört: als (willkürliches) Beispiel wählen wir die sogenannte *Gluconeogenese*, also die Bildung von Glukose in Zellen. Der Vorgang ist – wie die meisten Stoffwechselvorgänge – äußerst komplex und besteht aus einer ganzen Reihe von Einzelprozessen.

In unserem Zusammenhang ist jedoch nur Folgendes von Interesse: Zwar wäre es ausgeschlossen, die Gluconeogenese aus irgendwelchen Zuständen des Universums abzuleiten, die es *vor* der Existenz von Zellen gab, aber es lässt sich doch behaupten, dass alle Schritte, die für die Bildung von Glukose in Zellen durchlaufen werden müssen, als *biochemische Prozesse* beschrieben und verstanden werden können.

In dieser Hinsicht unterscheidet sich also die Gluconeogenese, die als Akzidens von Zellen erscheint, nicht von der Gravitation, dem Akzidens aller materiellen Objekte: Beide können durch Rückführung auf die jeweils darunter liegende Schicht des Seienden verstanden werden. Sie erscheinen als *Funktion* dieser Schicht.

Noch ein letztes Beispiel: Prozesse, die in neuronalen Netzen stattfinden, die nicht dafür geeignet sind, Geist hervorzubringen. (Dass solche neuronalen Netze existieren, kann vorausgesetzt werden. Man könnte z.B. das Verhalten der Roboter Hans und Susi durch neuronale Netze modellieren, und das wären dann neuronale Netze dieser Art.)

Die Akzidenzien solcher Prozesse lassen sich unter dem Begriff *Informationsverarbeitung* zusammenfassen. Informationsverarbeitung besteht aus Input-Output-Relationen. Sofern diese nicht intern – durch Vernetzung mit anderen solchen Relationen und die damit gegebene Rückkopplung – verändert werden, sondern immer in (annähernd) identischer Form bestehen bleiben, können sie als Funktionen der vorgegebenen Architektur des neuronalen Netzes sowie äußerer Bedingungen aufgefasst werden.

Das Verhalten von Tieren, die neuronale Netze dieser Art haben, lässt sich dann ebenfalls als Funktion dieser Architektur sowie äußerer Bedingungen beschreiben.

Auch in diesem letzten Beispiel gilt also: Die Analyse der Akzidenzien zeigt, dass sie als Funktionen der darunter liegenden, einfacheren Schichten des Seienden verstanden werden können.

Wir sind somit zu folgender Einsicht gelangt:

Bei allen evolutionären Übergängen zu neuen, komplexeren Schichten des Seienden – bis hin zu neuronalen Netzen, die keinen Geist hervorbringen – lassen sich die Akzidenzien, die in der jeweils neuen Schicht auftreten, als Funktionen von Akzidenzien einfacherer Schichten des Seienden beschreiben.

Das letzte Beispiel hat uns bereits nahe an das Reich des Geistes herangebracht. Unternehmen wir nun den letzten Schritt. Stellen wir die Frage:

Wodurch unterscheiden sich die Akzidenzien geistiger Zustände von den Akzidenzien des anderen Seienden? Folgt aus diesem Unterschied auch, dass sich – wie für die Erklärung der Verwandlung der Substanz gefordert –, beim Übergang vom Materiellen zum Geistigen die Akzidenzien auf eine andere Weise verändern als bei Übergängen, die im Bereich des Materiellen stattfinden?

Nach dem bisher Gesagten liegt die Antwort auf der Hand:

Betrachten wir ein neuronales Netz, das Geist hervorbringt. Als notwendige Bedingungen dafür, dass neuronale Muster, die etwas repräsentieren – wir haben sie als Attraktoren der neuronalen Dynamik bestimmt – zu geistigen Zuständen werden können, wurden bisher genannt:

- Die Existenz funktionell ungebundener Bereiche, deren Dynamik für die Strukturierung durch solche Attraktoren offen ist.
- Die Vernetzung der Attraktoren untereinander.

Repräsentationen können sich auf äußere Bedingungen oder auf Körperzustände beziehen. Aber auch der Informationsgehalt eines neuronalen Zustands selbst kann repräsentiert werden. Solche Metarepräsentationen sind z.B. dann erforderlich, wenn Handlungsalternativen gegeneinander abgewogen werden.

Der für unsere Fragen entscheidende Punkt ist folgender:

Aus dem Hebb'schen Gesetz folgt, dass die geistige Tätigkeit auf die neuronale Struktur zurückwirkt. Die geistige Tätigkeit verändert also ihre eigene neuronale Codierung. Sie verändert *sich selbst*.

Die Forderung, dass Repräsentations-Zustände untereinander vernetzt sind, ist gleichbedeutend mit dem Auftreten von Rückkopplungsschleifen: Zustand A beeinflusst Zustand B, Zustand B beeinflusst Zustand C, der wiederum auf Zustand A zurückwirkt usw. Einerseits werden durch solche Rückkopplungsschleifen bestehende Muster verstärkt, andererseits können dadurch aber auch Verbindungen zwischen Mustern entstehen, die zuvor nicht in Verbindung standen. Der Informationsgehalt der neuronalen Muster ändert sich damit: er wird in zunehmendem Maß von den *internen* Beziehungen zwischen den neuronalen Mustern bestimmt, während die ursprüngliche funktionelle Abhängigkeit in den Hintergrund tritt: die Repräsentations-Zustände entwickeln sich zu intrinsischen Bedeutungen.

Das trifft sogar auf Wahrnehmungen zu: auch wenn sie als Repräsentationen realer Objekte inhaltlich an diese gebunden bleiben – das neuronale Abbild eines zweimal in identischer Lage beobachteten Objekts in der primären Sehrinde wird in beiden Fällen nahezu identisch sein –, sind sie doch *als geistige Zustände* keineswegs auf diese Repräsentations-Funktion beschränkt. Wahrnehmung schließt all das ein, was sich an Informationsverarbeitung im zugehörigen kortikalen Feld zusätzlich zur Verarbeitung der reinen sinnlichen Information ereignet, und ein Halo von zugehörigen Assoziationen ist ebenfalls Teil von Wahrnehmungen.

Bei geistigen Zuständen, die nicht direkt an äußere Objekte gebunden sind, gibt es überhaupt keine prinzipielle Einschränkung für die Veränderungen, denen sie im Kreis der inneren Weiterverarbeitung unterworfen sind. Bei Gedankengängen sind stets Überraschungen möglich: neue Schlussfolgerungen ergeben sich, neue Begriffsbildungen werden notwendig, Irrtümer müssen korrigiert werden. Im Bereich der Phantasie ist die Veränderung bestehender oder die Schöpfung neuer intrinsischer Bedeutungen sogar das charakteristische Merkmal, und die Beziehung zu äußeren Gegenständen verblasst oder reißt gänzlich ab.

Daraus folgt das gesuchte Unterscheidungskriterium:

Intrinsische Bedeutung, das Akzidens geistiger Zustände, kann nicht als Funktion von Akzidenzien von einfacheren (atomaren, molekularen, biochemischen, neuronalen usw.) Schichten des Seienden verstanden werden.

Ein geistiger Zustand gewinnt seine Bedeutung durch seine Position im Netz der geistigen Zustände, d.h. im Netz der Bedeutungen. Auch wenn Wahrnehmungen und Urteile an die reale Welt gebunden

bleiben, ist es infolge der permanenten Veränderung durch Rückkopplungsschleifen unmöglich, irgendeine Art funktioneller Abhängigkeit der Bedeutung geistiger Zustände zu behaupten.

Kurz zuvor hatten wir festgestellt:

Bei allen evolutionären Übergängen zu neuen, komplexeren Schichten des Seienden – bis hin zu neuronalen Netzen, die keinen Geist hervorbringen – lassen sich die Akzidenzien, die in der jeweils neuen Schicht auftreten, als Funktionen von Akzidenzien einfacherer Schichten des Seienden beschreiben.

Dies ist also der Unterschied, der dafür verantwortlich ist, dass sich die Substanz des Seienden – und damit auch sein Wesen – erst beim Übergang von Materie zu Geist verwandelt, während sie bei den anderen Übergängen von einer Schicht zur nächsthöheren unverändert bleibt.

Liefert dieses Kriterium aber auch eine *Erklärung* dafür, *warum* sich die Substanz des Seienden beim Übergang von Materie zu Geist ändert? Ja. Folgendermaßen:

Substanz und Akzidens sind *untrennbar* verbunden: ein Ding ohne Eigenschaften – physikalisch ausgedrückt: ein Objekt, das mit nichts anderem wechselwirkt – existiert nicht.

Das *erste Akzidens* ist mit der *ersten Substanz* verbunden. Was lässt sich über komplexe Akzidenzien und die zu ihnen gehörende Substanz sagen?

Wenn komplexe Akzidenzien auf einfachere Akzidenzien reduzierbar sind, dann heißt das, dass sie zuletzt auch auf das erste und einfachste Akzidens zurückgeführt werden können. *Für uns* ist *Reduzierbarkeit* jedoch gleichbedeutend mit *ontologischer Identität*: Wenn B auf A reduzierbar ist, dann *ist* B eigentlich A. Wenn also ein komplexes Akzidens auf das erste Akzidens reduzierbar ist, dann *ist* es eigentlich das *erste Akzidens*, und dann ist es untrennbar mit der *ersten Substanz* verbunden.

Das bedeutet: Solange die Akzidenzien reduzierbar sind, bleibt die Substanz gleich. Ebenso gilt umgekehrt: Solange die Substanz gleich bleibt, sind die Akzidenzien reduzierbar.

Und daraus folgt: ***Wenn Akzidenzien auftreten, die nicht auf einfachere Akzidenzien zurückgeführt werden können, dann muss die zugehörige Substanz verändert sein.***

Mit der Bildung dieser Akzidenzien ist also zugleich eine Verwandlung der Substanz erfolgt.

Satz

Solange sich Akzidenzien höherer Komplexität als Funktionen von Akzidenzien geringerer Komplexität beschreiben lassen, bleibt die Substanz gleich. Wenn dieser funktionelle Zusammenhang verschwindet, dann ändert sich die Substanz. Für uns erscheint sie dann als neue, zweite Substanz.

Im Bereich der Materie ist die erste Bedingung erfüllt. Alles materielle Seiende kann als Attraktor der Dynamik von AGENS, als Muster raumzeitlicher Veränderungen aufgefasst werden. Dies ist das Wesen das materiellen Seienden.

Im Bereich des Geistes gilt die zweite Bedingung. Intrinsische Bedeutungen, die Akzidenzien geistiger Zustände, können auf keine Weise als Funktionen von Akzidenzien des zugrunde liegenden Seienden verstanden werden. Das Wesen des geistigen Seienden ist daher von dem des materiellen Seienden verschieden.

Damit ist die Frage vollständig geklärt, warum es bei der Entstehung von Geist aus Materie zugleich zu einer Verwandlung des Wesens des Seienden kommt.

Bemerkungen

1. Der Kern der Argumentation, mit der die Verwandlung des Wesens des Seienden begründet wird, ist die Unterbrechung der Verbindung zwischen den geistigen Akzidenzien und dem ersten Akzidens. Damit zerreißt zugleich die Verbindung der geistigen Akzidenzien mit der ersten Substanz, wodurch die Annahme einer zweiten Substanz erzwungen wird.

