Zeitschrift: Studia philosophica : Schweizerische Zeitschrift für Philosophie = Revue suisse de philosophie = Rivista svizzera della filosofia = Swiss journal of philosophy

Herausgeber: Schweizerische Philosophische Gesellschaft

Band: 37 (1977)

Artikel: Über Möglichkeiten und Grenzen einer Konfliktethik

Autor: Hügli, Anton

DOI: https://doi.org/10.5169/seals-883204

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: 10.12.2025

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

Studia Philosophica 37/1977

ANTON HÜGLI

Über Möglichkeiten und Grenzen einer Konfliktethik*

Nach einer neuerdings wieder zu Ansehen gekommenen Auffassung von Ethik haben moralische Prinzipien den Zweck, individuelle und gesellschaftliche Interessenkonflikte zu harmonisieren. Dieser Artikel versucht den Beweis zu führen, dass eine solche «Konfliktethik» notwendigerweise scheitern muss: nicht aus anthropologischen Gründen (Unvollkommenheit der menschlichen Natur usw.), sondern schon aus rein logischen Gründen. Der Beweis wird in drei Schritten geführt: Es werden zunächst die formalen Bedingungen und die minimalen Rationalitätskriterien herausgearbeitet, denen eine Konfliktethik in jedem Fall zu genügen hätte, und es wird dann der Bereich der überhaupt in Frage kommenden substantiellen moralischen Prinzipien abgesteckt. In einem dritten Schritt schliesslich wird aufgezeigt, dass keines der möglichen Prinzipien den minimalen formalen Kriterien genügen und jede Konfliktethik mithin nur neue Konflikte erzeugen würde.

Es gibt einen alten philosophischen Traum, der beinahe so alt ist wie die Philosophie selbst: der rationalistische Traum einer moralisch geordneten Welt. Nach ihm unterstehen wir alle einem ewigen Gesetz, das da war, ehe es geschriebene Gesetze oder einen Staat gab, das durch keinen Volksbeschluss geändert oder abgeschafft werden kann und das unverbrüchlich für alle Völker und zu allen Zeiten gilt. Denn dieses Gesetz sei nichts anderes als die göttliche Vernunft selbst, die der Natur innewohnt, durch alle Vernunftwesen hindurch geht und ihr Tun und Lassen regelt. «Es muss», wie der Stoiker Chrysipp zu Beginn seines Werkes über den Nomos sagt, «die Autorität sein, die bestimmt, was sittlich schön und was hässlich ist, muss Herr sein und Führer für die von Natur zur staatlichen Gemeinschaft veranlagten Wesen, und demzufolge die Richtschnur geben für das, was gerecht und ungerecht ist, indem es befiehlt, was getan werden soll, und verbietet, was nicht getan werden darf»[1].

Heraklit hat diesen Traum als erster klar ausgesprochen – alle menschli-

* Dieser Artikel geht auf einen Vortrag zurück, den der Autor am 26. Januar 1976 vor der Philosophischen Gesellschaft Zürich gehalten hat.

[1] Chrysipp bei H. v. Arnim: *Stoicorum veterum fragmenta* (Leipzig 1903) III, 314.

Korrespondenz: Dr. Anton Hügli, Blumenstrasse 16, D-44 Münster.

chen Gesetze nähren sich von dem einen göttlichen[2] –, in Platos Staat erhielt er seine klassische Gestalt, durch die Stoa wurde er zum Allgemeingut, die Aufklärung erhob ihn zum politischen Programm, Kant verklärte ihn zur Idee, Hegel versicherte uns seiner Gegenwart im modernen bürgerlichen Staat, und Marx verhiess uns seine Verwirklichung in der kommunistischen Gesellschaft. Und wie steht es mit diesem Traum heute? Nach einer Periode positivistischer Nüchternheit und existentialischer Verzweiflung hat man erneut wieder zu träumen begonnen[3]: Habermas schickt sich an, nach einer neuen Hegelschen Identität des Allgemeinen und des Besonderen in der postindustriellen Gesellschaft Ausschau zu halten[4], und die Erlanger Schule übt bereits – unter der kundigen Führung von Lorenzen – die ersten Schritte zur Konstruktion einer dauerhaften moralischen Ordnung ein[5].

Der Inhalt des Traums, von dem ich bisher so pauschal gesprochen habe, hat sich im Verlaufe der Zeit geändert; seine formalen Wesenszüge aber sind sich gleich geblieben. Die folgenden vier Grundbedingungen einer moralischen Ordnung gehören zu den unangefochtenen Glaubensartikeln rationalistischer Philosophie: (1) Die moralische Ordnung ist *universal:* sie gilt nicht nur für alle Menschen und Plätze, sondern auch für alle vernünftigen Wesen. (2) Eine moralische Ordnung ist *total:* sie kennt die Antwort auf jedes zwischen vernünftigen Wesen aufkommende moralische Problem. (3) Eine moralische Ordnung ist *harmonisierend:* in ihr sind die Wünsche und Ziele aller Wesen aufeinander abgestimmt: sie bilden einen geordneten Kosmos, ein Königreich der Zwecke. (4) Eine moralische Ordnung ist *autonom:* sie ist – wie das Corpus mysticum Christi – keine äussere Ordnung, sondern eine Ordnung der Herzen. Jeder kennt

[2] Heraklit, B 114. In: Diels/Kranz: *Die Fragmente der Vorsokratiker* (⁶1951) I, 176.

[3] Vgl. W. Schulz: *Philosophie in der veränderten Welt* (Pfullingen 1972) 736: «Der Begriff Ordnung ist für viele immer noch mit dem Merkmal der Metaphysik und des Konservativen behaftet und daher suspekt. Gleichwohl setzt sich de facto gegenwärtig ein neues Ordnungsdenken durch ...»

