

<b>Zeitschrift:</b>	Studia philosophica : Schweizerische Zeitschrift für Philosophie = Revue suisse de philosophie = Rivista svizzera della filosofia = Swiss journal of philosophy
<b>Herausgeber:</b>	Schweizerische Philosophische Gesellschaft
<b>Band:</b>	52 (1993)
<b>Artikel:</b>	Über die Unerforschlichkeit des Bösen nach Kant
<b>Autor:</b>	Lichtenberger, Hans Peter
<b>DOI:</b>	<a href="https://doi.org/10.5169/seals-883074">https://doi.org/10.5169/seals-883074</a>

#### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

#### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

#### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 02.02.2026

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

HANS PETER LICHTENBERGER

## Über die Unerforschlichkeit des Bösen nach Kant

Es gehört zu den meist als selbstverständlich hingenommenen Grundorientierungen des abendländischen Denkens über das Böse, es nicht nur aus dem Gegensatz zum Guten auffassen zu wollen, sondern insbesondere auch aus dem Gegensatz zur Vernunft. Die platonische – oder platonistische – Gleichung von «bonum» und «ratio» bleibt auch für Kant der Horizont seines Nachdenkens über das Böse. Faszinierend jedoch ist, dass Kant von der Dynamik seiner Untersuchung dahin gezogen wird, wider Willen diese Gleichung zu unterminieren. Die Kantische These vom Vernunftursprung des Bösen gibt nicht nur dem Bösen einen Rang, den es in Konzepten, die es als Mangel an Einsicht oder als Dominanz der Sinnlichkeit verstehen wollten, nicht haben konnte, – diese These wirft zugleich auf die Vernunft einen Schatten, der der späteren Frage vorarbeitet, ob denn die Vernunft noch Herr im eigenen Hause sei. So gerät die Kantische Abhandlung über das Böse in die Ambivalenz, die sie als ein Experiment erscheinen lässt, dessen Ergebnisse mehrdeutig sind: Sie ist nur möglich auf der Basis des kritischen Ansatzes, sie stellt sogar hinsichtlich des Verständnisses des Subjekts dessen Radikalisierung dar, aber sie stellt zugleich wichtige Elemente dieses Ansatzes in Frage. Der transzentalphilosophische Ansatz – die Frage nach den Bedingungen der Möglichkeit – endet bei seiner Durchführung angesichts des Bösen im Desaster.

Der Status der Religionsschrift, deren erstes Stück die Abhandlung über das Böse bildet, ist schwer zu bestimmen. Sie fällt aus der Unterscheidung zwischen Grundlegungsschriften, die sich auf die transzendentale Untersuchung eines Vernunftvermögens bzw. die apriorischen Voraussetzungen von Wissenschaft richten, und den Schriften zur angewandten Philosophie, die das Materiale empirischen Ursprungs mitverarbeiten, offensichtlich heraus. Sie muss als ein eigenständiges Stück Transzentalphilosophie genommen werden, da sie neue transzendentale Analysen zur Anthropologie vorträgt, die über den Kreis der drei Kritiken hinausreichen. Dabei ist zweifellos irritierend die Art und Weise, in der moralphilosophische Grundsatzewägungen kombiniert werden mit Momenten der empirischen Menschenkunde, der Rechts- und Geschichtsphilosophie sowie der Phänomenologie der Religion.

Auch die inhomogene literarische Gestalt des Textes sowie die lose Gliederung erschweren ein eindeutiges Verständnis.

Angesichts auffälliger Inkonsistenzen gerade in der Abhandlung über das Böse ist einer bequemen Deutung die Richtung gewiesen durch Goethes bekannte Sottise vom «beschlabberten Philosophenmantel», der die Christenmenschen zum Saumküsselfen herbeilocken solle<sup>1</sup>. Demnach verdankten sich die Unstimmigkeiten einer prekären und unaufrechten Amalgamierung von Transzentalphilosophie und christlicher Dogmatik, die der greise Kant in Erinnerung an seine pietistische Herkunft versucht habe. Eine solche Lesart wird hier genauso abgelehnt wie jene, die gegenüber den Kritiken ein Nachlassen der systematischen Kraft feststellen will<sup>2</sup>. Vielmehr möchte ich dafür plädieren, in dem scheinbar lockeren Aufbau des Werks sowie einer Textstruktur, die vermeintlich undiszipliniert zwischen transzentaler Argumentation, historischem Belegmaterial und pragmatischer Anthropologie pendelt, ein reflektiertes Vorgehen Kants zu sehen, dem Phänomen des Bösen in seiner Uneindeutigkeit und Unbegreiflichkeit gerecht zu werden. Kant ist hier auf ein Problem gestossen, das sich notwendig systematischer Präsentation entzieht. Schärfer noch, die Anerkennung des Bösen als eines Problems der Transzentalphilosophie muss dazu führen, die Grundlagen dieses Denkens selbst in Frage zu stellen. Dies aber muss zwangsläufig auf die Form der Darstellung zurückschlagen. Die Abhandlung über das Böse impliziert eine radikale Selbtkritik Kants an seinem moralphilosophischen Ansatz sowie an der damit verbundenen Postulatentheologie. Die Mehrdeutigkeit der Religionsschrift resultiert aus dieser Ambivalenz: Sie bedient sich der Basis des kritischen Ansatzes, um ihn in wichtigen Teilen zugleich fraglich zu machen. Die transzendentale Untersuchung nach den Bedingungen der Möglichkeit des Bösen untergräbt den bislang ge-