Ich habe hier vorausgesetzt, dass "Verbindung" mit "Reduzierbarkeit" (des Akzidens, nicht des konkreten Systemzustands) gleichzusetzen ist. Es wäre aber auch möglich, die Annahme der Reduzierbarkeit durch die schwächere Annahme der *Erklärbarkeit* zu ersetzen. Sie könnte ebenso als Unterscheidungskriterium zwischen den Akzidenzien der Qualia und materiellen Akzidenzien dienen, weil sich die Bedeutung geistiger Zustände nur aus geistigen Zusammenhängen erklären lässt und nicht aus irgendwelchen materiellen Akzidenzien.

Obwohl die schwächere Annahme logisch vorzuziehen wäre, habe ich mich für Reduzierbarkeit als Unterscheidungskriterium entschieden, und zwar aus folgendem Grund:

Die Verbindung von Substanz und Akzidens ist metaphysischer Art: *Untrennbarkeit*. Wenn komplexe Akzidenzien mit der ersten Substanz deshalb verbunden bleiben, weil zwischen ihnen und dem ersten Akzidens eine Verbindung durch die dazwischen liegenden Akzidenzien vermittelt wird, dann muss dies die engste Verbindung sein, die in einer Beschreibung der Natur zwischen Akzidenzien verschiedener hierarchischer Ebenen möglich ist, und das ist eben Reduzierbarkeit.

2. Der Inhalt des soeben abgeleiteten Satzes lässt sich auch so ausdrücken:

Was in der Beschreibung der Wirklichkeit als Rückführbarkeit der Attribute komplexer Objekte auf die Attribute einfacherer Objekte erscheint, äußert sich in der Wirklichkeit selbst als Identität der Substanz: die Substanz bleibt RAUMZEIT, was sich ereignet kann (letztlich) als *physikalischer Prozess* verstanden werden.

Was sich hingegen in der Beschreibung als formale Unabhängigkeit von Attributen zeigt, bedeutet ontologisch die Änderung der Substanz und die Verwandlung des Seienden: die Substanz wird EMPFINDUNG, das Geschehen muss als *geistiger Prozess* aufgefasst werden.

3. Der Grund für die Verwandlung des Seienden lässt sich auch auf einfache und intuitive Weise einsehen:

Wenn die funktionellen Abhängigkeiten neuronaler Zustände von ihren materiellen Voraussetzungen bei ihrer Umwandlung in geistige Zustände verblassen und schließlich ganz verschwinden, dann heißt das, dass sich hier ein Bereich des Universums vom Rest des Universums abkoppelt. Es entsteht ein neues, eigenständiges Universum, ein *Universum von Qualia*.

Bedeutungen sind die Akzidenzien der Entitäten dieses Universums. Sie müssen mit einer Substanz verbunden sein, und ihre Trennung vom Rest des Universums und dessen Substanz RAUMZEIT lässt vermuten, dass zu diesen neuen Akzidenzien eine neue Substanz gehört.

Substanz ist dasjenige, was Seiendem die metaphysische Qualität *Aktivität* verleiht. Man kann nun fragen: Was ist das, was den Entitäten des Universums der Bedeutungen *Aktivität* verleiht? Worauf gründet sich die Dynamik in diesem Universum? Die Antwort ist EMPFINDUNG. EMPFINDUNG ist das, was die Qualia antreibt. Also ist EMPFINDUNG die Substanz geistiger Zustände.

Bedeutung hingegen, dasjenige, was Gegenstand von Beschreibungen ist, also die formale Definition von geistigen Zuständen, ist *passiv*. Bedeutung ist *Information*, und Informationsverarbeitung allein ereignet sich – wie alles, was bloß Element einer Beschreibung oder eines Modells ist – niemals *von selbst*. Sie ist auf Aktivität von außen angewiesen.

Es muss aber wohl nicht eigens betont werden, dass dieses Universum der Qualia ein *inneres* Universum, ein Universum *im Kopf* ist. Die funktionelle Abkopplung des Geistes vom materiellen Universum, die sich im freien Flug der Gedanken und Vorstellungen äußert, bedeutet selbstverständlich nicht, dass sich der Geist, wie Esoteriker und Angehörige verschiedener Religionen meinen, tatsächlich von seinen räumlichen und zeitlichen Bindungen befreien kann. Er wird durch das neuronale Netz hervorgebracht, und dadurch bleibt er an materielle Bedingungen gebunden und in Raum und Zeit gefangen.

4. Da Substanz und Akzidenzien immer zusammengehören, ist es unbefriedigend, verschiedenen Schichten des Seienden von unterschiedlicher Komplexität zwar verschiedene Akzidenzien zuzuordnen, aber dieselbe Substanz. Allerdings besteht ein begriffliches Problem. Die komplexeren Akzidenzien können auf einfachere und schließlich auf das einfachste Akzidens zurückgeführt werden. Im wörtlichen Sinn kann das aber von den Substanzen, die zu diesen komplexeren Akzidenzien gehören, nicht behauptet werden, weil die Substanz begrifflich überhaupt nicht erfassbar ist, so dass der Begriff einer "abgeleiteten Substanz" in seiner eigentlichen Bedeutung unsinnig wäre.

Es ist aber möglich, einen solchen Begriff zu *definieren* und dadurch von seiner eigentlichen Bedeutung abzugrenzen. Definieren wir also:

Abgeleitete Substanz ist eine Substanz, die mit einem abgeleiteten Akzidens verbunden ist.

Jedes Seiende, das kein geistiges Seiendes ist, besteht dann aus einer abgeleiteten Substanz und abgeleiteten Akzidenzien. Nur der *Ursprung des Seienden* besteht *für uns* aus der ersten Substanz und dem ersten Akzidens.

Da wir vom Seienden immer nur die Akzidenzien verstehen, bleibt nicht nur ein Teil des Wesens des Seienden unseren Begriffen entzogen, sondern auch ein Teil jener Veränderung des Wesens des Seienden, die sich beim Aufstieg des Seienden zu immer komplexeren Formen ereignet. Alles, was wir tun können, ist, von der Seite der Akzidenzien her auf die Substanz und ihre Veränderungen zu schließen.

Es ist wichtig, stets in Erinnerung zu behalten, dass auch die Substanz RAUMZEIT des physikalischen Seienden nicht gedacht werden kann, und dass es schon allein aus diesem Grund unmöglich ist, ihre Verwandlung in die zweite Substanz EMPFINDUNG des geistigen Seienden zu denken. Wie soeben gezeigt wurde, kann aber bewiesen werden, dass diese Verwandlung stattfinden muss.

5.5. Kriterium für das Auftreten von Empfindungen

Aus den Überlegungen des vorigen Abschnitts ergibt sich ein Kriterium für das Auftreten von Qualia.

Die Existenz eines Quale setzt voraus, dass seine Bedeutung nicht aus seiner materiellen Beschaffenheit ablesbar ist. Diese Bedingung ist genau dann erfüllt, wenn neuronale Zustände, die etwas repräsentieren oder Input-Output-Relationen herstellen, miteinander vernetzt werden. Dann treten Rückkopplungsschleifen auf, innerhalb derer die Information, die in den neuronalen Zuständen codiert ist, in zunehmendem Maß durch die gegenseitigen Beziehungen der neuronalen Zustände bestimmt wird, während zugleich die ursprüngliche Abhängigkeit von äußeren Bedingungen abnimmt. Was anfangs *Repräsentation* war, wird zur *intrinsischen Bedeutung*.

Es lässt sich bei keinem neuronalen Muster feststellen, was es bedeutet. Das gilt sogar für Wahrnehmungen: hier kann äußerstenfalls bestimmt werden, was sie repräsentieren, aber nicht, was sie bedeuten. Entgegen allen hochfliegenden Hoffnungen der Neurowissenschaft wird es niemals möglich sein, jemandes Gedanken zu belauschen – es sei denn, er teilt freiwillig mit, was er jeweils denkt, und ermöglicht dadurch die Ermittlung der individuellen neuronalen Codierung dieser Gedanken. Aber auch das wird mit Sicherheit nur bei einfachen, standardisierten geistigen Prozessen möglich sein.

Wie schon mehrfach erwähnt, kann der soeben beschriebene Prozess der Entstehung von intrinsischer Bedeutung – und damit auch die Bildung von Qualia – nur dann stattfinden, wenn ein neuronaler Bereich vorhanden ist, der nicht funktionell festgelegt ist.⁴

Beim Menschen ist das das Großhirn. Seine funktionelle Ungebundenheit zeigt sich in seiner Plastizität: Wenn Areale ausfallen, die sich im Lauf der individuellen Entwicklung für bestimmte Aufgaben spezialisiert haben, dann können diese Aufgaben von anderen Regionen übernommen werden.

Auch andere Hirnstrukturen können aber das Kriterium der funktionellen Ungebundenheit erfüllen. Das Gehirn von Oktopoden ist ganz anders strukturiert als unser Gehirn, aber es ist anzunehmen, dass es neuronale Bereiche dieser Art enthält.

Die neuronalen Strukturen des sogenannten Zwischenhirns sind dagegen nicht dafür geeignet, Repräsentationen in einer Weise zu vernetzen, dass sie sich von ihrer ursprünglichen Funktion lösen und zu intrinsischen Bedeutungen entwickeln können. Gleichgültig, ob diese Funktion Teil eines

⁴ Solche Bereiche treten infolge von Mutationen auf, durch die bereits existierende neuronale Strukturen – etwa die des Zwischenhirns – vergrößert werden.

genetischen Programms ist oder durch äußere Bedingungen eingepägt – das Verhalten bleibt schematisch und stets auf den auslösenden Reiz bezogen. Repräsentationen, die in solche funktionellen Abläufe eingebunden sind, können sich nur in engen Grenzen verändern.

In Gehirnen, die – zusätzlich zu den stammesgeschichtlich noch älteren neuronalen Bereichen – nur Strukturen enthalten, die denen unseres Zwischenhirns ähneln, ist daher mit hoher Wahrscheinlichkeit die Entstehung von Qualia nicht möglich.

Die notwendige Bedingung für das Auftreten von Qualia lautet also:

Qualia treten in einem neuronalen Netz nur dann auf, wenn das Netz funktionell ungebundene Strukturen enthält, die die Vernetzung neuronaler Repräsentations-Zustände ermöglichen.

Das ist allerdings nur eine *notwendige* Bedingung. Lässt sich auch ein Kriterium, also eine notwendige und hinreichende Bedingung formulieren? Ich glaube ja. Es lautet:

Jedes Tier, das ein neuronales Netz hat, das funktionell ungebundene Strukturen enthält, erlebt Qualia.

Es ist zu beachten, dass diese Aussage nur für ein *Tier* gilt, und nicht für einen Roboter. Wir könnten ja dem Roboter Susi den Luxus eines zusätzlichen neuronalen Moduls gönnen, das wir mit den schon vorhandenen Neuronen einfach verbinden, ohne seine Funktionalität festzulegen. Dennoch ist auszuschließen, dass Susi Empfindungen haben wird.