[4] Vgl. etwa Habermas' programmatische Rede anlässlich der Verleihung des Hegelpreises der Stadt Stuttgart. In: J. Habermas/D. Henrich: *Zwei Reden* (Frankfurt a. M. 1974) 25-84; ausführlicher in: J. Habermas: *Zur Rekonstruktion des Historischen Materialismus* (Frankfurt a. M. 1976) 92-126.

[5] Vgl. etwa O. Schwemmer: *Philosophie der Praxis.* Versuch zur Grundlegung einer Lehre vom moralischen Argumentieren, in Verbindung mit einer Interpretation der praktischen Philosophie Kants (Frankfurt a. M. 1971). – Vgl. dazu meine kritische Besprechung: Zur Erlanger Philosophie der Praxis. In: *Studia Philosophica* 34 (1974) 199-222.

seinen Platz, und jeder folgt seinem eigenen Urteil, von niemandem geheissen und durch niemanden gezwungen.

Ich habe eine moralische Ordnung der Welt einen Traum genannt. Aber findet sich ein Beweis dafür, dass es keine solche Ordnung geben kann? Diese Frage ist hoffnungslos mehrdeutig. Was heisst, es «gebe» keine solche Ordnung? Dreierlei kann damit gemeint sein. (1) Die empirische Aussage, dass man bisher noch keine menschliche Gesellschaft gefunden habe, in der eine solche Ordnung verwirklicht ist. (2) Die empirische Unmöglichkeitsbehauptung, es werde – so wie die Natur des Menschen nun einmal sei – nie möglich sein, eine moralisch geordnete menschliche Gemeinschaft auszubilden. (3) Die logische Behauptung, eine moralische Ordnung sei nicht einmal als Ideal möglich, weil sich kein Normensystem konstruieren lasse, das die notwendigen Bedingungen einer solchen Ordnung auf konsistente Weise erfülle.

Es gehört zur Definition eines rationalistischen Philosophen, dass er an die logische Existenz einer moralischen Ordnung glaubt, und es ist eine philosophiegeschichtliche Tatsache, dass kein solcher Philosoph, nicht einmal Hegel, wenn man Ritters Interpretation folgen will[6], die moralische Ordnung faktisch schon für verwirklicht hielt. Das philosophische Hauptproblem der Rationalisten besteht daher darin, eine Erklärung dafür zu finden, warum das Ideal noch nicht Wirklichkeit geworden ist, und die Frage zu beantworten, ob es je wirklich werden könne. Nach der Erklärung für die moralische Unvollkommenheit der menschlichen Welt brauchten sie allerdings nicht lange zu suchen. Der Mensch und seine Welt sind nicht gut genug für die Moral. Was die moralische Ordnung braucht, sind Vernunftwesen, die nicht nur vollkommene ethische Einsicht besitzen, sondern auch vollkommenes Wissen und vollkommene Selbstlosigkeit. Aber den Menschen mangelt es an dem einen oder dem andern oder gar an mehrerem zugleich. Woran genau es ihnen fehlt, darüber gehen freilich die Ansichten auseinander und dementsprechend auch darüber, ob die menschliche Unzulänglichkeit je überwunden werden könne und mit welchen Mitteln sie zu überwinden sei. Falls wir Wesen sind von der Art der Miltonschen Areopagiten und der Millschen und Popperschen Liberalen, fehlbar in unserem Urteil und von unvollkomme-

[6] Vgl. Hegel und die französische Revolution (1956); abgedruckt in: J. Ritter: *Metaphysik und Politik*. Studien zu Aristoteles und Hegel (Frankfurt a. M. 1969) 183-255.

nem Wissen, aber voll selbstlosen Wahrheitsdranges, dann wird uns nur unaufhörliche Diskussion und Kritik zu Besserem führen. Falls wir aber aus demselben krummen Holz geschnitzt sind wie der kantische Mensch, mit vollkommenem Urteil zwar und vollkommenem Wissen, aber von unvollkommener Selbstlosigkeit, so hilft womöglich nur eines: die strenge Zucht des Gesetzes. Und falls die mangelnde Selbstlosigkeit nicht auf die menschliche Natur, sondern auf die gesellschaftlichen Verhältnisse zurückzuführen ist, so führt der einzige Weg zur Besserung über gesellschaftliche Veränderung und Revolution.

Ich habe indes nicht die Absicht, nun alle anthropologischen und geschichtsphilosophischen Kombinationen durchzurechnen, sondern möchte mich vielmehr dem ihnen zugrundeliegenden rationalistischen Grunddogma selber zuwenden: Sind die Mängel, die man der menschlichen Natur anzulasten pflegt, nicht vielleicht bei der Idee einer moralischen Ordnung selbst zu suchen? Wenn menschliche Vernunft und moralische Ordnung nicht zueinander passen wollen, liegt es dann tatsächlich an der Unvollkommenheit unserer Vernunft und nicht eher an der Unvollkommenheit der moralischen Ordnung?

Von den traditionellen rationalistischen Voraussetzungen her lässt sich diese Frage nicht einmal stellen, geschweige denn beantworten. Wer die Moral vor dem Forum der Vernunft zur Rechenschaft ziehen will, so würde ein klassischer Rationalist etwa argumentieren, der nimmt stillschweigend an, dass es eine aussermoralische oder zumindest eine moralisch neutrale Vernunft geben könne. Aber ist das Moralisch-Richtige je etwas anderes als das Vernünftige und das Vernünftige je etwas anderes als das Moralisch-Richtige?

Diese seltsame Identitätsthese enthält, genau besehen, nicht nur eine, sondern gleich zwei Behauptungen. (1) Es gibt eine Norm, die alle rationalen Wesen qua rationale Wesen anerkennen, d. h. die ihr Handeln bestimmt und der sie den Vorrang geben, wenn immer sie mit einem andern Ziel kollidiert. (2) Was rationale Wesen wollen, ist per definitionem das, was moralischerweise getan werden soll. – Beide Behauptungen sind gleichermassen fragwürdig.