1 Vgl. Goethe, Brief an Herder vom 7. 6. 1793.

2 Zur detaillierten Auseinandersetzung mit der Forschungsgeschichte fehlen hier Raum und Gelegenheit. Über die Rezeption bis Anfang der 70er Jahre dieses Jahrhunderts informieren zwei Literaturberichte: H.-O. Kvist, Das radikale Böse bei Immanuel Kant, in: *Makarios-Symposium über das Böse*, hg. von W. Strothmann, Wiesbaden 1983, S. 237-288; R. Malter, Zur Kantliteratur 1970-72. Neue Bücher zu Kants Rationaltheologie und Philosophie der Religion, in: Sonderheft der *Kant-Studien* 1974. *Akten des 4. Internationalen Kant-Kongresses* Mainz 1974, S. 155-177. An neueren Arbeiten sind zu nennen: K. Konhardt, Die Unbegreiflichkeit der Freiheit. Überlegungen zu Kants Lehre vom Bösen, in: *Zeitschr. f. phil. Forschung* 42 (1988), S. 397-416; H. Hoping, *Freiheit im Widerspruch. Eine Untersuchung zur Erbsündenlehre im Ausgang von Immanuel Kant*, Innsbruck-Wien 1990; G. E. Michalson, *Fallen freedom. Kant on radical evil and moral regeneration*, Cambridge 1990. Besonders perspektivenreich: Ch. Schulte, *radikal böse. Die Karriere des Bösen von Kant bis Nietzsche*, 2. Aufl. München 1991.

sicherten Bauplan der Vernunft. Die aus der Frage nach dem Bösen erwachsende Religionsauffassung gewinnt damit neues Gewicht: Sie dient der Sicherung der Vernunft angesichts ihrer aktuellen Selbstzerstörung als ratio perversa.

Der Versuch, den Begriff des Bösen hinreichend zu bestimmen und dieses in seinen Ursprüngen aufzuklären, stösst an die Grenzen von Theorie und Darstellung. Kant hat dies sehr wohl bedacht und durch ein kunstvolles Geflecht von Antithesen, die eine oberflächliche Lesart als Inkonsistenzen brandmarken könnte, dokumentiert. Die nach dem Durchlaufen eines differenzierten Untersuchungsganges zum Resultat werdende These von der Unerforschlichkeit des Bösen drückt sich methodisch in unaufgelösten antinomischen Bestimmungen aus<sup>3</sup>.

## I.

Das Böse ist kein explizites Thema der vorangegangenen moralphilosophischen Hauptschriften Kants wie auch der nachfolgenden *Metaphysik der Sitten*. Doch war Kant für das Faktum, dass beständig vom Sittengesetz abgewichen wird, in der Grundlegung und Durchführung seiner Ethik keineswegs blind. Im Gegenteil, der vielberufene Rigorismus seiner Moralphilosophie ist von der Tatsächlichkeit des Bösen so tief durchdrungen, dass er geradezu als Reaktion auf die Einsicht gelten kann, dass die schärfste Anfechtung der Möglichkeit von Ethik überhaupt im Faktum des Bösen besteht.

Es war Kants kopernikanische Wende in der Moral, dass die Bestimmung des Guten nur eine formale sein kann, dass sie bar ist jeder inhaltlichen Vorgabe, aus welcher Quelle sie auch stammen möge. Es ist lediglich die reine Form des Wollens des vernünftigen Subjekts, nach der eine Handlungsweise als gut zu gelten hat. Der kategorische Imperativ beurteilt Handlungen nach der möglichen allgemeinen Gesetzmässigkeit ihrer Maxime. Das Prinzip,

3 Dies wird weiter unten erläutert. – Es wäre reizvoll, die antinomische Begriffsstruktur des Bösen zu vergleichen mit der Methodik der Antinomienproblematik in den Kritiken. Dabei springt ins Auge: Während in den Kritiken die Antinomien – pauschalisierend gesagt – gelöst bzw. erklärt werden durch die Dimensionsunterscheidung von empirischer und intelligibler Welt, bleiben die Antinomien im Begriff des Bösen unaufgehoben stehen. Zwar lassen sie sich auch auf die genannten Pole beziehen, doch umreissen sie nur in ihrer Unaufgelöstheit den Begriff des Bösen. Dies indiziert, dass auch die Zuordnung der beiden Welten in der Religionsschrift eine andere sein muss. Davon später.

nach dem Handlungsmaximen als gut beurteilt werden, gibt somit zugleich ein formales Kriterium für das Böse als einer Abweichung von der Norm des Sittengesetzes. In der formalen Bestimmung des Guten wie des Bösen sah Kant durchaus ein «Paradoxon der Methode»: «dass nämlich der Begriff des Guten und Bösen nicht vor dem moralischen Gesetze (dem es dem Anschein nach so gar zum Grunde gelegt werden müsste), sondern nur [...] nach demselben und durch dasselbe bestimmt werden müsse.» (*KpV* A 110)<sup>4</sup>