Bei einem Tier kann jedoch vorausgesetzt werden, dass seine Art bereits ihre Lebensfähigkeit bewiesen hat, und das ist eine *sehr* starke Voraussetzung. Sie schließt eine Menge technischer Voraussetzungen ein, von denen wir nur wenige *genau* und einige *überhaupt nicht* kennen.

Die Behauptung, dass ein Tier mit einem solchen neuronalen Netz Empfindungen hat, stützt sich auf folgendes Argument:

Es muss lebensfähige Vorfahren dieser Tierart gegeben haben, in deren Gehirn die großhirnähnliche neuronale Struktur noch nicht oder nur in sehr geringem Maß vorhanden war. Vermutlich wäre der von uns betrachtete Organismus also auch ohne die "freie" neuronale Struktur lebensfähig, und das neuronale Netz enthielte alle Funktionen, die für die angemessene Regulation seines Verhaltens erforderlich sind. Wenn nun aber die neue, zunächst funktionsfreie Struktur hinzukommt, dann entstehen *unausweichlich* Metarepräsentationen und vernetzte Repräsentationen: die Information über die Umwelt und die Körperzustände *muss* ja auf irgendeine Weise in die neue Struktur gelangen und

dort weiterverarbeitet werden; – allerdings wiederum nur dann, wenn die neuronale Erweiterung die Fortsetzung eines neuronalen Gewebes ist, das seine Funktionsfähigkeit schon bewiesen hat, und nicht einfach ein Neuronenhaufen. Diese Bedingung, in der wiederum zahlreiche technischer Voraussetzungen zusammengefasst sind, ist aber bei einem Tier sicher erfüllt.

Damit ist aber schon gezeigt, dass es Empfindungen hat.

Ein einfaches ergänzendes Argument

Die Existenz von Empfindungen setzt voraus, dass *jemand* da ist, der empfindet.

Betrachten wir die Empfindung *Schmerz*: die sinnliche Information kann nur dann zur Empfindung *Schmerz* werden, wenn es ein *Subjekt* gibt, das die sinnliche Information zur Kenntnis nimmt.

Es ist nicht notwendig, genauer zu bestimmen, was mit den Begriffen "jemand" oder "Subjekt" gemeint ist. Es genügt, zu sehen, dass in einem neuronalen Netz, in dem nur automatisierte Prozesse stattfinden, die entweder genetisch vorprogrammiert oder durch Umweltbedingungen eingepägt sind, kein Platz für dieses postulierte Subjekt ist.

Die Annahme eines Subjekts setzt voraus, dass Reiz und Verhalten nicht in jedem Fall – sei es in der Form eines Reflexes oder in der Form eines erlernten Programms – fest miteinander verbunden sind. Es muss auch Fälle geben, wo die sinnliche Information nicht direkt zum stereotypen Verhalten führt, sondern einer *weiteren Verarbeitung* zugeführt wird.

Das neuronale Netz muss also zu *Metarepräsentationen* fähig sein.

Eine einzelne Metarepräsentation genügt aber offensichtlich nicht, um die Annahme eines *Subjekts* zu begründen. Dafür ist es erforderlich, dass die Metarepräsentationen als Erinnerungen abgespeichert und miteinander vernetzt werden. Erst dann ist es gerechtfertigt, anzunehmen, da existiere *jemand*, der empfindet.

Dieses einfache Argument führt also abermals zu der notwendigen Bedingung, die gerade eben abgeleitet worden ist: Für das Auftreten von Empfindungen ist die Existenz einer funktionell ungebundenen neuronalen Struktur erforderlich, die die Vernetzung neuronaler Repräsentations-Zustände ermöglicht.

Es ist aber festzuhalten, dass es sich um ein *strukturelles Argument* handelt, und dass es daher, wie alle strukturellen Argumente, ungeeignet ist, die *metaphysische Tatsache* der Verwandlung neuronaler Zustände in Qualia zu begründen. Es kann nur zur Bestimmung notwendiger Bedingungen dienen.

5.6. Wer oder Was hat Empfindungen?

Das Kriterium erlaubt es, eine Grenze zwischen *Automaten* und *empfindenden Wesen* zu ziehen. Die konkrete Einordnung setzt allerdings eine genaue Kenntnis der Strukturen des jeweiligen neuronalen Netzes und seiner Leistungen voraus. Wie schon erwähnt, wurden die Intelligenzleistungen von Vögeln lange Zeit unterschätzt, weil sie dafür nicht die selbe neuronale Struktur gebrauchen wie Säugetiere. Inzwischen ist aber bekannt, dass einige Vogelarten hochintelligent sind. Vögel haben zweifellos Empfindungen.

Was ist mit Fischen? Die Entdeckung zwischenhirnähnlicher Strukturen ist *kein* Hinweis auf Empfindungen. Aber es fragt sich, ob das komplexe Verhalten und die Lernfähigkeit einiger Fischarten nicht auf die Existenz von neuronalen Strukturen hinweisen, die das Kriterium der *funktionellen Ungebundenheit* erfüllen. Vielleicht gibt es Fischarten, bei denen das der Fall ist.

Manche Fragen lassen sich aber mit Hilfe des Kriteriums eindeutig entscheiden:

Nehmen Bienen Farben wahr? Nein. Unterschiedliche Farben werden in ihren neuronalen Netzen zwar verschieden repräsentiert, aber die Repräsentation bleibt reine Information. Es findet keine Umwandlung in EMPFINDUNG statt.

Empfinden Krustentiere Schmerz? Nein. Sie sind nicht empfindungsfähig. Die Beziehung zwischen *Robert* und seiner Garnele wird einseitig bleiben.

Diesseits der Grenze, die durch das Kriterium gezogen wird, gibt es keine Empfindungen. Was ist aber auf der anderen Seite? Was ist, wenn die anfänglich funktionsfreie großhirnähnliche Struktur *sehr klein* ist? Sind dann die Empfindungen irgendwie "blasser"? Ist *rot* weniger rötlich? Ist *Schmerz* weniger schmerzhaft?

Einerseits ist zu bedenken, dass sich die Unbeschreibbarkeit von Empfindungen auf deren Abstufungen überträgt. Es ist nicht möglich, Abstufungen von etwas zu beschreiben, was nicht beschreibbar ist.

Andererseits haben wir aber zu Empfindungen dieses besondere Verhältnis, dass wir sie zwar nicht durch Beschreibungen erfassen können, aber dennoch genau wissen, was sie sind, weil sie uns *unmittelbar* – als sie selbst – gegeben sind. Deshalb meine ich, dass die Vorstellung "blasserer" oder "schwächerer" Empfindungen und eines "dumpfen" Bewusstseins eine geeignete Annäherung an das Wesen derjenigen Qualia ist, die Tiere erleben, bei denen die für das Auftreten von Qualia erforderlichen funktionell ungebundenen neuronalen Strukturen wenig ausgeprägt sind.

Ein wichtiger Aspekt der Schlussfolgerungen des vorigen Abschnitts ist, dass auch in neuronalen Netzen, die zur Bildung von Qualia fähig sind, nicht von Beginn an Qualia vorhanden sind. Die Verwandlung von Materie in Geist, von Repräsentation in intrinsische Bedeutung, von einem neuronalen Muster in ein Quale ist ein Entwicklungsprozess.

Daraus folgt z.B., dass Säuglinge nach der Geburt keine Farben *wahrnehmen*; Auch später, wenn sie schon dazu fähig sind, Farben zu *unterscheiden*, können sie zunächst noch keine *Farbempfindung* haben – eben deshalb, weil sich die Information noch nicht zu einem Quale entwickeln konnte.

5.7. Versuch einer begrifflichen Annäherung an die Verwandlung der Substanz

Die Tatsache, dass die erste Substanz RAUMZEIT sich in die zweite Substanz EMPFINDUNG verwandelt, wirkt zunächst befremdlich – selbst dann, wenn die Notwendigkeit dieser Verwandlung eingesehen werden kann. Dieses Befremden kann auf verschiedene Weise reduziert werden.

Erste Annäherung

Der erste Schritt ist jedenfalls die Einsicht, dass der Begriff RAUMZEIT hier nicht mit dem mathematisch-physikalischen Begriff "Raumzeit" identisch ist. Dieser ist nur *definiert*, er besteht also nur aus den Eigenschaften, durch die er definiert ist, während der Substanzbegriff RAUMZEIT dasjenige bezeichnet, was diese Eigenschaften *hat*, dasjenige also, was die Raumzeit *ohne* diese Eigenschaften "ist", wobei die Anführungszeichen darauf hinweisen, dass der ontologische Status der RAUMZEIT nicht "Existenz" sein kann.

Was sich verwandelt, ist also nicht die physikalische "Raumzeit" – diese Vorstellung wäre absurd –, sondern die undenkbbare erste Substanz RAUMZEIT.

Zweite Annäherung

Wenn wir versuchen, *Raumzeit* als Substanz zu denken, wird sie zu AGENS. Was wir von AGENS wissen, ist, dass ihm die metaphysische Qualität *Aktivität* zukommt.⁵

Materielle Objekte sind raumzeitliche Muster. Es ist daher möglich, ihre Substanz durch das Begriffspaar [Raumzeit, Aktivität] zu definieren:

Materielle Substanz := [Raumzeit, Aktivität]

Auch geistige Zustände sind raumzeitliche Muster, wobei aber zu beachten ist, dass der Begriff "sind" in dieser Aussage keine definitorische Bedeutung hat, wie das in der Aussage "Materielle Objekte sind raumzeitliche Muster" der Fall ist.

Somit kann die zweite Substanz durch das Begriffspaar [Raumzeit, Empfindung] definiert werden:

Geistige Substanz := [Raumzeit, Empfindung]

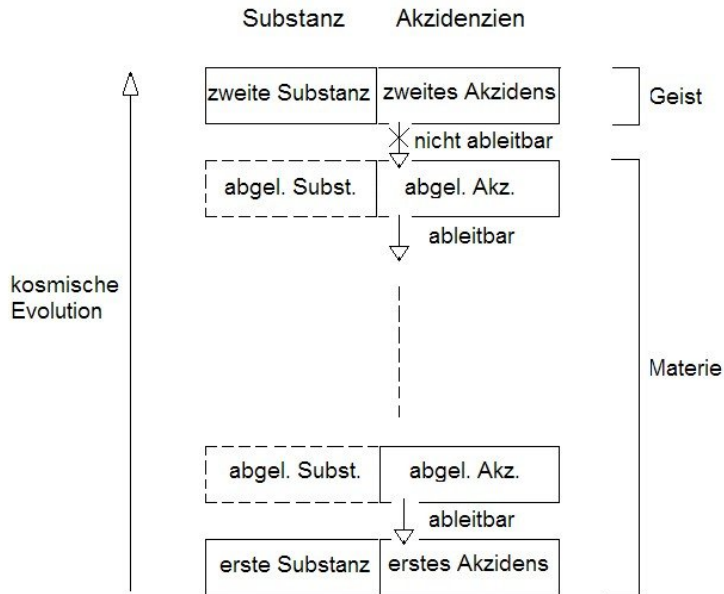
Hier erscheint die Verwandlung des Wesens des Seienden als Verwandlung der metaphysischen Qualität: *Aktivität* wird zu *Empfindung*. Die erste Substanz ist dasjenige, was die Akzidenzien der Materie *aktiv* macht, die zweite Substanz ist dasjenige, was die Akzidenzien des Geistes *aktiv* macht.