Die erste Behauptung – ich möchte sie die aprioristische These nennen – besagt, die Anerkennung gewisser Normen sei eine notwendige und hinreichende Bedingung dafür, dass ein Wesen rational genannt werden könne. Ihre Zweifelhaftigkeit wird offenkundig, wenn man sich, wie dies

etwa C. D. Broad in bezug auf Kant getan hat[7], die zwei folgenden Fragen vorlegt: (a) Gibt es eine Norm, die jedes rationale Wesen anerkennen würde? (b) Gibt es eine Norm, deren Gültigkeit aus dem Begriff eines rationalen Wesens erschlossen werden kann? Die Antwort auf die zweite Frage ist zweifellos negativ: Ein rationales Wesen ist, grob gesprochen, ein Wesen, das fähig und willens ist, gewisse intellektuelle Operationen zu verstehen und auszuführen, z. B. Begriffe zu bilden, Aussagen zu machen, Schlüsse zu ziehen, Begründungen zu geben oder Rechtfertigungen beizubringen. Aber aus dem Begriff eines solchen Wesens folgt keineswegs, dass es irgendeinen Schluss als richtig, irgendeine Aussage als wahr oder irgendeine Norm als gültig anerkennen müsste, geschweige denn einen bestimmten Schluss, eine bestimmte Aussage oder eine bestimmte Norm. Positiv beantworten lässt sich hingegen die erste Frage: Eine Aussage oder eine Norm, *falls* sie wahr oder gültig ist, würde von einem rationalen Wesen auch als wahr oder gültig anerkannt werden. Ein rationales Wesen ist nicht dadurch definiert, dass es die Wahrheit des euklidischen Satzes einsieht, und dieser Satz ist nicht darum wahr, weil er von einem vernünftigen Wesen anerkannt wird, sondern umgekehrt: weil der euklidische Satz wahr ist, wird ihn ein vernünftiges Wesen auch anerkennen. Dasselbe gilt für die Anerkennung von Normen. Es ist durchaus möglich, dass es gültige Normen (was dies auch immer heissen mag) gibt, und aus dieser Tatsache folgte dann, dass sie von vernünftigen Wesen als gültig anerkannt würden, aber es gibt keine Norm, deren Gültigkeit, wie die aprioristische These behauptet, aus dem Begriff eines vernünftigen Wesens erschlossen werden könnte[8].

Noch offenkundiger ist die Fragwürdigkeit des zweiten Teils der rationalistischen Identitätsthese – was vernünftige Wesen wollen, sei das, was die Moral uns zu tun gebiete. Diese Behauptung ist ein typisches Beispiel

[7] Vgl. *Five Types of Ethical Theory* (London 1930) 125ff.

[8] Dieses Argument richtet sich auch gegen die sogenannte transzendentale Begründung ethischer Normen. So versucht etwa Apel die Grundnorm der «wechselseitigen Anerkennung» damit zu begründen, dass vernünftige Wesen «alle der sprachlichen Kommunikation fähigen Wesen» als Personen anerkennen müssten, «da sie in all ihren Handlungen und Äusserungen virtuelle Diskussionspartner sind». In: *Transformationen der Philosophie. II: Das Apriori der Kommunikationsgemeinschaft* (Frankfurt a. M. 1973) 400. Aber selbst wenn es zutrifft, dass die «logische Geltung von Argumenten» nur innerhalb einer Argumentationsgemeinschaft überprüft werden kann (S. 399), so folgt daraus keineswegs, dass jeder der sprachlichen Kommunikation Fähige als Mitglied dieser Gemeinschaft anerkannt werden müsste, sondern nur, dass keiner, der zur Überprüfung von Argumen-

jenes Fehlschlusses, der von Hume erstmals aufgedeckt[9] und von Moore als naturalistischer Fehlschluss bezeichnet wurde[10]. Das rein faktische oder deskriptive Urteil «ein rationales Wesen will *x*» wird gleichgesetzt mit dem Sollensurteil «*x* soll getan werden». Aber die Frage, was der Fall ist, ist völlig verschieden von der Frage, was der Fall sein soll, und die Antwort auf die erste Frage beantwortet nicht schon die zweite. Es mag zwar tatsächlich zutreffen, dass ein rationales Wesen immer nur das will, was getan werden soll, doch aus der Tatsache, dass es *x* will, folgt ebensowenig, dass *x* getan werden soll, wie aus der Tatsache, dass ein rationales Wesen *x* glaubt, auch schon folgte, dass *x* wahr ist. Dass rationale Wesen tatsächlich wollen, was getan werden soll, liesse sich nur durch zwei getrennte Untersuchungen erweisen: Man müsste erst feststellen, was getan werden soll, und dann nachweisen, dass ein rationales Wesen de facto auch tun würde, was getan werden soll. Mit einem Wort: Die These, dass das Vernünftige auch das Moralisch-Richtige sei, ist keine Voraussetzung, mit der man beginnen, sondern allenfalls ein Resultat, mit dem man enden kann; sie ist, technisch gesprochen, keine analytische, sondern eine synthetische Aussage.

Zusammenfassend also: Der Inhalt moralischer Prinzipien lässt sich nicht aus dem ableiten, was vernünftige Wesen «immer schon» tun. Rationale Wesen als solche wollen und tun überhaupt nichts. Sie bedürften deshalb auch keiner moralischen Ordnung und noch weniger einer Harmonisierung ihrer Zwecke. Das Problem einer moralischen Ordnung stellt sich nur für Wesen, die nicht bloss rational, sondern auch sinnlich und körperlich sind, verwundbar durch andere und fähig zu verwunden, weil sie in einer gemeinsamen körperlichen Welt zu handeln haben, in der sie sich gegenseitig daran hindern können, das zu sein, zu tun oder zu haben, was sie zu sein, zu tun oder zu haben wünschen.