Das Bild der Kantischen Ethik ist massgeblich davon geprägt, dass in den moralphilosophischen Grundschriften primär nach den Bedingungen der Möglichkeit des Guten gefragt wird. Doch auch dort bereits begleitet das Problem des Bösen die Untersuchung wie ein zunächst nicht wahrgenommener Schatten oder wie eine mühsam verdrängte Anfechtung. Dies zeigt sich zunächst schon in der Konzeption des kategorischen Imperativs. Die Tatsache, dass das Sittengesetz uns in der Form eines Imperativs entgegentreten muss, macht deutlich, dass wir nicht über einen «heiligen Willen» verfügen, sondern dass die Neigung zum Abweichen ständig präsent ist. Die problematische Erörterung der «Triebfedern» der reinen praktischen Vernunft (*KpV* A 126ff.) fasst klar ins Auge, dass zwischen der Reinheit des Sittengesetzes und der Motivation eines endlichen Willens eine kaum zu überbrückende Kluft besteht. Schliesslich ergänzt das Konzept des höchsten Gutes den reinen Willen, der ohne die Vorgabe materialer Zwecke, die sowohl moralisch als auch sinnlich sind, orientierungslos bleiben müsste. Das höchste Gut lässt sich – bezogen auf die sinnliche Natur des Menschen – geradezu als ein den moralischen Rigorismus auf nicht ganz einwandfreie Weise ergänzendes Gegengewicht zur Versuchung des Bösen verstehen. In all diesen Fällen jedoch kann die Abweichung vom Sittengesetz mit dem Einfluss sinnlicher Neigungen oder sonstiger heteronomer Antriebe erklärt werden.

Es wird in der Folge noch deutlich hervortreten, weshalb die Auffassung, nach der das Böse nur eine Abweichung vom Guten ist, der Realität des Bösen nicht gerecht wird. Diese Auffassung kennzeichnet den platonischen Traditionstrom des abendländischen Denkens, dem auch Kant sich in seinen Grundlagen zugehörig wusste. Ein Denken, das im Bösen den Mangel des Guten erblickt, verfügt über Erklärungsmuster, die in ihrer Erklärungsleistung zugleich das Böse bagatellisieren. Sie werden der – unmittelbar gewissen – Realität des Bösen nicht gerecht. Es kennzeichnet den Realitätssinn

4 Kants Schriften werden nach der Ausgabe von Wilhelm Weischedel zitiert: I. Kant, *Werke in 10 Bänden*, Darmstadt 1968. Die Zitate werden nach der Originalpaginierung lokalisiert.

Kants, dass er in weiser theoretischer Inkonsequenz das jedem Menschen gewisse, aber nicht theoretisch deduzierbare Faktum des Bösen gegen jede philosophisch-systematische Domestizierung ernst nimmt. Er fragt daher nach den Bedingungen der Möglichkeit des Bösen mit der gleichen Insistenz, mit der er zuvor nach den transzendentalen Voraussetzungen der Newtonschen Naturwissenschaft gefragt hat. Allerdings stellt sich nicht die gleiche Befriedigung einer Lösung ein.

Die Heteronomie-Vorstellung läuft Gefahr, das Böse durch seine Erklärbarkeit moralisch nicht nur zu neutralisieren, sondern sogar zu entschulden. Vielmehr muss das Böse, der moralischen Zurechenbarkeit wegen, als ein Akt der Freiheit begriffen werden. So wie das Gesetz einzig in der Form der Maxime begründet ist, so kann auch das Böse nur einer bewusst gewählten Maxime entstammen. D.h., ich entschliesse mich, anderen Antriebskräften als dem Sittengesetz dominierenden Einfluss auf meine Maximenbildung zu gewähren. Das aber bedeutet, dass das Böse in der Freiheit wurzelt – nicht in der Heteronomie, und dass es seinen Ursprung in der Vernunft hat, nicht in der Sinnlichkeit.

Diese Verankerung des Bösen in der Maximenbildung hat eine doppelte Konsequenz: Das Böse ist strikt auf das autonome Individuum bezogen und es kann nur als moralisches Böses, nicht aber als physisches Übel gedacht werden. Als Handlung von Kollektiven, als Struktur gesellschaftlicher Verhältnisse oder als krankhafte Persönlichkeitsstörung tritt es so wenig auf wie als sinnloses Naturgeschehen oder Leiden. Diese Einengung im Begriff des Bösen ist dem Verzicht auf die Theodizeefrage geschuldet. Subjekt und Urheber des Bösen ist allein der verantwortliche Mensch, kein Gott, keine Ordnung des Kosmos. Die Versuche in der Theodizee unterliegen der Gefahr, das Böse durch ihre Erklärungen zu verharmlosen.

Der Konflikt zwischen Gut und Böse spielt sich nicht zwischen verschiedenen Vermögen des Menschen ab oder zwischen seiner phänomenalen und noumenalen Natur, sondern er ist ein Widerstreit der praktischen Vernunft mit sich selbst. Als ein solcher aber muss er für dieses Vernunftkonzept ruinös werden<sup>5</sup>.

5 Auch wenn Kant in Fortsetzung der Religionsschrift vom Sieg des guten Prinzips über das Böse spricht, so ist jener Zwiespalt nicht aufgehoben, da das Böse unaustilgbar bleibt. Es macht die Würde des moralischen Individuums aus, unter dem Sittengesetz stets den Kampf gegen das ihm innewohnende Böse zu führen.

## II.

Kants Rede vom Bösen nimmt zwar die Entgegensetzung von phänomenaler und noumenaler Welt in Anspruch, doch unterläuft sie diese auch. Als Erklärungsmodell wird sie nicht in Anspruch genommen. Vielmehr ist ein seinem theoretischen Status nach schwer zu bestimmendes Wechselspiel zwischen Erfahrungsgesetzmäßigkeit der Wirklichkeit des Bösen und seiner intelligiblen Wurzel in der Freiheit zu konstatieren, ein Changieren zwischen transzendentaler Begründungssuche und empirischer Anthropologie. Das Böse wird so zum Indiz, dass die reinliche Unterscheidung und Zuordnung beider Welten nicht funktioniert.