Es ist unmittelbar einzusehen, dass Empfindung tatsächlich genau das leistet: Bedeutung ohne *Empfindung* ist gleichgültig. Empfindung ist die Basis jeder Wertung und jeder Motivation. Sie ist das, was uns antreibt.

Dritte Annäherung

Um die kosmische Evolution im Schema von Substanz und Akzidenzien zu begreifen, ist es notwendig, das Verhältnis von Ableitbarkeit bzw. Nicht-Ableitbarkeit einerseits und Wesensgleichheit bzw. Wesensänderung andererseits zu bestimmen. Zunächst eine Skizze zur besseren Übersicht:

⁵ Es ist zu beachten, dass *Aktivität* kein Akzidens ist. Wenn wir den *Ursprung des Seienden*, um ihn für unser Denken zugänglich zu machen, in Substanz und Akzidens teilen und als Aktivität von AGENS bezeichnen, dann wird hier zwar *Aktivität* zum Akzidens, aber gerade dadurch verschiebt sich das, worin sich das Akzidens "Aktivität" von der metaphysischen Qualität *Aktivität* unterscheidet – d.h. das, was dem Akzidens *fehlt* – in den Begriff der Substanz. Substanz ist dann das, was die Akzidenzien *aktiv* macht.



Von dieser Skizze ist uns nur die rechte Seite zugänglich, die Seite der Akzidenzien. Die linke Seite handelt von genau demjenigen, was allen Beschreibungen fehlt und was nicht gedacht werden kann, von der Substanz.

Deshalb ist *für uns* das Verhältnis einer Seinsschicht zur darunter liegenden Seinsschicht ausschließlich durch die Alternative *ableitbar* oder *nicht ableitbar* gegeben, d.h. durch das Verhältnis der Akzidenzien. Ableitbarkeit bzw. Reduzierbarkeit ist aber eine derart enge Beziehung zwischen dem, was abgeleitet ist und dem, woraus es abgeleitet ist, dass es *für uns* so scheint, als wäre das Abgeleitete und das Ursprüngliche *eigentlich* dasselbe. Diese Identität gilt jedoch nur für Beschreibungen; In der Wirklichkeit ist jedes Akzidens untrennbar mit einer zugehörigen Substanz verbunden, und im Bereich der Substanz kann von Ableitbarkeit nicht die Rede sein. Der Begriff *abgeleitete Substanz* ist nur auf die in 5.4. unter Bemerkung 4 definierte Weise zu verstehen: als Substanz, die zu einem abgeleiteten Akzidens gehört.

Seiendes ist immer Akzidens *und* Substanz. Wenn nun der Begriff der Ableitbarkeit im Bereich der Substanz nicht zulässig ist, dann ist damit zugleich die Vorstellung aufgehoben, dass Seiendes, dessen Akzidenzien ableitbar sind, ein ableitbares Seiendes ist. Die Identität zwischen dem Abgeleiteten und

dem, woraus es abgeleitet ist, die im Bereich der Beschreibungen behauptet werden kann, ist auf die Wirklichkeit nicht übertragbar. Sie erweist sich als Täuschung, die nur *für uns* existiert, und der wir genau dann unterliegen, wenn wir Wirklichkeit und Beschreibung auf die Weise gleichsetzen, wie das in der derzeitigen Naturwissenschaft der Fall ist.

Auf dieselbe Weise, wie sich Wirklichkeit und Beschreibung unterscheiden, sind auch Evolution und Ableitung verschieden. Es wäre durchaus angemessen, anzunehmen, dass sich das Wesen des Seienden bei *jedem* evolutionären Schritt zu einer komplexeren Seinsschicht wandelt. Das ändert jedoch nichts an der Tatsache, dass *für uns* die Substanz gleich bleibt, solange die Akzidenzien ableitbar sind, und dass sie erst dann verwandelt erscheint, wenn die Akzidenzien nicht ableitbar sind. Deshalb ist es gerechtfertigt, die Wirklichkeit in einen materiellen und einen geistigen Bereich einzuteilen.

Vierte Annäherung

Zuletzt ist es in einem gewissen Maß auch möglich, den Begriff der geistigen Substanz auf folgende Weise direkt zu rekonstruieren:

Es wird ein Begriff der Substanz eines geistigen Zustands benötigt, der sich nicht auf die physikalische oder irgendeine andere Ebene *unterhalb* der geistigen Ebene bezieht, sondern der sich als zum geistigen Zustand selbst gehörig ausweist und der geeignet erscheint, den Akzidenzien der geistigen Zustände – den intrinsischen Bedeutungen – *Aktivität* zu verleihen.

Wir wissen ja schon, welcher Begriff dafür geeignet ist: *Empfindung*. Aber die Frage ist jetzt, wie weit die Substanz EMPFINDUNG *als Begriff* zugänglich ist, wie weit wir EMPFINDUNG sozusagen *von außen* verstehen könnten, wenn wir sie nicht *von innen*, also aus unserem eigenen Erleben kennen.

Stellen wir uns also vor, wir wüssten alles über neuronale Netze, aber wir wüssten nicht, was EMPFINDUNG ist, und unsere Aufgabe wäre es, herauszufinden, was die Substanz geistiger Zustände ist. Wie weit kämen wir?

Wir müssen auf der Seite der Akzidenzien beginnen. Was finden wir vor? Intrinsische Bedeutungen. Sie folgen zeitlich aufeinander und sind miteinander in einem virtuellen Raum vernetzt. (Ich vermeide den Ausdruck Parameterraum, weil ich meine, dass die sich ständig verändernde Dynamik der vernetzten Bedeutungen nicht auf einen Parameterraum übertragen werden kann. Es gibt keine gesetzmäßige Beziehung zwischen den physikalisch-chemischen und neuronalen Parametern und den sich wandelnden Bedeutungen der Muster, weil sich in diesem Szenario lokale und globale Gesetzmäßigkeiten auf nicht formalisierbare Weise ergänzen. Mehr darüber später.)

Die Frage ist: Was verleiht den Bedeutungen *Aktivität*? Die Antwort muss ein Begriff sein, der auf derselben Ebene liegt wie Bedeutung, ein Begriff, der mit dem Begriff Bedeutung ebenso ein Paar bildet wie die Begriffe Masse und Gravitation.

Bei der Frage: "In welchen Systemen können Empfindungen auftreten?" haben wir festgestellt, dass da *jemand* sein muss, der empfindet. Dasselbe gilt für Bedeutungen: von Bedeutungen zu reden ist nur sinnvoll, wenn da *jemand* ist, *für den* sie etwas bedeuten.

Damit haben wir auf der Seite der Akzidenzien folgende Voraussetzungen: Die Akzidenzien geistiger Zustände sind Bedeutungen. Das, was sie bedeuten, bedeuten sie *für ein Subjekt*. Sie stellen sich diesem Subjekt unmittelbar und als Ganze dar, also nicht in der Gestalt von Mengen von Variablenwerten.

Wie könnte ein Substanz-Begriff beschaffen sein, der geeignet ist, mit diesem Akzidens-Begriff ein Paar zu bilden – eine untrennbare Einheit, wie Masse und Gravitation?

Was wir schon wissen, ist, dass es in jedem neuronalen Netz, das zu einem lebensfähigen Wesen gehört, ein System gibt, das die neuronalen Zustände *bewertet*. Aus diesem System muss die Regulation der Dynamik der globalen neuronalen Zustände hervorgegangen sein, bevor sie zu geistigen Zuständen wurden und ihre Dynamik selbst bestimmten.

Für das Subjekt müssen die Bewertungen genauso unmittelbar gegeben sein wie die Bedeutungen, also abermals nicht in der Form von Mengen von Variablenwerten, sondern in der Form von *Qualitäten*.

Was ist eine *Qualität für ein Subjekt*?

Wenn wir den gesuchten Begriff nicht schon wüssten, dann würden wir ihn nun genau *so* definieren: als "Qualität für ein Subjekt".

Qualität für ein Subjekt ist geeignet, mit dem Akzidens *Bedeutung für ein Subjekt* eine untrennbare Einheit zu bilden. *Qualität für ein Subjekt* ist die Substanz des geistigen Seienden.

Ich glaube, damit wäre die maximal mögliche gedankliche Annäherung an die letztlich keinem Begriff zugängliche Substanz EMPFINDUNG erreicht.

5.8. Philosophische Zombies

"Philosophische Zombies" sind hypothetische Wesen, die zu dem Zweck ersonnen wurden, die Frage des Verhältnisses zwischen neuronalen Zuständen und Qualia und damit zugleich das Problem der Unterscheidung zwischen Automaten und empfindenden Wesen zu verdeutlichen.

Die Physiologie von Philosophischen Zombies gleicht der von Menschen. Zombies haben also dasselbe Gehirn wie Menschen. Ihnen fehlt jedoch Empfindung. Sie erleben keine Qualia.

Ist es denkbar, dass solche Zombies existieren?

Von unserem Standpunkt aus ist sofort zu sehen, dass ihre Existenz unmöglich ist:

Die Voraussetzung, dass ein Zombiergehirn in physiologischer Hinsicht mit dem eines Menschen identisch ist, schließt die Identität von Strukturen *und* Prozessen ein. Das bedeutet, dass die neuronalen Zustände eines Zombiergehirns auf dieselbe Weise miteinander vernetzt sind und sich durch Interaktion verändern wie die Zustände eines Menschengehirns. Daraus folgt aber, dass der Prozess der Verwandlung von einem *physikalischen Seienden* in ein *Quale* stattgefunden hat.

Also sind die neuronalen Zustände – die raumzeitlichen Erregungsmuster – des Zombiergehirns ebenfalls Qualia.

Wenn man annähme, der Zombie habe keine Empfindungen und die neuronalen Muster des Zombiergehirns seien *nichts als* Informationsverarbeitung, dann wäre das also genauso absurd wie die Annahme, die Erde könne auch ohne Gravitation existieren. Es würde bedeuten, vom Seienden gedanklich seine *Substanz* zu entfernen, d.h. dasjenige, was den Akzidenzien die metaphysische Qualität *Aktivität* verleiht. Substanz und Akzidens sind aber untrennbar, sie sind immer *eins*. Bei der Erde erscheint uns das selbstverständlich, im Fall der Qualia jedoch erscheinen solche absurden gedanklichen Konstruktionen wie Neurozombies möglich – aber nur solange man keinen Begriff davon hat, was Qualia sind.

Wenn man ein neuronales Erregungsmuster, das zum Quale geworden ist, aus dem Netz der Bedeutungen herausnehmen könnte, dem es seine Existenz verdankt, dann allerdings wäre es tatsächlich *nichts als* ein neuronales Muster. *Als geistiges Seiendes* hätte man es dann zerstört, es hätte sich in ein physikalisches Seiendes zurückverwandelt.