Unter diesen und nur unter diesen Bedingungen kann die Frage nach den rationalsten Prinzipien der Konfliktlösung überhaupt aufkommen;

ten beiträgt oder beitragen kann, von ihr ausgeschlossen werden darf. Ob tatsächlich jeder, der sprechen kann, diese Bedingung auch schon erfüllt, ist jedoch eine empirische und keine apriorische Frage, und die faktischen Argumentationsgemeinschaften in Universitäten, Instituten und Laboratorien mit ihren Initialriten, Geheimsprachen und esoterischen Ritualen bieten jedenfalls – ob man dies nun bedaure oder nicht – eher das Bild einer geschlossenen als das einer offenen Gesellschaft.

[9] Vgl. *A Treatise of Human Nature*, hg. L. A. Selby/Bigge (Oxford 1888) III/1, 1, S. 469f.

[10] Vgl. *Principia Ethica* (Cambridge 1903) Kap. IV.

denn als rational handelnde Wesen werden sie die für sie beste Form der Konfliktlösung nicht nur für den Augenblick, sondern auf lange Sicht festzulegen versuchen.

Unter diesen Bedingungen wird freilich eine der historisch wirksamsten Antworten auf das Problem einer moralischen Ordnung von vornherein hinfällig: der spätestens in den Religionskriegen des 17. Jahrhunderts sich zersetzende naturrechtliche Gedanke der Antike, dass die Natur einem jeden seinen Platz und seine Aufgabe in dieser Welt zuweise und dass es keine Konflikte mehr gebe, wenn jeder das Seine tue und seine Rolle erfülle. Die Deduktion von Prinzipien aus einer «immer schon» bestehenden und nur noch zu entdeckenden idealen Ordnung der Natur fiele nicht nur unter das Verdikt des naturalistischen Fehlschlusses; sie machte auch unverständlich, wie vernünftige Wesen – und solche haben wir ja vorausgesetzt – miteinander in Konflikt bleiben können. Entweder haben sie die richtige Einsicht noch nicht gewonnen, und dann stehen sie mindestens bis dahin vor demselben Problem wie bisher, oder die Ordnung gibt nicht her, was sie verspricht. So oder so also: Die Frage nach der besten Lösung ihrer Konflikte bleibt weiterhin bestehen. Mehr noch: Es leuchtet auch keineswegs ein, warum diese Lösung in irgendeiner Art von moralischer Ordnung gesucht werden soll. Bevor man fragt, ob eine moralische Ordnung die beste Form der Konfliktlösung sein könne, muss man sich erst fragen, ob sie überhaupt eine Form der Konfliktlösung sein könne. Kann sie es nicht, so beantwortet sich – nach dem Sollen-impliziert-Können-Prinzip – auch die Frage, ob sie es soll. Prüfen wir daher ihr Können. Gibt es moralische Prinzipien, die eine dauerhafte Konfliktlösung garantierten?

I

Was macht eine moralische Ordnung überhaupt aus? Eine moralische Ordnung im Sinne unserer Bedingungen besteht in einem System von Normen oder Prinzipien: jener Prinzipien, auf die sich rational handelnde Personen unter den Bedingungen der Interaktion und der gegenseitigen Verwundbarkeit zur universellen Regelung ihrer Konflikte verpflichten würden. Sich einer moralischen Ordnung unterwerfen heisst, in Anlehnung an Kant gesprochen, sich auf eine Verhaltensweise festlegen, die man in jeder denkbaren Konfliktsituation und Position zu befolgen ge-

denkt und die man von allen befolgt sehen möchte, gleichgültig, welche sonstigen Absichten, Wünsche oder Ziele man jeweils auch haben wird. Aus dieser Bestimmung ergeben sich die wesentlichen formalen Merkmale eines moralischen Prinzips:

(1) Es ist handlungsleitend oder präskriptiv in dem Sinn, dass es nicht sagt, was ist, sondern was sein soll. Wer ein moralisches Prinzip äussert, verpflichtet sich selbst, direkt oder indirekt, auf eine Vorschrift oder Direktive in bezug auf aktuelle oder mögliche Entscheidungen. Entspricht sein Verhalten nicht dem Prinzip, so spricht dies, im Unterschied zu einem deskriptiven Prinzip, nicht gegen das Prinzip, sondern gegen sein Verhalten.

(2) Das Prinzip ist universalisierbar: Wer behauptet, er sei moralisch verpflichtet, *x* zu tun, sagt damit zugleich, dass auch jeder andere verpflichtet sei, *x* zu tun, ausser es sei eine relevante Differenz zwischen ihm und dem andern oder zwischen seiner Situation und der Situation des andern.

(3) Das Prinzip ist ein letztes Prinzip: Bei einem allfälligen Konflikt zwischen der Forderung des moralischen Prinzips und der eines andern Handlungsprinzips gilt die Forderung des moralischen Prinzips[11].

Nach dieser begrifflichen Klärung kann ich meine Frage etwas präziser fassen: Gibt es, wie die Anhänger der Idee einer moralischen Ordnung behaupten, ein universales Prinzip, das von rational handelnden Personen als höchste handlungsleitende Norm zur Regelung ihrer Konflikte anerkannt würde?