Obzwar das Böse nicht aus der Gegensetzung von phänomenaler und noumenaler Welt erklärt werden kann, dient diese doch als Explikationsmittel. Weder ist das Böse auf der Seite apriorischer Begründungen festzumachen noch auf der Seite pragmatischer Anthropologie. Die Uneindeutigkeit des Bösen liesse sich allenfalls in einer Grauzone zwischen intelligibler und sinnlicher Welt lokalisieren. Ausgeschlossen ist die Variante heteronomer Bestimmung, nach der die Sinnlichkeit die Oberhand über die Vernunftstruktur gewinne; nicht diskutiert wird in dem Kantischen Ansatz die heute näherliegende Alternative, wonach das Böse in einer hybriden Überschreitung der Vernunft über ihre Naturbasis bestehen könnte.

Die neue grundlegende Antithese, aus der alle weiteren Bestimmungsmomente des Bösen entstammen, ist die von «ursprünglicher Anlage zum Guten» und «natürlichem Hang zum Bösen» (*RiG* B 19ff.). Hier wird auf beiden Seiten die Zusammengehörigkeit beider Welten, obzwar als eine prinzipiell störbare, gedacht.

Kant gibt drei zusammengreifende Anlagen des Menschen zum Guten an, die notwendig insgesamt zur Bestimmung des moralischen Wesens gehören: die Tierheit im Menschen, die Anlage für die Menschheit als eines vernünftigen Wesens und drittens die Anlage für die «Persönlichkeit als eines vernünftigen und zugleich der Zurechnung fähigen Wesens» (*Rig* B 15; vgl. B 18f.). Ohne Zweifel liegt der Hauptakzent auf dem dritten Punkt, denn «die Anlage für die Persönlichkeit ist die Empfänglichkeit der Achtung für das moralische Gesetz als einer für sich hinreichenden Triebfeder der Willkür» (B 18).

Kants Schema der Anlagen zum Guten reicht von der Animalität bis zur Moralität. Während die ersten beiden Anlagen zugleich in «Laster» umschlagen können, kann auf die dritte «schlechterdings nichts Böses gepropft werden» (B 18). Damit wird klar: Keine natürliche Anlage ist von sich aus

Wurzel des Bösen. Das Böse liegt nicht in der Natur, sondern im freien Umgang mit dem, was wir von Natur sind. Sofern die menschliche Natur durch die Anlage zum Guten bestimmt ist, kann das moralisch Böse nicht analytisch im Begriff der menschlichen Natur vorhanden sein. Doch macht der die Potentialität markierende Begriff der Anlage zugleich auch deutlich: Obzwar wir zum Guten veranlagt sind, werden wir doch nur wirklich gut durch einen bewussten Akt der Freiheit. Bei aller Kontinuität der drei Anlagen ist damit die Anlage der Persönlichkeit ausgezeichnet vor den beiden anderen, die eher unserer sinnlichen und endlichen Konstitution geschuldet sind. So ist die Möglichkeit des Bösen zwar in dem Anlageschema als Missbrauch angelegt, sie folgt jedoch nicht aus ihm. Die Unterscheidungen von Vernunft und Sinnlichkeit, Freiheit und Natur, Autonomie und Heteronomie entfallen als vertraute Modelle für die Herkunft des Bösen, da prinzipiell die natürlichen Anlagen solche zum Guten sind.

Kants anthropologische Lehre von der Anlage zum Guten reformuliert zwar in gewisser Weise die klassische Sichtweise der Zweiweltentheorie, relativiert sie aber auch zugleich. Denn wenn sowohl die natürliche wie die intellektuelle Anlage des Menschen gut ist, so lässt sich aus diesem Ansatz in keiner Weise die Möglichkeit der Wirklichkeit des Bösen herleiten. Damit entfällt die Chance, aus der Natur des Menschen einen präzisen Begriff des Bösen zu gewinnen.

Doch bedarf die unbestreitbare Tatsächlichkeit des Bösen einer anthropologischen Erklärung. Das Böse ist als Faktum gegeben und dem begreifenswollen Bemühen vorausgesetzt. Doch ist es nicht auf die gleiche Weise gegeben wie Gegenstände theoretischer Erkenntnis. Bereits die Feststellung seiner Faktizität führt ins Dilemma:

- a) Entweder gehen wir aus von der unmittelbaren Erfahrung des Bösen, z. B. sinnloser Grausamkeit, und verfügen nur über eine unmittelbare Evidenz, die begründungslos unserm Gefühl oder Urteilsvermögen entspringt;
- b) oder wir beurteilen Handlungen nach der Form der ihnen zugrundeliegenden Maxime. Dann bewegen wir uns ausschliesslich im Bereich der Willensbildung, nicht aber auf der Ebene der empirischen Handlung. Maximen sind aber grundsätzlich der sinnlichen Erfahrung entzogen.

Damit ergibt sich eine Spannung zwischen dem aktual Bösen, für das wir über keine Kriterien verfügen, und einem virtuell Bösen, für das wir über keine Erfahrung verfügen.

Diese Spannung bezeichnet die aporetische Situation, einen hinlänglichen Begriff des Bösen zu gewinnen. Kant versucht, diese inhomogenen Momente

in den Begriff eines «Hangs zum Bösen» zu vereinigen, einen Begriff, der darauf angelegt ist, den Abgrund zwischen sinnlicher Erfahrung und moralischer Beurteilung zu überbrücken. Der Begriff des «Hanges» muss also so angelegt sein, dass er eine Vereinbarkeit unserer empirischen Beschaffenheit mit der Theorie der Freiheit aufweist. Damit gerät der «Hang zum Bösen» zum genauen Gegenstück des kategorischen Imperativs, sofern dieser eben dieselbe Vereinbarkeit fordert: Handle so, dass du als empirisches Wesen deinem intelligiblen Charakter entsprichst. Diese Forderung wird im Hang verkehrt: Die empirische Beschaffenheit tritt als Maxime an die Stelle des intelligiblen Charakters. Wie dies aus Freiheit möglich ist, ist die entscheidende Frage.