Es ist aber klar, dass man ein *globales* neuronales Erregungsmuster nicht von seiner neuronalen Umgebung trennen kann, ohne es auch physikalisch zu verändern. Die Struktur könnte dabei vielleicht kurzzeitig erhalten bleiben, aber keinesfalls die Dynamik. Diese würde sich augenblicklich – beginnend mit dem Zeitpunkt der Trennung – verändern und nach kurzer Zeit zusammenbrechen.

5.9. Künstliche Intelligenz

Verstehen

Kann eine Maschine, die Zeichenketten nach bestimmten Regeln in andere Zeichenketten umwandelt, etwas *verstehen*?

Diese Frage lässt sich mit erstaunlicher Klarheit durch ein Gedankenexperiment von John Searle beantworten, das den Namen "Chinesisches Zimmer" trägt. Die folgende Kurzversion hat Searle 1999 formuliert:⁶

"Imagine a native English speaker who knows no Chinese locked in a room full of boxes of Chinese symbols (a data base) together with a book of instructions for manipulating the symbols (the program). Imagine that people outside the room send in other Chinese symbols which, unknown to the person in the room, are questions in Chinese (the input). And imagine that by following the instructions in the program the man in the room is able to pass out Chinese symbols which are correct answers to the questions (the output). The program enables the person in the room to pass the Turing Test⁷ for understanding Chinese but he does not understand a word of Chinese."

Falls man akzeptiert, dass dieser Ablauf möglich ist – und dazu ist man gezwungen, wenn man *künstliches Verstehen* für möglich hält – dann ist dieses Argument gegen den Turing-Test stichhaltig. Das Chinesische Zimmer *ist* ja eine Version des Turing-Tests: Zwar ist hier der unbekannte Kommunikationspartner ein Mensch, aber er versteht sicher kein Chinesisch. Da er aber den Test besteht, müsste ihm die Fähigkeit, chinesisches zu verstehen, zuerkannt werden.

⁶ John Searle, *The Chinese Room*, in R.A. Wilson and F. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA: MIT Press, 1999.

⁷ Dieser nach Alan Turing benannte Test besteht darin, dass Menschen durch schriftliche Kommunikation mit einem nicht sichtbaren Partner feststellen sollen, ob dieser Partner ein Computer ist oder ein Mensch. Falls die Entscheidung nicht möglich ist, werden dem Partner menschliche Fähigkeiten zuerkannt.

Searle fügt dann hinzu: "The point of the argument is this: if the man in the room does not understand Chinese on the basis of implementing the appropriate program for understanding Chinese then neither does any other digital computer solely on that basis because no computer, qua computer, has anything the man does not have."

In dieser Form ist das Argument allerdings nicht korrekt: Wenn das Chinesische Zimmer als Computer-Analogie aufgefasst wird, dann ist der Mann darin offensichtlich nicht der ganze Computer, sondern er übernimmt nur die Aufgaben des Datentransports und der Transformation. Auch der härteste Verfechter der Existenz künstlicher Intelligenz würde aber kaum behaupten, dass Prozessor und Motherboard allein schon etwas verstehen könnten.

Deshalb wurde Searle entgegnet⁸, dass nicht *der Mann*, sondern *das System* Chinesisch versteht.

Searle konterte mit dem Argument, der Mann könne ja die Daten und das Programm auswendig lernen, und dann *wäre* er das System.

Ich breche die Diskussion des Gedankenexperiments an dieser Stelle ab, weil ich davon überzeugt bin, dass mit der letztgenannten Variante der Beweis erbracht ist, dass Computer nicht verstehen, was sie tun: Die Aktionen des Manns entsprechen genau denen eines Computers – er transformiert Zeichenketten nach vorgegebenen Regeln in andere Zeichenketten⁹ – und er versteht *nicht*, was die Zeichenketten bedeuten.

Etwaige Zweifel daran, ob das Gedankenexperiment überhaupt realisierbar wäre, richten sich nicht gegen das Argument, sondern in jedem Fall nur gegen die Annahme, dass Verstehen simuliert werden kann. Die Stärke des Arguments liegt ja gerade darin, dass das Maximum dessen vorausgesetzt wird, was in einer Simulation erreichbar ist: es wird angenommen, dass die Antworten, die der Mann gibt, ohne sie zu verstehen, von denen eines Menschen, der das, was er antwortet, auch versteht, nicht unterschieden werden können.

Das heißt: Sogar dann, wenn die KI ihr Ziel erreicht hätte, Kommunikation so zu modellieren, dass sie sich von menschliche Kommunikation nicht unterscheidet, würde der Simulation Verstehen fehlen.

⁸ Unter anderem von Georges Rey: *What's Really Going on in Searle's "Chinese Room"*, *Philosophical Studies* 50, 169–185. 1986.

⁹ Um auch die Möglichkeit eines lernfähigen Programms zu berücksichtigen, kann man einfach hinzufügen, dass der Mann auch die Zeichenketten der Regeln in seinem "book of instructions" abhängig vom Input transformiert.

Eine weitere Stärke des Gedankenexperiments liegt darin, dass es eine einfache und intuitive Definition von Verstehen enthält: Man kann sich mit dem Mann auf chinesisches *und* auf Englisch unterhalten. Englisch versteht er, Chinesisch versteht er nicht. Verstehen ist das, wodurch sich diese beiden Aktivitäten unterscheiden. Jeder von uns kennt diesen Unterschied. Wir *verstehen* ihn.

Searles Argument beweist also, dass selbst einer perfekt gelungenen Computersimulation menschlicher Kommunikation das Verständnis dessen fehlt, *was* simuliert wird.

Das Argument enthält jedoch keine Antwort auf die Frage, *warum* dieser Unterschied zwischen Mensch und Computer besteht. Gerade die Deutlichkeit, mit der sich das Fehlen von Verstehen in Searles Gedankenexperiment zeigt, lässt die Klärung dieser Frage dringend und längst überfällig erscheinen.

Warum Computer keinen Geist hervorbringen; die *formale* Begründung

In den Kapiteln 3, 4 und 5 ist *Geist* auf naturwissenschaftliche Weise definiert worden; seine Existenz ist mit naturwissenschaftlichen Methoden begründet und die Verwandlung des Wesens des physikalischen Seienden in ein Quale mit metaphysischen Argumenten erklärt worden. Also muss die Antwort darauf, warum Computer nicht verstehen, was sie tun, in den Schlussfolgerungen dieser Kapitel enthalten sein. Das soll im Folgenden demonstriert werden.

Klären wir zunächst eine Frage, die sich sofort aufdrängt: *Wenn Menschen sprachlich kommunizieren, sind sie dann nicht ebenfalls Maschinen, die Zeichenketten nach vorgegebenen Regeln transformieren?*

Die Antwort ist *nein*. Im Kapitel 3. *Willensfreiheit* ist gezeigt worden, dass die geistige Tätigkeit weder feststehenden Regeln folgt noch Regeln, die sich nach feststehenden Meta-Regeln ändern.

(Im Kapitel 4. *Das veränderte Bild der Wirklichkeit* habe ich begründet, dass diese Behauptung keinen Widerspruch zu der Annahme bildet, dass an jedem Ort und zu jedem Zeitpunkt durch das fundamentale Gesetz festgelegt ist, wie die Zukunft aus der Gegenwart entsteht.)

Das nun folgende Gedankenexperiment dient der Hervorhebung desjenigen Unterschieds zwischen Mensch und Computer, aus dem hervorgeht, warum Computersimulationen geistiger Leistungen kein Verstehen enthalten.

Sei M ein menschliches neuronales Netz. Die Kombination der Werte aller Variablen von M zu einem bestimmten Zeitpunkt ist ein *Zustand* von M .

Wir betrachten eine Zustandsfolge von M in einem Zeitraum Δt , der sich von $t = 0$ Sekunden bis, sagen wir: $t = 1000$ Sekunden erstreckt. Nennen wir diese Folge D .

Die Zeitintervalle zwischen zwei aufeinander folgenden Zuständen können beliebig klein gewählt werden. Nehmen wir an, ihre Länge sei 10^{-24} Sekunden. Demnach enthält unsere Folge D 10^{27} Zustände von M . (Da es sich um ein Gedankenexperiment handelt, kann die Dauer der Zeitschritte aber auch weiter verringert werden.) Ferner legen wir fest, dass die Variablenwerte mit einer Genauigkeit von 1024 Bits dargestellt sind. (Auch diese Genauigkeit kann natürlich beliebig hoch angenommen werden.)¹⁰

Als Voraussetzung für die folgende Argumentation werden die Schlussfolgerungen benötigt, die im Kapitel über Willensfreiheit erarbeitet worden sind. Hier eine kurze Zusammenfassung:

Das neuronale Netz besteht aus mehreren übereinander geschichteten Ebenen von miteinander wechselwirkenden Entitäten ansteigender Komplexität. Deshalb sind die Naturgesetze zur Beschreibung der Dynamik des Netzes nicht ausreichend. Es müssen weitere Gesetze – sogenannte Strukturgesetze – berücksichtigt werden. Die Gesetze der hierarchisch obersten Schicht, also die geistigen Gesetze, sind dominant, d.h. die Dynamik der geistigen Zustände bestimmt die neuronale und molekulare Dynamik in einem höheren Maß als das umgekehrt der Fall ist. (Im 4. Kapitel wurde diese Tatsache als *Kausalität von oben* bezeichnet.)

Die Gesetze des Geistes sind jedoch, im Gegensatz zu den Gesetzen der anderen Ebenen, nicht feststehend. Die geistige Tätigkeit wirkt auf sich selbst zurück: sie ändert ihre eigene neuronale Codierung und damit auch das Netz der Bedeutungen, und daraus folgt zugleich eine Änderung der Übergangsregeln der Abfolge geistiger Zustände.

Jetzt beziehen wir in unsere Überlegungen die folgende Turing-Maschine ein:

¹⁰ Der Grund für die Einführung von D ist, dass D direkt mit einem Computer bzw. mit einer Turing-Maschine verglichen werden kann.

(Ich werde hier keine Definition der Turingmaschine geben. Um das Folgende zu verstehen, genügt es, zu wissen, dass eine Turing-Maschine alles berechnen kann, was ein beliebiger Computer berechnen kann.)

T sei die Turing-Maschine, die die Zustandsfolge D im Zeitraum $\Delta t/2$, der bei $t = 0$ Sekunden beginnt und bis $t = 500$ Sekunden reicht, *exakt* wiedergibt.¹¹ (Die Existenz von T kann vorausgesetzt werden.)

Die Zustände von T stimmen also (in der gewählten Näherung) mit den Zuständen von M im Zeitraum $\Delta t/2$ überein.

Die Dynamik von T ist berechenbar. Es existiert also eine Funktion – nennen wir sie f_T –, die es erlaubt, aus dem Anfangszustand von T – das ist die anfängliche Sequenz der Zeichen 0 und 1 auf dem Band von T, die dem Zustand von M am Beginn des Intervalls Δt entspricht – die Folge der Zustände von T im Intervall Δt zu berechnen.¹²

Unsere erste Frage ist: *Was geschieht **nach** dem Ende des Intervalls $\Delta t/2$?*

Die Antwort folgt aus dem Unterschied zwischen M und T:

Das Netz M produziert nach Voraussetzung aus sich selbst heraus – in Rückkopplungsschleifen – neue Regeln des Übergangs von einem Zustand in den nächsten. Was *nach* dem Ende von $\Delta t/2$ geschieht, ist also in den *bis dahin* geltenden Regeln nicht enthalten.