Welche Prinzipien gewählt würden, hängt wesentlich davon ab, welche Eigenschaften und Fähigkeiten die moralischen Gesetzgeber von sich als Adressaten dieser Prinzipien voraussetzen dürfen; denn zwischen der Situation, in der allgemeine Prinzipien diskutiert und formuliert, und der Situation, in der sie appliziert werden, besteht ein ebenso grosser Unterschied wie zwischen der Situation des Schreibtischstrategen und der des

[11] Mit diesen drei Bedingungen übernehme ich faktisch die metaethische Theorie des universalen Präskriptivismus von R. M. Hare: *Freedom and Reason* (Oxford 1963), aber ohne dessen sprachanalytischen Voraussetzungen. Ich behaupte nicht, meine Definition eines moralischen Prinzips sei eine korrekte Analyse der alltagssprachlichen Bedeutung von «sollen», sondern bloss, es sei eine adäquate Analyse eines bestimmten philosophischen Verständnisses von Ethik, das ich mit dem Begriff einer moralischen Ordnung zu kennzeichnen versuchte. Dass sich Hares Ergebnisse dabei so mühelos übertragen lassen, beweist allenfalls, in welchem Ausmass die Alltagssprache, oder vielleicht eher Hare selbst, im Banne dieser philosophischen Tradition steht.

Offiziers in der Schlacht. Was helfen die ausgeklügeltsten Prinzipien, wenn ihre Adressaten in der konkreten Situation nicht fähig oder nicht willens sind, die sich aus ihnen ergebenden Imperative abzuleiten. Mit der Möglichkeit, dass sie nicht willens sein könnten, sich an die von ihnen gesetzten Prinzipien zu halten, muss schon deshalb gerechnet werden, weil aus der Tatsache, dass alle, aufs ganze gesehen, den Zustand bevorzugen, in dem jeder sich an die vereinbarten Prinzipien hält, keineswegs folgt, dass es für den Einzelnen in der konkreten Situation nicht noch eine bessere Alternative geben könnte: den Zustand, in dem alle ausser ihm sich an die vereinbarten Prinzipien halten.

Welche Annahmen also sollen über die Adressaten der moralischen Ordnung getroffen werden? Die Antwort ergibt sich aus dem Ziel meiner Untersuchung. Um zu prüfen, ob es rationale Prinzipien der Konfliktregulierung geben könne, müssen wir von der Annahme idealster Voraussetzungen ausgehen; denn je vollkommener die Personen, die die moralischen Prinzipien zu applizieren haben, desto minimaler die Bedingungen, denen diese Prinzipien genügen müssen, und je minimaler die Bedingungen, desto aufschlussreicher das Resultat: Wenn sich kein Prinzip finden lässt, das minimalsten Bedingungen genügt, so wird sich auch kein Prinzip finden lassen, das schärferen Bedingungen zu genügen vermag.

Ideal sind die Voraussetzungen genau dann, wenn in jeder Konfliktsituation Einigkeit darüber besteht, wie der Konflikt zu entscheiden ist und jeder von jedem annehmen darf, dass er sich dem Entscheid beugen wird. Auf welchem Wege wir uns diese Voraussetzungen realisiert denken, ob durch die Annahme einer Gesellschaft von moralisch und intellektuell vollkommenen Personen oder durch die Annahme vollkommenster politischer und sozialer Institutionen, spielt dabei keine Rolle. Um des Argumentes willen soll den Adressaten der moralischen Prinzipien all das an Wissen, Urteilskraft und gutem Willen zugestanden werden, was wir an realen Menschen so häufig zu vermissen pflegen. Die Voraussetzung vollkommenen Wissens stellt sicher, dass jeder Adressat die für seine Entscheidungen erforderlichen Informationen hat; die Voraussetzung vollkommener Urteilskraft garantiert, dass jeder die Fähigkeit besitzt, universale Prinzipien in der jeweiligen konkreten Situation anzuwenden, und die Voraussetzung des guten Willens verbürgt, dass jeder sich selbst dann normgerecht verhält, wenn dies seinem Eigeninteresse zuwiderläuft, und dass jeder weiss, dass auch jeder andere sich so verhält. Die Voraussetzung

des guten Willens könnte freilich auch durch die Annahme ersetzt werden, dass die Konfliktentscheidung politischen oder rechtlichen Instanzen übertragen werden kann, welche genügend Macht und Mittel haben, die Entscheidung in jedem Fall durchzusetzen, aber weder die Macht noch die Möglichkeit, ihre Autorität für eigene Zwecke zu missbrauchen.

Meine volle Frage lautet also jetzt: Gesetzt, rationale, voll informierte, voll urteilsfähige, moralisch gesinnte oder zumindest legal handelnde Personen hätten sich, unter den Bedingungen der Interaktion und der gegenseitigen Verwundbarkeit, auf ein universales moralisches Prinzip zur Lösung oder Beseitigung ihrer künftigen Konflikte zu verpflichten; welches sind die minimalen Bedingungen, die sie an ein solches konfliktethisches Prinzip stellen würden? Gibt es ein Prinzip, das diese Bedingungen auf konsistente Weise erfüllt?

Prüfen wir zuerst den ersten Teil meiner Frage: Welche Bedingungen hätte das gesuchte konfliktethische Prinzip zu erfüllen?

Die ersten drei Bedingungen ergeben sich aus dem Umstand, dass das gesuchte Entscheidungsverfahren für alle denkbaren Konfliktparteien und in jedem Konfliktfall gültig sein soll.

1. *Bedingung der Einzigkeit:* Für jeden Konflikt gibt es ein und nur ein Entscheidungsverfahren.

Falls es mehr als ein Verfahren gäbe, bestünde die Möglichkeit, dass es zu verschiedenen Konfliktlösungen käme. Die eine Lösung könnte dann im Interesse der einen Partei, die andere im Interesse der andern Partei sein. Dies führte zu einem neuen Konflikt über das zu verwendende Verfahren, und es gäbe kein Verfahren, um diesen Konflikt zu beenden[12].

2. *Bedingung der Vollständigkeit:* Für jeden denkbaren Konflikt muss es eine Lösung geben.
3. *Bedingung der Eindeutigkeit:* Für jeden Konflikt darf es stets eine und nur eine Lösung geben.

Diese Bedingung, die sich mit demselben Argument wie Bedingung 1 begründen lässt, schliesst auch alle inkonsistenten Lösungen aus, denn wenn es nicht verschiedene Lösungen geben darf, kann es auch keine sich ausschliessende Lösungen geben.