Der Hang ist die bewusste Verkehrung der Maxime, in der das Subjekt seine Einheit in den beiden Welten zugunsten der natürlichen herstellt. Von der «Anlage» unterscheidet er sich dadurch, dass er die Zweiheit voraussetzt und zugleich zu unterlaufen sucht. «Unter einem Hange (propensio) verstehe ich den subjektiven Grund der Möglichkeit einer Neigung (habituelle Begierde, concupiscentia), sofern sie für die Menschheit überhaupt zufällig ist. Er unterscheidet sich darin von einer Anlage, dass er zwar angeboren sein kann, aber doch nicht als solcher vorgestellt werden darf, sondern auch (wenn er gut ist) als erworben oder (wenn er böse ist) als von den Menschen selbst sich zugezogen gedacht werden kann.» (RiG B 20f.)

In diesem Zitat sind die wesentlichen Punkte des Hanges angesprochen worden, die in der Folge noch entfaltet werden: subjektiv-zufällig, aber in der menschlichen Natur verwurzelt, d. h. allgemein; angeboren, aber erworben, bzw. zugezogen; natürlich, aber aus Freiheit. Hinzu kommt: Erfahrungstat-  
sache, aber intelligibler Ursprung. Diese Gegensatzpaare sollen nun kurz erläutert werden. Sie stehen insgesamt unter dem Zeichen, das antinomische Verhältnis von Natur und Freiheit, phänomenal und noumenal, überbrücken zu sollen.

Es ist eine Folge der Unmöglichkeit, das Böse in seiner empirischen Manigfaltigkeit einzugrenzen und zu bestimmen, die zu seiner Verlagerung in das Innere des Subjekts führt. Im Hang tritt ein Konstitutionsmoment des Subjekts auf, das sich weder der reinen moralischen Autonomie noch der natürlichen Beschaffenheit des Menschen verdankt, aber doch zu beiden in Beziehung gesetzt werden muss. Die dabei auftretenden Schwierigkeiten sind geeignet, die Idee des Menschen als einer moralischen Person zu erschüttern.

Der Hang als die Wurzel des Bösen wird gefunden im transzentalen Rückschluss von der Faktizität des Bösen auf die Bedingungen seiner Möglichkeit. Kants transzendentale Wende radikalisiert sich letztlich darin, dass

er das Böse ausschliesslich den Leistungen des Subjekts zuschreiben muss. Da die Anlagen des Menschen nur Anlagen zum Guten sind, muss der Hang – im Gegensatz zur Anlage – als «Tat» aufgefasst werden (*RiG* B 25f.).

Doch auch diese Kennzeichnung gerät ins Ambivalente: Einerseits muss der Hang den Charakter der Tat haben, um aus freier Willkür zu entspringen, andererseits scheint dieser Charakter der geforderten Allgemeinheit des Hangs zuwiderzulaufen. Dieser Widerspruch wird durchaus reflektiert: «Dagegen versteht man unter dem Begriffe eines Hanges einen subjektiven Bestimmungsgrund der Willkür, der vor jeder Tat vorhergeht, mithin selbst noch nicht Tat ist» (B 25).

Von Tat muss also in einem doppelten Sinne die Rede sein: als die böse Handlung, die ein Individuum in der Zeit vollzieht, und als intelligible Tat, die jenseits von Zeitbestimmungen in der Vernunft stattfindet als der Entschluss der freien Willkür zum Bösen. Sie ist keine Disposition oder Gewohnheit, noch durch heteronome Quellen verursacht. Sie ist nicht zeitlich fixierbar, sondern hat ihren Ursprung in der Vernunft. «Eine jede böse Handlung muss, wenn man den Vernunftursprung derselben sucht, so betrachtet werden, als ob der Mensch unmittelbar aus dem Stande der Unschuld in sie geraten wäre.» (B 42) Obwohl diese intelligible Tat nicht in Terminen des Empirischen beschrieben werden kann, muss sie als ein Faktum der menschlichen Existenz gelten. Die Figur der intelligiblen Tat macht das Böse denkbar, aber sie erklärt es nicht. Könnte sie dieses leisten, so höbe sie die Freiheit auf. Der postulierte Vernunftursprung des Bösen bleibt damit so unerforschlich wie das Faktum der Freiheit. Die in der Explikation dieses Sachverhalts ins Spiel gebrachte Zweiweltenlehre dient weniger einer Deduktion des Bösen als vielmehr seiner Verrätselung.

Auch in weiteren Punkten führt Kant das notwendige Scheitern von Problemlösungsversuchen vor. Ist der Mensch von Natur aus böse, so ist die Allgemeinheit dieser intelligiblen Tat in jedem Exemplar der Gattung aufzuzeigen. Da das Böse aus dem Begriff des Menschen nicht *a priori* abzuleiten ist, bleibt nur ein Hinweis auf die Erfahrung. Dieser aber ist kein Beweis. Zwar ist richtig, dass wir von bösen Handlungen – (woher wissen wir, dass diese Handlung böse ist?) – auf böse Maximen schliessen können, doch bleibt der Erfahrungsbeleg für diese These notwendig ungenügend. Dieses Versagen aber ist prinzipiell notwendig: Ein zureichender Beweis verbietet sich wegen der Freiheit des Menschen, die die Zufälligkeit und Zurechenbarkeit des Menschen impliziert. Es ist also lediglich der Erfahrungsdruck – nicht nur in der Weltbeobachtung, sondern auch in der Selbstwahrnehmung –, der die Annahme der Allgemeinheit des Bösen motiviert. Die transzendentale Re-

flexion, die zur intelligiblen Tat führt, erreicht jedoch nur seine Möglichkeit und Zufälligkeit.