Hingegen sind die Regeln von T fixiert. Die Turing-Maschine T ist an die vorher festgelegten Regeln gebunden;

Daraus folgt, dass nach dem Ende von $\Delta t/2$ – wenn wir die Turingmaschine T weiter laufen lassen – die Zustände von T nicht mehr den in D enthaltenen Zuständen von M entsprechen können. Vielmehr muss von da an eine wachsende Differenz zwischen den Zuständen von M und T auftreten.

Wenden wir uns nun wieder dem Problem der Simulation von Verstehen zu. Es hat nun eine etwas deutlichere Gestalt angenommen:

¹¹ Jeder Zustand von M, der durch T simuliert wird, ist eine Folge der Zeichen 0 und 1 auf dem Band von T. Ich werde eine solche Zeichenfolge, sofern sie einen der Zustände von M repräsentiert, als *Zustand* von T bezeichnen, und das, was üblicherweise "Zustand" der Turingmaschine genannt wird, werde ich *interner Zustand* nennen.

¹² Hier könnte die Frage auftauchen, warum nicht anstelle der Funktion f_T digitalisierte Versionen der Naturgesetze verwendet werden, um die Zustandsfolge von T zu berechnen. Das ist aber nicht möglich, weil die nichtlineare Rückkopplung die *exakte* Berechnung späterer Zustände von M aus Anfangsbedingungen und solchen Gesetzen verhindert. (Ich erinnere an den Vergleich mit einer großen Anzahl gravitierender Körper: auch hier existiert kein exaktes Verfahren zur Berechnung der Zukunft aus Gesetz und Anfangsbedingungen.)

Das System M bringt Geist hervor, M versteht, was geschieht. Nun ist aber die Folge der Zustände des Systems T in beliebiger Näherung identisch mit der Folge von Zuständen von M im Zeitraum $\Delta t/2$. T ist also eine perfekte Simulation von M, die während dieses Intervalls die Zustände von M – z.B. auch die Zustände, die die Sprachausgabe steuern – reproduzieren kann.

Heißt das, dass auch T während dieser Zeit etwas versteht?

Zunächst ist Folgendes zu klären:

Was bedeutet eigentlich die Identität der Zustände der beiden Systeme M und T? Ist T eine digitalisierte Version von M? Sind M und T – in diesem Sinn – identisch bis auf die unterschiedliche physikalische Realisierung?

Wenn das der Fall wäre, dann müsste der Turing-Maschine Verstehen zugestanden werden. Das haben wir aber vorhin – bei der Diskussion des chinesischen Zimmers – ausgeschlossen. Außerdem stünde diese Annahme im Widerspruch zu der Tatsache, dass sich die Zustände der beiden Systeme nach dem Ende von $\Delta t/2$ zunehmend unterscheiden. Wäre T einfach eine digitalisierte Version von M, dann würden M und T auch weiterhin übereinstimmen.

Also sind zwar die Zustände von M und T innerhalb von $\Delta t/2$ identisch, aber T ist dennoch keine digitalisierte Version von M. Da aber "Zustand" als Menge *aller* Variablenwerte des Systems zu einem bestimmten Zeitpunkt definiert ist, folgt daraus, dass der strukturelle Unterschied zwischen M und T in den *Übergangsregeln* liegen muss, denen die Zustandsfolgen der beiden Systeme gehorchen.

Was wissen wir eigentlich über *diesen* Unterschied zwischen M und T?

In M existieren mehrere Ebenen von miteinander wechselwirkenden Entitäten. Alle Ebenen weisen eine eigene Dynamik auf, und zwischen den verschiedenen Ebenen besteht ein komplexes Abhängigkeitsverhältnis. In verschiedenen Situationen kann die Dynamik einer bestimmten Ebene in den Vordergrund treten: bei Reflexhandlungen etwa die Dynamik der neuronalen Ebene, bei Defekten des Transmittersystems die Dynamik der chemischen Ebene. Für uns von Interesse ist jedoch die Dynamik der Attraktoren des Netzes, die wir als geistige Zustände identifiziert haben.

In T wird dagegen die Folge der Zustände Schritt für Schritt, ein Zeichen nach dem anderen, durch die Operation [Lesen, Schreiben, Band bewegen, Wechsel des internen Zustands] erzeugt.

Damit sind wir nun bei der konkreten Formulierung der Frage angelangt, ob Computer *Verstehen* simulieren können. Sie lautet:

Kann die Dynamik eines neuronalen Netzes, die auf der Interaktion der Attraktoren des Netzes beruht, durch schrittweise, nach vorgegebenen Regeln durchgeführte Transformation von Variablenwerten simuliert werden?

Wie zu erwarten war, ist die Antwort *nein*. Der Grund ist, dass die Beschreibung der Entstehung und der Veränderung von Ordnung durch *globale Variable* erfolgt, und dass diese Beschreibung, wie in Kapitel 4 gezeigt wurde, nicht auf die Beschreibung durch *lokale Variable* zurückgeführt werden kann. Dasselbe gilt aber auch umgekehrt. Der Zusammenhang der beiden Arten der Beschreibung ist *nicht formalisierbar*.

Wie erwähnt, wirkt aber die geistige Tätigkeit auf die neuronale Codierung zurück, was sich wiederum auf die geistige Tätigkeit auswirkt. Das heißt: Die Dynamik der globalen Variablen ändert die Dynamik der lokalen Variablen und umgekehrt.

Wenn die Turing-Maschine T eine vollständige Simulation des menschlichen neuronalen Netzes M darstellen soll – mit anderen Worten: wenn M und T tatsächlich isomorph sein sollen –, dann muss T sowohl die lokale als auch die globale Dynamik rekonstruieren. Das ist aber unmöglich, weil beide zwar miteinander zusammenhängen, aber dieser Zusammenhang eben nicht durch einen Algorithmus ausgedrückt werden kann. Wenn sie aber einzeln, jede für sich, rekonstruiert werden, dann sind sie nach kurzer Zeit falsch, weil die Auswirkungen der jeweils anderen Dynamik nicht berücksichtigt werden.

Die Interaktion von Attraktoren kann daher nicht durch eine schrittweise Berechnung von Variablenwerten simuliert werden, wie sie in Turing-Maschinen erfolgt. Die Attraktoren – die geistigen Zustände – müssen *als solche* vorhanden sein, um die zugehörige Zustandsdynamik zu erzeugen. Anders gesagt: Nur die neuronalen Erregungsmuster *als Ganze* bringen die Dynamik des neuronalen Netzes hervor. Ihre Existenz ist eine notwendige Voraussetzung dieser Dynamik.

In einer Turing-Maschine bzw. in einem Computer *haben* sie aber keine Existenz. Auch die Annahme von Parallelverarbeitung und die Annahme von mehrfach hierarchisch übereinander geschichteten Programmstrukturen ändern daran nichts. Gleichgültig, wie komplex (künftige) Computersimulationen geistiger Leistungen auch sein mögen, wie viele Ebenen von Metarepräsentationen sie enthalten und wie viele Prozessoren parallel arbeiten – es bleibt das unverrückbare Faktum, dass alle Veränderungen sich *schrittweise und voneinander getrennt* vollziehen und kein *ganzes, autonomes Seiendes* formen.

Und das bedeutet, dass die Dynamik geistiger Prozesse – also die Dynamik der Attraktoren des Netzes – nicht im Computer rekonstruiert werden kann.

Ich fasse zusammen: Die Attraktoren des Zustandsraums des neuronalen Netzes existieren in einem Computer nicht. Ihre Simulation ist nicht möglich, weil die Dynamik ihrer Interaktionen ein *eigenständiges Element* der Wirklichkeit und ihrer Beschreibung darstellt und nicht auf die Dynamik lokaler Entitäten reduzierbar ist. Daher kann das Zusammenwirken der Dynamik der lokalen Entitäten (der Neurone) und der globalen Entitäten (der Attraktoren, d.h. der geistigen Zustände) nicht durch eine Kette von Transformationen von Variablen abgebildet werden.

Letztlich besagt dieses Argument, dass ein geistiger Zustand *ein unteilbares Seiendes* ist. Geistige Zustände sind Einheiten von *Bedeutung und EMPFINDUNG*. Als solche sind sie unteilbar und können nur *als Ganze* miteinander interagieren, und diese Bedingung überträgt sich auf die physikalische Gestalt, auf das neuronale Muster, den Attraktor, der die notwendige Voraussetzung für die Existenz des geistigen Zustands ist.

Damit ist auch geklärt, inwiefern M und T nicht isomorph sind. Die Entitäten, von denen die Dynamik von M *eigentlich* bestimmt wird, existieren in T überhaupt nicht, weder als sie selbst noch in simulierter Form. Also können auch die Übergangsregeln der Zustände von M in T nicht existieren.

Auch wenn die Übergangsregeln von T eine vollständige und beliebig genaue Simulation der Zustände von M in einem bestimmten Zeitintervall ermöglichen, kann doch mit Sicherheit behauptet werden, dass diese Regeln die *kausalen Beziehungen* – jene Beziehungen, von denen die Entwicklung des wirklich existierenden Systems M *tatsächlich* abhängt, wie z.B. *Gründe* – nicht enthalten.

Daraus geht hervor, dass die Dynamik eines Systems nicht mit einer Funktion gleichgesetzt werden darf, aus der sich eine Zustandsfolge des Systems in einem gegebenen Zeitraum berechnen lässt. Zwischen beiden besteht eine unaufhebbare ontologische Differenz.

Das ist also der Grund, warum zwar einerseits in T ein Algorithmus implementiert werden kann, der eine Folge von Zuständen erzeugt, die den Zuständen von M in einem gewählten Zeitraum entsprechen, warum aber andererseits dieser Algorithmus nicht über diesen Zeitraum hinaus gilt: Die Struktur des Zustandsraums des neuronalen Netzes M wird durch das Beziehungsgeflecht geistiger Zustände bestimmt, und es ist nicht möglich, diese Struktur durch einen Algorithmus zu reproduzieren, weil die Beziehungen der geistigen Zustände nicht formalisierbar sind.

Diese Aussage ist auch der erste Schritt zur Begründung der metaphysischen Behauptung, dass T nicht versteht, was geschieht.

Den zweiten Schritt bildet die folgende, in den Abschnitten 5.3 und 5.4 dieses Kapitels abgeleitete Aussage:

Die Verwandlung physikalischer Zustände – neuronaler Muster, die Attraktoren der globalen neuronalen Dynamik sind und etwas repräsentieren – in Qualia kann sich nur dann vollziehen, wenn auch diese Muster selbst wiederum in einem so hohen Grad miteinander vernetzt sind, dass ihr Informationsgehalt durch diese interne Vernetzung bestimmt wird und nicht als Funktion anderer – innerer oder äußerer – Gegebenheiten aufgefasst werden kann.