[12] Die Einzigkeitsbedingung verlangt nicht, dass in jedem Konfliktfall ein und dasselbe Verfahren anzuwenden sei, sondern nur, dass niemals mehrere Verfahren zur Anwendung kommen dürfen, ohne dass feststeht, welches die letzte Entscheidungsinstanz ist.

Die restlichen zwei Bedingungen folgen aus der Tautologie, dass ein begehrendes Wesen seine Wünsche und Bedürfnisse so weit wie möglich zu befriedigen sucht.

4. *Bedingung der Neutralität:* Eine Konfliktlösung darf sich nicht ändern, wenn die Präferenzen der verschiedenen Konfliktparteien vertauscht würden (wenn z. B. A B's und B A's Präferenzen hätte). Technisch gesprochen: Eine Konfliktlösung muss invariant sein in bezug auf jede mögliche Permutation der Präferenzen der jeweiligen Konfliktparteien.

Diese Bedingung soll verhindern, dass eine der Parteien gegen die Interessen der andern den Ausgang des Konfliktes diktieren kann oder in irgendeiner andern Form bevorzugt oder privilegiert wird. Ihre Notwendigkeit lässt sich leicht einsehen: Jeder möchte zwar, dass er selbst oder die von ihm bevorzugten Individuen oder Gruppen bei einem möglichen Konflikt den Vorrang haben, aber keiner kann hoffen, dass diese Privilegierung von den andern akzeptiert würde. Die einzige Möglichkeit, die eigenen Interessen zu sichern, ist daher die, ein Entscheidungsverfahren zu wählen, das die Interessen eines jeden auf gleiche Weise berücksichtigt. – Die Diskrimination bestimmter Individuen oder Gruppen könnte jedoch selbst dann nicht zur allgemeinen Bedingung des gesuchten Entscheidungsverfahrens werden, wenn ihr faktisch alle, Diskriminierte und Diskriminierende, Opfer und Nutzniesser, in gleichem Masse zustimmten, denn als allgemeine Bedingung müsste sie nicht nur in bestimmten, sondern in allen Konfliktfällen anwendbar sein, also nicht nur, wenn die privilegierten oder diskriminierten Gruppen oder Individuen zufälligerweise Konfliktpartei sind, sondern auch, wenn sie es nicht sind.

Keines weiteren Beweises bedarf die folgende und letzte Bedingung. Sie verbietet, dass den jeweiligen Konfliktparteien eine Lösung aufgezwungen wird, die gegen die Interessen aller Beteiligten verstösst.

5. *Bedingung der Effizienz:* Falls alle Parteien eines Konflikts die Lösung *A* der Lösung *B* vorziehen würden, darf die Konfliktentscheidung nicht in der Lösung *B* liegen.

Jedes dieser fünf Postulate umschreibt eine notwendige, aber nicht hinreichende Bedingung für die Rationalität eines konfliktethischen Prinzips. Als rational möchte ich ein Prinzip dann und nur dann bezeichnen, wenn es alle diese fünf Bedingungen erfüllt.

Mit dieser Definition ist der Rahmen unserer Untersuchung abgesteckt,

und wir können nun endgültig zur eigentlichen Frage übergehen: Gibt es ein rationales konfliktethisches Prinzip? Eine Antwort auf diese Frage ist zweifellos nur möglich, wenn wir die Gewissheit haben können, alle denkbaren Prinzipien in Betracht gezogen zu haben. Ich werde daher zunächst versuchen, einige Klassifikationen vorzunehmen, um so die Klasse der überhaupt in Frage kommenden Prinzipien einzugrenzen.

Die möglichen Konfliktlösungsverfahren lassen sich in zwei Hauptgruppen unterteilen: mittelbare und unmittelbare Verfahren. Von *mittelbaren* Verfahren spreche ich dann, wenn das Prinzip bloss das Verfahren nennt, nach dem Konflikte gelöst werden sollen, und erst die Anwendung des Verfahrens den Konflikt entscheidet; von einem *unmittelbaren* Verfahren hingegen, wenn der Konflikt aufgrund des Prinzips allein entschieden werden kann, d.h. wenn die Anwendung des Prinzips selbst schon das Verfahren darstellt. Zu den mittelbaren Verfahren gehören die meisten der gängigen Konfliktlösungsverfahren: Konfliktentscheidung durch die Ausübung oder Androhung von Gewalt, durch Abstimmung oder Verhandlung, durch eine übergeordnete Autorität oder durch die Verwendung von Zufallsmechanismen wie das Los oder die Lotterie. Beispiele für unmittelbare Verfahren hingegen finden sich selten: Am typischsten wohl ist der Appell an die Spielregeln zur Entscheidung über die Zulässigkeit oder Unzulässigkeit von Handlungen oder Zügen innerhalb eines Spiels.

Die drei formalen Bedingungen moralischer Prinzipien – und daran zeigt sich ihre Formalität – sind an sich mit jedem Konfliktlösungsverfahren verträglich, vorausgesetzt, dass es universalisiert und zum Prinzip erhoben wird. Die Devise etwa, Konflikte nur durch Gewalt oder durch die Androhung von Gewalt zu lösen, ist ein ebenso mögliches moralisches Prinzip wie der Grundsatz, Konflikte stets durch das Los zu entscheiden.