Das gleiche gilt für die Antithese «zugezogen – angeboren» (B 8). Um der Verantwortlichkeit des Menschen willen, die in der Autonomiephilosophie begründet ist, muss das Böse als selbst zugezogen betrachtet werden. Es resultiert aus dem freien Entscheid, die Rangordnung der Maximen zu verkehren. Dieser Entscheid unterliegt keiner Notwendigkeit, er könnte auch anders ausfallen. Doch «radikal» ist das Böse nur, wenn es als unlösbar in der menschlichen Natur verwurzelt gesehen wird. Diesen Aspekt bezeichnet der metaphorische Ausdruck «angeboren». Er kollidiert zwangsläufig damit, dass das Böse nur der freien Zuständigkeit geschuldet ist (vgl. B 27). Eine transzendentale Herleitung für das angeborene Böse jedoch ist ein Ding der Unmöglichkeit. Die Universalität des Bösen, soviel auch für sie sprechen mag, ist weder durch Erfahrung noch durch transzendentale Reflexion zu erhärten.

### III.

Ob Kant selbst die Lehre vom radikal Bösen als Ergänzung oder Kritik seiner Moralphilosophie verstanden hat, sei dahingestellt. Da diese Lehre aus offenen Problemen der Moralphilosophie erwächst und zugleich wichtige Grundlagen derselben in Frage stellt, scheint einiges dafür zu sprechen, sie als Kantische Selbstkritik zu nehmen.

Die aufgeführten Begriffspaare in der Bestimmung des Bösen bezeichnen die Aussichtslosigkeit des Untersangens, das Böse terminologisch und systematisch präzise einzugrenzen. Es ist anzunehmen, dass dieses antithetische Verfahren bewusst gewählt ist, um das Scheitern der Vernunft an der Unerforschlichkeit des Bösen zu demonstrieren. Weder lässt es sich in der transzentalen Fragestellung erreichen noch als Erfahrungsdatum in seiner Allgemeinheit dartun. Die leitenden Begriffe pendeln zwischen den Sphären von Freiheit und Natur, intelligibler und empirischer Welt, ohne sie indes wirklich vermitteln zu können. Der Hang zum Bösen ist sowohl empirisches Faktum als auch intelligible Tat, ohne dass beide Aspekte in ein geklärtes Verhältnis rücken. Das Böse dementiert den früheren Anspruch auf die Koordination oder gar Harmonie beider Seiten.

Damit ist ein Grundpfeiler der Kantischen Systematik brüchig geworden. Hatte das Konzept des höchsten Guts die Systemeinheit noch in der Übereinstimmung von Natur und Freiheit postuliert, so fragt Kant nun nach dem

Ursprung ihrer tatsächlichen Unvereinbarkeit. Da auch im subjektiv sittlichen Willen das Böse als Maxime nie auszuschliessen ist – da es ja der Freiheit des Willens sich verdankt – droht die gesamte moraltheologische Postulatenlehre illusionär zu werden. Das Böse, obzwar aus Freiheit entsprungen, doch nicht vernunftgemässer Gesetzlichkeit unterliegend, entzieht dem Vernunftglauben an eine moralisch geleitete Zweckordnung der gegenwärtigen oder der zukünftigen Welt den Boden. Mit dem Wegfall der scharfen Unterscheidung von phänomenaler und noumenaler Welt hält das Reich der Natur maskiert Einzug im Reich der Freiheit: Wenn heteronome Motive aus Freiheit gewählt werden können, wenn Naturbestimmtheit aus freiem zurechenbarem Entschluss ergriffen wird, dann ist die reinliche Sphärentrennung hinfällig – damit aber auch die Möglichkeit einer geordneten Synopse beider Bereiche. Der Konflikt von Gut und Böse hat die frühere Diastase von Natur und Freiheit überlagert. Er ist aber nicht deckungsgleich mit ihm, denn auf beiden Seiten – der natürlichen Anlage zum Guten wie dem natürlichen Hang zum Bösen – findet sich das Verhältnis von Natur und Freiheit.

Dass durch das Böse das bisher angestrebte Verhältnis von Natur und Freiheit aus den Fugen gerät, zeigt sich auch darin, dass das Böse geradezu gegenbildlich zum höchsten Gut gefasst wird. Die Einheit aus unterschiedlichen Prinzipien, die im höchsten Gut gedacht wurde, tritt im Bösen in charakteristischer Verkehrung auf.