Wenn nun aber dieses wandlungsfähige Attraktor-Netz in einem Computer nicht existiert, weil seine Dynamik nicht durch Transformation von Variablenwerten simuliert werden kann, dann folgt daraus, dass ein Computer keinen Geist hervorbringt. Seine Zustände sind keine Qualia. Seine Dynamik bleibt reine Informationsverarbeitung und wird nie zu Geist.

Verstehen ist aber ein geistiger Prozess. Also gibt es in Computern kein Verstehen.

Es bleibt jedoch offen, wie weit es gelingen kann, bestimmte geistige Leistungen durch (geistlose) Simulationen zu rekonstruieren. Der Raum des Möglichen scheint aber in fast allen Fällen wesentlich enger zu sein als das Ausmaß der jeweiligen Hoffnungen; Nur selten bestätigt sich der anfängliche Optimismus über die ersten Schritte der Verwirklichung hinaus. Mit dem Verlassen begrenzter, formalisierbarer Subwelten sinken die Leistungen der Simulationen meist auf ein unakzeptables Niveau.

Es bleibt zuletzt die Frage, ob es überhaupt Geist geben kann, der nicht durch die natürliche Evolution entstanden, sondern von uns geschaffen ist. Gefragt ist also nicht die Möglichkeit der Erzeugung von künstlicher *Intelligenz*, sondern von künstlichem *Geist*.

Ich sehe keinen prinzipiellen Grund, die Existenz eines von uns geschaffenen Geistes auszuschließen.¹³ Es ist allerdings zurzeit unklar, wie ein System, das Geist hervorbringt, auf andere als auf biologische Weise realisiert werden könnte.

Außerdem lässt sich auch nicht abschätzen, wie viel evolutionäre Erfahrung – die bei der Erzeugung von künstlichem Geist technisch umgesetzt werden müsste – in die Chemie und Struktur eines menschlichen neuronalen Netzes und des zugehörigen Körpers eingegangen ist.

¹³ Allerdings folgt im nächsten Abschnitt (5.10. Der metaphysische Unterschied zwischen Wirklichkeit und Simulation) eine prinzipielle Einschränkung.

Es ist aber anzunehmen, dass es wesentlich mehr ist, als diejenigen meinen, die gegenwärtig über eine Optimierung des Menschen nachdenken.

Bemerkung:

Letztlich gehen alle formalen und technischen Argumente, warum die Dynamik natürlicher Systeme nicht exakt simuliert werden kann, auf die Tatsache zurück, dass die Natur kein algorithmisches System ist; Sie erzeugt die Zukunft aus der Gegenwart nicht durch die Anwendung eines Algorithmus, sondern allein aus den an jedem Ort gegebenen differenziellen Bedingungen.

Wie im vierten Kapitel ausgeführt, kann die Art, wie die Natur die Zukunft erzeugt, nicht auf eine Beschreibung oder ein Modell übertragen werden. Deshalb ist eine genaue und vollständige Beschreibung oder Modellierung der Wirklichkeit unmöglich.

Damit ist Simulationen eine prinzipielle Grenze gesetzt.

In denjenigen Bereichen der Wirklichkeit, wo sich die Natur durch die Bildung von Ordnung und den daraus folgenden Strukturgesetzen algorithmischer Beschreibbarkeit nähert, ist diese Grenze ohne Bedeutung. Ein bekanntes Beispiel sind atomare und molekulare Vorgänge. Hier ist in manchen Fällen eine Beschreibung möglich, die sich von der Wirklichkeit nicht mehr messbar unterscheidet.

Es gibt aber auch Bereiche, die sich der Beschreibbarkeit weitgehend entziehen.

Und im Fall von Geist geht dadurch bei der Simulation genau das verloren, was eigentlich simuliert werden sollte.

Warum Computer keinen Geist hervorbringen; die *metaphysische* Begründung

Die metaphysische Erklärung, warum Computer keine Qualia hervorbringen können, ist viel kürzer und einfacher als die formale:

Zustände von Computern sind Listen, die die Zeichen 0 und 1 enthalten. Ein Computerzustand ist also nichts anderes als eine *Zahl*, und die Dynamik eines Computers ist eine Folge von Transformationen von Zahlen.

Notwendige Bedingung für die Verwandlung eines physikalischen Zustands in ein Quale ist, dass seine Akzidenzien nicht als Funktion einfacherer Akzidenzien oder als Funktion äußerer oder innerer (nicht-geistiger) Bedingungen aufgefasst werden können. Nur dann kann die zweite Substanz EMPFINDUNG entstehen.

Diese Aussage muss an die Verhältnisse im Computer angepasst werden: die Akzidenzien sind dasjenige, woraus sich die Dynamik des Systems zusammensetzt, also das Programm bzw. die Transformationsregeln. Wenn das Programm *lernfähig* ist, können sich die Transformationsregeln ändern. Es gibt jedoch keine "einfacheren" Transformationsregeln, sondern nur "frühere".

Damit lässt sich die notwendige Bedingung für das Auftreten von Qualia in einem Computer formulieren:

Die Transformationsregeln dürfen keine Funktionen früherer Transformationsregeln, vorhergehender Zustände und des aktuellen Inputs sein.

Jeder Zustand des Computers ist aber eine Funktion des Inputs, des vorhergehenden Zustands und der darauf angewendeten Transformationsregeln, und dasselbe gilt für jedes Set von Transformationsregeln. Es ist eine Kette funktioneller Abhängigkeiten, die nie unterbrochen wird und bis zum Anfangszustand und dem anfänglichen Programm zurückführt.

Und daraus folgt, dass die Verwandlung in ein Quale niemals stattfindet. Es gibt keinen *Geist in der Maschine*.

5.10. Der metaphysische Unterschied zwischen Wirklichkeit und Simulation

Alle bisherigen Argumentationen über die Begrenztheit von Simulationen waren zumindest teilweise formaler Natur. Es existiert jedoch auch ein rein metaphysisches Argument, das auf dem Unterschied zwischen Wirklichkeit und Simulation beruht und aus dem sich eine prinzipielle Grenze für Simulationen ableiten lässt. Es lautet folgendermaßen:

Die Wirklichkeit besteht aus *Substanz und Akzidenzien*. Substanz ist das, *wovon* die Wirkungen des Seienden ausgehen, und zugleich das, was den Akzidenzien die metaphysische Qualität *Aktivität* verleiht.

Von einem wirklich existierenden System können auf eine Beschreibung oder ein Modell des Systems jedoch *nur die Akzidenzien* übertragen werden. Deshalb fehlt den Beschreibungen und Modellen diese Aktivität; *Von sich aus* sind sie *passiv*, sie bestehen nur aus Akzidenzien.

Eine Simulation ist aber nichts anderes als ein Modell der Wirklichkeit, das auf irgendeine Weise *aktiviert* wird. Auch eine Simulation ist also nicht von sich aus aktiv, sondern bedarf einer äußeren Aktivierung

Betrachten wir zur Illustration das System Sonne, Erde und Mond:

Es kann ein Modell der drei Körper angefertigt werden, ein sogenanntes *Tellurium*. Durch geeignete mechanische Vorrichtungen (Zahnräder, Wellen, Ketten usw.) kann die Bewegung von Erde und Mond annähernd nachgeahmt werden, aber nur dann, wenn dem Modell von außen Aktivität zugeführt wird, z.B. durch eine Kurbel oder einen Elektromotor. *Von sich aus* ist es *passiv*.

Die Bewegungen der Körper können auch im Computer berechnet und dann auf einen 3D-Schirm übertragen werden. Aber auch das in der Form von Gleichungen und Anfangsbedingungen und den zugehörigen näherungsweise Lösungsverfahren im Computer implementierte Modell ist *passiv*; erst wenn der Strom eingeschaltet wird, beginnt die Simulation zu laufen.

Definieren wir also:

Eine Simulation ist die Rekonstruktion der Dynamik eines Systems A in einem anderen System B, das sich von A auf folgende Weise unterscheidet:

Die Dynamik von A ergibt sich aus der Substanz und den Akzidenzien der Objekte von A, d.h. die Objekte von A sind von sich aus und ihrem Wesen gemäß aktiv und erzeugen so die Dynamik von A.

Die Dynamik von B ergibt sich jedoch nicht aus der Substanz und den Akzidenzien der Objekte von B, sondern aus der von uns geplanten und ausgeführten Konstruktion von B.

Metaphysisch ausgedrückt: Die Objekte von B dienen nur als materielle Basis konstruierter Akzidenzien, die somit ohne zugehörige Substanz auftreten. Da diesen Akzidenzien die Substanz fehlt, sind sie nicht von sich aus aktiv. Die Aktivität muss dem System B daher von außen zugeführt werden.

Diese Definition lässt sich wieder am Beispiel des Telluriums veranschaulichen:

Im wirklichen System Sonne, Erde und Mond bewegen sich die Objekte *von selbst* und ihrem Wesen gemäß, also aufgrund ihrer Gravitation.

In der Simulation ist die wesensgemäße Aktivität der Modellkörper – die Gravitation, die sich als ihr Gewicht äußert – zwar vorhanden, aber sie trägt nichts zur simulierten Dynamik bei. Dafür ist die *von uns* erstellte Konstruktion des Telluriums verantwortlich – eben die Zahnräder, Wellen, Ketten usw. Das Modell bewegt sich nicht von selbst, sondern nur, wenn es angetrieben wird. Erst dann imitieren die Modellkörper die Bewegungen der Himmelskörper.

Aus dem soeben bestimmten Unterschied zwischen Wirklichkeit und Simulation folgt der wichtige

Satz

Bei der Simulation eines Seienden geht das Wesen dieses Seienden verloren.

Ein Tellurium imitiert zwar die Bewegungen der Himmelskörper, aber *nicht durch Gravitation*. Die metaphysische Einheit der Substanz *Masse* und des Akzidens *Gravitation* – das *Wesen des Seienden* im wirklich existierenden System – verschwindet in der Simulation. Hier erfolgt die Bewegung aufgrund konstruierter, *masseloser* Akzidenzien.

Ein Computerprogramm kann geistige Leistungen nachahmen, aber *nicht durch Geist*. Die metaphysische Einheit der Substanz *EMPFINDUNG* und des Akzidens *Bedeutung* – das *Quale* im wirklich existierenden neuronalen Netz – verschwindet im Fall des Computers. Die Imitation der geistigen Leistungen erfolgt aufgrund konstruierter, *empfindungsloser* Akzidenzien.

Wenn wir also ein System herstellen wollen, das Geist hervorbringt, dann kann dieses System keine *Simulation* von Geist sein.

Um es noch deutlicher zu sagen:

Es gibt keine Simulation von Geist, weil in der Simulation das Wesen des Seienden verloren geht. Vom Quale, dessen Wesen die Einheit von EMPFINDUNG und Bedeutung ist, bleibt in der Simulation nur die Bedeutung übrig, d.h. der Informationsgehalt.

Durch die Ebene der konstruierten Akzidenzien wird die metaphysische Einheit von Substanz und Akzidenzien aufgehoben, die in keiner Beschreibung enthalten ist, und über die nicht vollständig erfassbare Dynamik der Wirklichkeit wird ein Algorithmus gebreitet, eine *nur aus Akzidenzien* bestehende Schicht der Wirklichkeit. Die Substanz EMPFINDUNG hat hier keinen Platz.