Sobald man aber die Rationalitätsbedingungen mit ins Spiel bringt, scheiden die mittelbaren Verfahren unweigerlich aus. Entscheidungsverfahren, die nach irgendwelchen undurchschaubaren Kriterien Entscheidungen produzieren, verstossen gegen mindestens zwei dieser Bedingungen: die der Effizienz und die der Neutralität. Denn es besteht keine Gewähr, dass Lösungen erzeugt werden, die auch bei einer Vertauschung der Präferenzen gleich blieben und niemals gegen die gemeinsamen Interessen der Konfliktparteien verstiessen. Falls es rationale Entscheidungsverfahren überhaupt gibt, können sie deshalb nur unter den unmittelbaren Verfahren gefunden werden. Dieses Ergebnis besagt allerdings nicht, dass

Konflikte nicht weiterhin mittelbar gelöst werden könnten, sondern nur, dass in letzter Instanz immer unmittelbar entschieden werden muss, wenn ein Konflikt durch die anderen Verfahren nicht entschieden werden kann oder diese Verfahren selbst zum Gegenstand von Konflikten werden.

Unmittelbare Konfliktlösungsverfahren erfordern von den Akteuren jene Art von praktischer Überlegung, der sie – als rational Handelnde – auch ausserhalb jeder Konfliktsituation immer schon folgen. Die konfliktethischen Prinzipien haben daher dieselbe Funktion wie die Prinzipien rationaler Wahl. Hier wie dort geht es darum, aus den jeweils offenstehenden Handlungen jene auszuwählen, die die gegebenen Ziele der Handelnden am «besten» zu realisieren verspricht: bei den Prinzipien rationaler Wahl die im Augenblick gegebenen individuellen Ziele, bei den konfliktethischen Prinzipien die langfristigen Ziele aller. Ein kurzer Blick auf die Prinzipien rationaler Wahl ist daher wohl am geeignetsten, uns über das Wesen der gesuchten konfliktethischen Prinzipien Klarheit zu verschaffen.

II

Welches sind die Prinzipien eines rational Handelnden? Dass diese Frage nicht so leicht zu beantworten ist, beweist die Tatsache, dass dazu ganze Wissenschaftszweige ins Leben gerufen worden sind: die Wahrscheinlichkeitstheorie etwa und insbesondere die erst in den letzten Jahrzehnten entwickelte rationale Entscheidungstheorie und mathematische Spieltheorie[13]. Für unsern Zweck genügt es, einige der elementarsten Grundbegriffe rationaler Wahl zu erörtern. Die schwierige Frage, ob die

[13] Das Standardwerk ist nach wie vor: J. von Neumann/O. Morgenstern: *Theory of Games and Economic Behavior* (Princeton 1944, ²1947). Eine der besten Expositionen der neueren Entwicklung der Spiel- und Entscheidungstheorie geben R. D. Luce und H. Raiffa: *Games and Decisions* (New York/London/Sidney 1957). Eine gute Einführung unter Berücksichtigung der neuesten Literatur gibt auch W. Stegmüller: *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie IV/1: Personelle Wahrscheinlichkeit und rationale Entscheidung* (Berlin/Heidelberg/New York 1973) 285-385. – Lesenswert ist auch R. C. Jeffrey: *Logik der Entscheidungen* (Wien/München 1957). – Zur Anwendung der Entscheidungstheorie vgl. etwa W. Krelle: Anwendung der Entscheidungstheorie auf soziale Probleme. In: H. Scholz (Hg.): *Die Rolle der Wissenschaft in der modernen Gesellschaft* (Berlin 1969) 159-171; H. Hax (Hg.): *Entscheidungen bei unsicheren Erwartungen.* Beiträge zur Theorie der Unternehmung (Köln/Opladen 1970); G. Gäfgen: Formale

Prinzipien rationaler Wahl als präskriptiv oder deskriptiv zu nehmen seien[14], wird sich für uns nicht stellen: Da wir rational Handelnde als Akteure voraussetzen, können diese Prinzipien als Beschreibung ihres faktischen Verhaltens verstanden werden.

Ein rational Handelnder im Sinne der Entscheidungstheorie muss im wesentlichen den folgenden Rationalitätskriterien genügen: (1) Er muss die Konsequenzen der ihm jeweils offenstehenden Handlungsalternativen kennen, d. h. er muss wissen, was geschieht, wenn er nicht handelt, und was geschehen wird, wenn er handelt. (2) Er muss eine Präferenzordnung für seine Handlungskonsequenzen erstellen können, d. h. er muss von jeder Handlungskonsequenz sagen können, ob er sie – im Vergleich mit einer beliebigen andern – als besser, schlechter oder gleich gut einstuft. Die Präferenzordnung muss transitiv sein (wenn die Konsequenz *a* mindestens so gut ist wie *b*, und *b* mindestens so gut wie *c*, so muss *a* mindestens so gut sein wie *c*), und sie muss reflexiv sein (jede Alternative muss mindestens so gut ist wie *b* und *b* mindestens so gut wie *c*, so muss *a* mindestens so ordinal wie kardinal ausgedrückt werden, d. h. man kann die Konsequenzen entweder mit Hilfe eines gewöhnlichen Vergleichs ordnen (*x* ist besser als / gleich gut wie *y*), oder man kann jeder Konsequenz eine bestimmte Zahl auf einer Präferenzskala zuordnen und dadurch auch die Intensität der Präferenzen ausdrücken.

Wenn eine Konsequenz höher in der Präferenzordnung steht als eine andere, so wird von ihr gesagt, dass sie einen höhern Wert oder, in der üblichen Terminologie ausgedrückt, einen höhern *Nutzen* habe. Die Hauptmaxime eines rational Handelnden lautet daher kurz: Wähle jene Handlung, die den grössten Nutzen bringt! Oder noch kürzer: Maximiere den zu erwartenden Nutzen!

Um Missverständnissen vorzubeugen: Ein in diesem Sinn rational Handelnder muss keineswegs ein kalt berechnender Egoist sein, bar jeder Sympathie und jeden Mitgefühls. Er kann durchaus auf das Wohl anderer bedacht sein, und unter seinen Zielen können sich die schönsten Ideale

Theorie des politischen Handelns: Wissenschaftliche Politik als rationale Wahl von Strategien. In: H. Maier u. a. (Hg.): *Politik und Wissenschaft* (München 1971) 209-257. – Eine gute Kritik der ethischen Relevanz der Entscheidungs- und Spieltheorie vor allem aufgrund pragmatischer Überlegungen findet sich in O. Höffe: *Strategien der Humanität.* Zur Ethik öffentlicher Entscheidungsprozesse (Freiburg/München 1975) Teil 1.