Es war das grundlegende Konstruktionsprinzip für das höchste Gut, dass das moralische Moment der natürlichen Erfüllung vorangeht. Allein aus dieser Reihenfolge kann sich die Forderung einer nicht-absurden Welt ergeben. Das Böse, das seinen Ursprung im Vermögen zum Guten, der Freiheit, hat, kann daher durch keine material neuen Bestimmungskomponenten ausgezeichnet werden. Es ist keine dualistische Gegenmacht zum Guten, sondern resultiert aus dem Prinzip des Guten selbst. So lässt sich das Böse formal umschreiben als die verkehrte Rangordnung der sittlichen und natürlichen Elemente des Handelns bzw. der Maximenbildung. «Also muss der Unterschied, ob der Mensch gut oder böse sei, nicht in dem Unterschiede der Triebfedern, die er in seine Maxime aufnimmt (nicht in dieser ihrer Materie), sondern in der Unterordnung (der Form derselben) liegen: welche von beiden er zur Bedingung der andern macht. Folglich ist der Mensch (auch der beste) nur dadurch böse, dass er die sittliche Ordnung der Triebfedern in der Aufnahme derselben in seine Maximen umkehrt.» (B 34)

Das Böse ist also nicht nur seiner Verfassung nach der genaue Gegenbegriff zum höchsten Gut; es stellt auch seine sachliche Möglichkeit als Einheit von Tugend und Glückseligkeit, Freiheit und Natur in Zweifel. So verfügt auch

die *RiG* nicht mehr über den starken Begriff des höchsten Guts, wie er im kritischen System entwickelt wurde. Der praktisch motivierte Vernunftglaube an die Wirklichkeit des Guten verliert angesichts der Universalität des Bösen jeglichen Halt. Damit schwächt sich das Moment der Glückseligkeit ab, da es von vornherein unter dem Verdacht steht, durch Selbstliebe die Ordnung der Maximen zu verkehren. Die mit dem Sittengesetz verbundene Hoffnung auf Glückseligkeit ist obsolet geworden. Das höchste Gut ist reduziert auf die sittliche Aufgabe der Herbeiführung einer ethischen Gemeinschaft, einer «Gesellschaft nach Tugendgesetzen und zum Behuf derselben» (*RiG* B 129). Die bisherige Konzeption des höchsten Gutes wird wegen ihrer eudämonistischen Konnotationen, die sich vom Verdacht des Bösen nicht trennen lassen, ausgeschieden. Das höchste Gut ist nur noch moralischer Gemeinschaftszweck. «Jede Gattung vernünftiger Wesen ist nämlich objektiv, in der Idee der Vernunft, zu einem gesellschaftlichen Zwecke, nämlich der Beförderung des höchsten als eines gemeinschaftlichen Guts bestimmt.» (*RiG* B 135) Diese Beförderung des höchsten Guts ist zugleich der Prozess der Überwindung des Bösen, der wegen dessen Radikalität nie vollendbar ist. Auch wenn Kant hypothetisch an dieses Konzept des höchsten Guts beiläufig die Gottesidee anschliesst, so hat sie doch jede konstitutive Funktion verloren.

Es gehört zu den verwundernswerten Zügen der Kantischen Religionsschrift, dass gerade in ihr der Gottesgedanke nahezu völlig zurücktritt. Ihr Moralkonzept und ihre Vorstellung vom höchsten Gut können nicht mehr als Brücke zur Idee Gottes dienen.

Ein weiterer, bereits gestreifter Aspekt der Lehre vom Bösen kann die neue Kraftlosigkeit des Gottesbegriffs illustrieren. Im gesamten kritischen System diente der Gottesbegriff der Aufgabe einer Einheit von Natur und Freiheit. Diese Grundstruktur des gesamten Kantischen Ansatzes wird in der *RiG* hinfällig. Die Aufteilung des Menschen in ein Sinnes- und ein Verstandeswesen, die als platonische Tradition auch für Kant systembildend war, hat ihre Geltung angesichts des Bösen verloren. Der Mensch ist der existierende Widerspruch: Als ganzer gehört er zur Natur, als ganzer steht er unter dem Gesetz der Freiheit. Die Antithese von «Anlage» und «Hang», insbesondere die von «angeboren» und «selbst zugezogen», artikuliert diesen Zwiespalt. Zwar bleibt das Problem von Natur und Freiheit bestehen, doch kollabiert diese Unterscheidung, die der früheren Ethik so viel Evidenz verlieh, an der Tatsache des Bösen. Im Ursprung des Bösen – hier zeigt es sich wieder als Gegenbegriff zum höchsten Gut – sind Natur und Freiheit auf fatale Weise geeint und ununterscheidbar. Der «Hang» zum Bösen ist ein «natürlicher»,

der doch nur der Freiheit entspringt. Diese Sachlage ist angetan, den Begriff der Freiheit selbst zu untergraben. Damit stürzt ein weiterer Grundpfeiler der Kantischen Philosophie, nämlich die analytische Verbindung von Freiheit und Sittengesetz. Diese hatte in den moralphilosophischen Hauptschriften das Ziel, die Möglichkeit guten Handelns zu sichern, das Böse liess sich nur abstrakt-negativ als Abirren von der Norm, als Heteronomie der Neigungen qualifizieren. Es gehörte nicht in den Bereich der sittlichen Entscheidungsfreiheit. Zwei Motivkreise könnten Kant zur Revision dieser Auffassung angeregt haben.

Der eine dürfte im ergebnislosen Ringen mit der Triebfedern-Problematik liegen; der zweite – damit zusammenhängend – in einer verstärkten Reflexion auf die inneren Verwirklichungsbedingungen des Sittengesetzes, aus welcher hervorgeht, dass im konkreten Falle nie eindeutig entscheidbar ist, ob eine Handlung autonomen oder heteronomen Maximen entspringt. Mit anderen Worten: gefragt ist nicht nur die Erklärung moralischen Handelns, einer Aufhellung bedarf insbesondere die Möglichkeit nicht-moralischen Handelns.