Im Fall der Gravitation erscheint die Tatsache des Verschwindens der Substanz selbstverständlich: gleichgültig, wie perfekt der Zusammenhang der Objekte, deren Bewegungen die gravitative Dynamik nachahmen sollen, auch konstruiert sein mag – was die Objekte antreibt, *ist nicht* die (abgeleitete) Substanz *Masse*. Die Ursache ihrer Dynamik ist nicht Gravitation, und sie kann auch niemals zu Gravitation werden. Das Simulations-Szenario enthält keine Gravitation.

Im Fall von Geist verhält es sich genauso. Gleichgültig, wie genau die Simulation ist – was auch immer die Entitäten der Simulation antreibt, *ist nicht* die Substanz EMPFINDUNG, und es kann auch nicht dazu werden. Die Simulation enthält keine Empfindung. *Sie enthält keinen Geist.*

Korollar

Die Simulation von Geist ist nicht möglich.

Bei der evolutionären Entfaltung der Natur geht die metaphysische Qualität *Aktivität* von jeder Schicht des Seienden auf die nächsthöhere, komplexere über. Im letzten Schritt entsteht Geist durch die Verwandlung der ersten Substanz in die zweite. Aber diese Verwandlung der Substanz ereignet sich nur *für uns*. *An sich* sind Substanz und Akzidens in jeder Schicht des Seienden *ungetrennt*, und sie entfalten sich *als Eines*.

Bei einer Simulation sind die Akzidenzien, aus denen die Dynamik gebildet ist, substanzlos. Für die Entstehung von Geist wäre es jedoch erforderlich, dass zu diesen Akzidenzien eine Substanz gehört, die sich in EMPFINDUNG verwandelt. Diese Substanz kann aber nicht einfach "entstehen". Damit

sich die Verwandlung ereignen könnte, müsste ebendiese Substanz im System schon vorhanden sein. Die Substanz, die dort tatsächlich vorhanden ist – die der Elemente der Simulation – ist aber nicht die zu den Akzidenzien der Simulation gehörende Substanz, und deshalb kann sie sich nicht in EMPFINDUNG verwandeln und mit den Akzidenzien vereint zum *Quale* werden. Die Simulation bringt also keine Empfindung hervor, mit anderen Worten: *sie empfindet nichts*.

Empfindung ist kein Akzidens, keine Eigenschaft eines Systems oder seiner Elemente. Sie kann nicht *konstruiert* werden, sondern nur aus der wesensgemäßen Aktivität des Seienden hervorgehen.

Das bedeutet, dass es bei der Schaffung von Geist notwendig ist, die Ebene substanzloser, konstruierter Akzidenzien zu vermeiden. Die Objekte, deren Dynamik zur Entstehung von Geist führen soll, müssen dies aufgrund ihres Wesens leisten, d.h. durch ihre eigenen Akzidenzien und mittels der Aktivität, die diesen Akzidenzien durch ihre Substanz verliehen wird.

Es ist somit nicht möglich, Geist durch eine auf Silizium oder Materialien mit vergleichbaren Eigenschaften basierende Technologie zu erzeugen. Geist kann von uns nur in Form einer "künstlichen Evolution" geschaffen werden. Dadurch wird die Schaffung von Geist in eine entfernte, vielleicht unerreichbare Zukunft verschoben.

Addendum

Die Hoffnung, Computer könnten irgendwann in der Zukunft zu Empfindungen fähig sein, beruht auf der Verwechslung eines *wirklich existierenden Systems* mit der *zeitlichen Entwicklung seines Zustands*.¹⁴

Diese Verwechslung zeigt sich schon am Begriff selbst: im alltäglichen Gebrauch bezieht sich der Ausdruck "Zustand" auf etwas Existierendes. Der physikalische "Zustand" eines Systems ist jedoch bloß eine Liste von Zahlen – die Werte der Systemvariablen –, und eine Computersimulation ist die sukzessive Transformation dieser Zahlen. Es ist dabei gleichgültig, auf welche Weise die Liste codiert wird und welches Speichermedium verwendet wird. Ebenso gleichgültig ist es, wie die Transformation durchgeführt wird und auf welcher Art von Maschine sie stattfindet. Es bleiben doch immer nur *Zahlen*, die transformiert und gespeichert werden. Es bleibt immer eine *Zustandsfolge*, eine zeitlich veränderliche Liste von Zahlen, und sie wird niemals identisch mit dem *wirklich existierenden System*.

¹⁴ Zur Erinnerung: Als *Zustand* eines physikalischen Systems bezeichnet man die Kombination der Werte aller Variablen des Systems zu einem bestimmten Zeitpunkt.

Die in diesem Abschnitt gegebene Definition der Simulation als einer Dynamik aus konstruierten Akzidenzien ohne Substanz wird durch dieses Bild der Computersimulation konkretisiert und veranschaulicht. Hier wird verständlich, was *konstruierte, substanzlose Akzidenzien* sind, die eine materielle Basis benötigen, und was es im Gegensatz dazu bedeutet, dass das wirkliche System sich aufgrund der *wesensgemäßen Aktivität* seiner Elemente entwickelt, die aus der *untrennbaren Einheit von Substanz und Akzidenzien* hervorgeht:

In der Simulation sind die Variablenwerte auf einer materiellen Basis gespeichert. Sie werden zu einem Rechenwerk transportiert, neu berechnet und dann zurücktransportiert, um als Output für ein Momentbild der Simulation zu dienen. Die Variablen stehen hier nicht für die Akzidenzien des Speichermediums oder des Rechenwerks – sie sind vom Rechenwerk und vom Speicher trennbar und gehen mit diesen nur zum Zweck der Simulation eine vorübergehende Verbindung ein. Rechenwerk und Speicher sind also nicht die zu diesen Variablen gehörende Substanz, sondern nur deren materielle Basis. Eine materielle Basis kann sich aber nicht in EMPFINDUNG verwandeln – das kann nur eine Substanz, die untrennbar mit Akzidenzien verbunden ist, aus denen komplexere, nicht-ableitbare Akzidenzien entstehen.

Das wirklich existierende System ist nicht bloß eine zeitlich veränderliche Liste von Zahlen. Im wirklichen System gibt es keine Zahlen, die in einen Speicher geschrieben, transformiert oder auch wieder gelöscht werden können, sondern *Eigenschaften von Dingen*, die mit diesen Dingen untrennbar verbunden sind und aus denen sich die Dynamik des Systems entfaltet. Im Fall menschlicher neuronaler Netze führt das zur Bildung nicht-ableitbarer Akzidenzien, zur Verwandlung der zugehörigen Substanz und damit zur Entstehung von Geist.

Existenz ist mehr als eine *Zustandsfolge mit Übergangsregeln*. Intuitiv ist das vermutlich (fast) jedem Menschen klar. Wäre die Verwechslung von Wirklichkeit und Beschreibung in unserer naturwissenschaftlich-technisch ausgerichteten Kultur nicht schon so weit fortgeschritten, dann könnte die Vorstellung der Computersimulation geradezu als Paradigma für die Empfindungslosigkeit von Simulationen dienen, weil es ja offensichtlich ist, dass die Gesamtheit der Akte der Variablenberechnung und -speicherung niemals eine Empfindung hervorbringen kann.

Solange man aber von dieser intuitiv einsichtigen Tatsache keinen Begriff hat, ist sie letztlich *unbegründbar*. Genau das war bisher der Fall, und daraus resultierte die Unklarheit und Verwirrung bei der Frage, ob Simulationen verstehen und empfinden können. Erst durch die begriffliche Bestimmung und Analyse des Unterschieds zwischen Wirklichkeit und Beschreibung sowie zwischen Wirklichkeit und Simulation, die wir hier durchgeführt haben, wird Klarheit geschaffen; Daraus lässt sich ableiten, dass Computer nicht zu Empfindungen fähig sind und dass sie niemals Geist hervorbringen können, gleichgültig, in welchem Maß ihre Leistung noch gesteigert wird.

Zuletzt noch eine Bemerkung zu künstlichen neuronalen Netzen. Wenn sie durch *Software* auf einem konventionellen Rechner realisiert werden, dann gilt für sie alles, was soeben über Computersimulationen gesagt wurde. Was ist aber mit künstlichen Netzen, die bereits als solche konstruiert, also durch *Hardware* realisiert sind? Können sie die zuvor bestimmte Grenze für Simulationen durchbrechen? Sind sie zur Bildung von Geist fähig?

Nein. Unser metaphysisches Argument greift auch in diesem Fall: Immer dann, wenn die Elemente eines Systems *von uns konstruiert* sind, ist damit eine Ebene *substanzloser Akzidenzien* geschaffen. Somit fehlt die Substanz, die sich in die zweite Substanz EMPFINDUNG verwandeln könnte. Die Entstehung von Empfindung und Geist findet nicht statt.

Das ist ein rein metaphysisches Argument. Lässt es sich auch *formal* interpretieren? Ich denke ja: Die Wirklichkeit ist nicht determiniert; die Zukunft ist nicht in der Gegenwart enthalten. Daher existiert *Kausalität von oben*, und das bedeutet, dass sich durch Selbstorganisation komplexe, eigenständige Entitäten bilden können, die eine neue Schicht der Wirklichkeit formen, mit einer eigenen Dynamik, die nicht aus der Dynamik einer darunter liegenden Schicht ableitbar ist. Somit können diese komplexen Entitäten als *Ursache* des Geschehens aufgefasst werden – genau so, wie es bei *geistigen Zuständen* der Fall ist.

In einem von uns konstruierten System bleibt dagegen die Dynamik immer durch *Kausalität von unten* vorgegeben. Die zeitliche Entwicklung des Systemzustands ist vollständig durch Regeln festgelegt. Komplexe Aggregate von einfacheren Elementen bilden keine neue, eigenständige Systemebene. Ihre Dynamik ist durch die von uns konstruierte Dynamik des Systems bestimmt. *Geist* als eigenständiges Phänomen kann also nicht auftreten. Es ist nicht möglich, Geist zu *konstruieren*.¹⁵

Der Übergang von einem wirklich existierenden System zu einer Simulation, d.h. zu einem konstruierten System, ist gleichbedeutend mit einer *Verarmung* des Systems. Der Reichtum an Möglichkeiten der Gestaltbildung, der im *Wesen der Wirklichkeit* begründet liegt – das darin besteht, *Substanz und Akzidens* zu sein – verschwindet in der Konstruktion. Sie besteht *nur aus Akzidenzien*, also fehlt ihr die metaphysische Qualität *Aktivität*, und das bedeutet, dass eine Konstruktion – ebenso wie eine Beschreibung – die Weise, wie die Wirklichkeit die Zukunft erzeugt, nicht nachahmen kann.

Das Gewebe der Wirklichkeit ist von unendlicher Feinheit. Es kann durch unsere Konstruktionen nicht nachgebildet werden.

¹⁵Das gilt auch für den Fall, dass die Konstruktion nicht-algorithmisch ist. Auch dann wird durch die metaphysische Argumentation das Entstehen von Geist ausgeschlossen. Das metaphysische Argument ist also stärker als das formale.