Hierauf hatten die ethischen Hauptschriften keine befriedigende Antwort. Sie sind beherrscht von der Identität von Freiheit und autonomer Gesetzgebung des Willens. Das Moralgesetz als ein Fakt der Vernunft steht im Wechselverhältnis mit dem ebenfalls unableitbaren Begriff der Freiheit: «Also ist ein freier Wille und ein Wille unter sittlichen Gesetzen einerlei.» (GMS BA 98) Auf diese Weise ist Freiheit einzig als Freiheit zum Guten denkbar. Wenn die Freiheit unter dem Sittengesetz die einzige wirkliche Freiheit ist, verwandelt sich nicht dann der Freiheitsbegriff in einen Automatismus, der alternative Entscheidungen nicht zulässt, bzw. diese in den nicht zurechenbaren Bereich der Naturkausalität abschiebt? Lässt sich unter diesem Freiheitsverständnis die Aufnahme einer dem Sittengesetz widerstrebenden Maxime überhaupt denken?

Kants These: «Freiheit und eigene Gesetzgebung des Willens sind beides Autonomie, mithin Wechselbegriffe» (GMS BA 104f.), destruiert in der Konsequenz wider Willen den Freiheitsbegriff. Es fiele nicht nur nicht-moralisches Handeln unter das Verdikt der Unfreiheit, sondern auch moralisches Handeln verlöre seine Zurechenbarkeit, wenn der Mensch nicht auch anders sich entscheiden könnte. Sind Freiheit und Sittengesetz Wechselbegriffe, dann ist die Freiheit der Wahl aufgehoben.

Die Zurechenbarkeit nicht-moralischen Handelns kann einzig dann gedacht werden, wenn das analytische Verhältnis von Freiheit und Sittengesetz aufgebrochen wird. Die Faktizität des Bösen macht diesen Schritt unver-

meidbar. Damit löst die Lehre vom zurechenbar radikal Bösen das frühere Konzept heteronomer Bestimmung ab. Es muss selbst noch der Freiheit geschuldet sein, dass ich heteronomen Maximen den Vorrang vor dem Sittengesetz einräume. Das Böse besteht darin, dass ich mich aus freien Stücken zur Aufgabe moralischer Freiheit entscheide. Damit aber gerät der Begriff der Autonomie aus der Balance: die Autonomie ist nicht mehr durch die moralische Gesetzgebung gesichert, sondern gerät zur freien und unbegreiflichen Willkür. Erst in solcher Ungebundenheit ist sie Freiheit zum Guten wie zum Bösen.

Kant hat die Anerkennung der unbegreiflichen Realität des Bösen mit der Auflösung der Identität von Freiheit und Sittengesetz bezahlt. Dieser Preis ist sehr hoch, da er die Grundlage seiner Ethik zerstört. Letztlich besagt dieses Verständnis von Freiheit, dass Kant über kein zwingendes Argument mehr verfügt, weshalb ich mir das Sittengesetz auferlegen soll. Die Wahl einer «guten Gesinnung» wird so rätselhaft wie das Böse selbst, der freie Wille wird zum Geheimnis.

Die Lehre vom Bösen stellt eine Selbtkritik Kants dar, wie sie fundamentaler kaum vorstellbar ist. Sie folgt aus einer Radikalisierung des subjektivitätstheoretischen Ansatzes, in der die früher sorgsam austarierte Zuordnung von Gott, Welt, Mensch zugunsten der Subjektivität aufgegeben wird. Damit wird die Subjektivität sich nicht durchsichtiger: Vermochte das frühere System die Unbegreifbarkeit des Subjekts noch an den Widerlagern von Welt und Gott zu befestigen, so ist nunmehr das Subjekt mit dem unerforschlich Bösen als einem Konstituens seiner selbst alleingelassen.

Es sind nahezu alle Bauelemente des Kantischen Systems, die durch das Böse ihre Stabilität verlieren: das Verhältnis von Natur und Freiheit, von sinnlicher und intelligibler Welt, die Postulatentheologie mit ihrer Gotteslehre, schliesslich der Begriff moralischer Autonomie selbst.

Es gehört zu den tiefsten Überzeugungen Kants, dass die Welt nicht absurd ist. Das höchste Gut als Einheit von Tugend und Glückseligkeit war entworfen als Gegenpol zu diesem aus dem Weltlauf sich aufdrängenden Eindruck. An diese Voraussetzung wird nun gerührt. Macht nicht die Allgemeinheit und Unerforschlichkeit des Bösen die Welt als ganze unbegreiflich? Wo kann das Vertrauen auf die Rationalität noch einen Verankerungspunkt finden?

Die Lehre vom Bösen hat Sprengkraft für den gesamten Ansatz der Kantischen Transzentalphilosophie. Sie macht dies implizit, nicht programmatisch deutlich. Ob Kant selber sich die destruktiven Folgen seines Neuansatzes in vollem Umfang eingestanden hat, ist kaum zu entscheiden. Es hat eher den Anschein, als habe Kant in dieser Schrift in den Abgrund

geblickt – und sei wieder zurückgewichen. In den folgenden Schriften lassen sich kaum Spuren dieser neuen Einsichten ausmachen. Die *Metaphysik der Sitten* von 1797/98 macht eher den Eindruck, als wolle sie den Stand der *RiG* widerrufen. So beteuert die Einleitung erneut die von der *RiG* aufgegebene Identität von Sittengesetz und Freiheit<sup>6</sup>. Die bis ins Kasuistische reichende Institutionenlehre von Moral und Recht mit ihrer starken Betonung der Legalität kann als ein Versuch betrachtet werden, das Böse, das explizit nicht erwähnt wird, nicht nur durch moralische Gesinnung, sondern vor allem durch äussere Ordnung im Zaum zu halten.

6 Zu dieser erneuten Revision vgl. G. Prauss, *Kant über Freiheit als Autonomie*, Frankfurt/Main 1983, S. 111f.