[14] Vgl. etwa W. Stegmüller, a. O. (Anm. 13) 298-306.

und die hehrsten Prinzipien befinden. Das Wort «Nutzen» hat zwar einen gewissen hedonistischen Beigeschmack und ruft Erinnerungen wach an den Glücksbegriff von Bentham, Mill und Sidgwick. Aber die Maxime «Maximiere deinen Nutzen» hat mit Glück, Vergnügen und ähnlichem mehr nicht das Geringste zu schaffen. Sie gilt auch für einen rigorosen Kantianer, der keinen andern Wunsch hat als den, die Pflicht um der Pflicht willen zu tun. Die Handlung mit dem höchsten Nutzen ist dann eben jene, die ihm die Pflicht gebietet.

Berechtigter dagegen scheint ein anderer Vorwurf: Das Prinzip der Nutzenmaximierung, so wird gelegentlich eingewendet, sei tautologisch oder zumindest von einer nicht zu überbietenden Simplizität. Dieser Eindruck ändert sich freilich rasch, wenn man sich das erste Rationalitätskriterium – die Kenntnis der Handlungskonsequenzen – etwas genauer ansieht. Nur in den seltensten Fällen können wir die Konsequenzen unserer Handlungen mit Sicherheit voraussagen; in den meisten Fällen wissen wir – selbst bei bester Information – nicht, ob der Sachverhalt, den wir mit unserem Handeln herbeizuführen oder zu verhindern suchen, ohne unser Zutun tatsächlich eintreten oder unterbleiben würde. Die Entscheidungstheorie unterscheidet daher drei Arten von Entscheidungen: Entscheidungen unter Sicherheit, Entscheidungen unter Risiko und Entscheidungen unter Unsicherheit. Bei Entscheidungen unter Sicherheit glaubt der Handelnde, die Konsequenzen seiner Handlungen mit Sicherheit voraussagen zu können. Bei Entscheidungen unter Risiko besteht zwar keine Gewissheit, aber der Handelnde ist immerhin in der Lage, den möglichen Umständen und den möglichen Konsequenzen seines Tuns gewisse Wahrscheinlichkeiten beizumessen. Bei Entscheidungen unter Unsicherheit kennt er die Situation so schlecht, dass er die verschiedenen möglichen Konsequenzen seiner Handlungen nicht einmal probalistisch beurteilen kann.

Für Entscheidungen unter Sicherheit mag die Formel «Maximiere den zu erwartenden Nutzen» als Handlungsmaxime ausreichen, für die andern Entscheidungsarten bedarf sie jedoch der Interpretation; denn die Schwierigkeit liegt hier darin, die Handlung mit dem höchsten möglichen Nutzen erst zu finden. Für Entscheidungen unter Risiko hat sich eine allgemein anerkannte Standardlösung herausgebildet: Man ordne jeder der möglichen Handlungskonsequenzen einen Wahrscheinlichkeitswert zu und wähle jene Handlung, bei der das mathematische Produkt von

Nutzwert und Wahrscheinlichkeitswert am grössten ist[15]. Über das rationalste Verfahren bei Entscheidungen unter Unsicherheit besteht dagegen wenig Einhelligkeit. Auf unserer Suche nach den möglichen konfliktethischen Prinzipien werden wir uns jedoch vor allem mit dieser Entscheidungsart beschäftigen müssen.

Jede Entscheidung in einer Konfliktsituation ist, in gewisser Hinsicht, eine Entscheidung unter Unsicherheit. Dies zeigt sich am besten anhand einer Analyse von Konfliktsituationen. Zu einem Konflikt kommt es unter den folgenden zwei Bedingungen: (1) Es sind mindestens zwei Akteure (Individuen, Gruppen oder Nationen) im Spiel, deren Handlungen derart miteinander verflochten sind, dass das, was der eine tun kann, von dem abhängt, was der andere tun wird, und umgekehrt. Die möglichen Handlungskonsequenzen eines jeden sind daher durch drei Dinge bestimmt: die Handlungssituation, die Handlungen der andern Akteure und die eigenen Handlungen. (2) Die beiden Akteure haben verschiedene Präferenzordnungen für die Konsequenzen ihrer möglichen Handlungskombinationen. Wären ihre Präferenzordnungen identisch, so käme es nicht zu einem Konflikt, sondern zu einer spontanen Kooperation: Jeder will genau das, was der andere auch will.

Gesetzt nun, jeder kenne die Präferenzordnung des andern und jeder wisse, dass der andere die seine kennt und dass der andere weiss, dass er dies weiss, usw. Wie kann der einzelne Akteur aufgrund dieser Information auf rationalste Weise entscheiden, was er tun soll?

Nehmen wir zum Beispiel an, A und B seien zwei Tomatenproduzenten, die den Tomatenmarkt im Lande X beherrschen. Beide stehen vor der Frage, ob sie die bisherige Produktion senken oder beibehalten sollen. Schränken beide die Produktion ein, wird der Tomatenpreis derart in die Höhe getrieben, dass sie beide einen doppelten Gewinn abschöpfen können. Produziert nur einer weniger, steigt der Preis zwar, aber nicht genug, um sein Weniger an Produktion wettzumachen. Er verliert, während der andere dank seiner uneingeschränkten Produktion zu einem dreifachen Gewinn kommt. Jeder möchte seinen Gewinn erhöhen, aber keiner will zum grossen Verlierer werden. Richtig entscheiden könnte jeder im Grunde nur, wenn er wüsste, wie der andere entscheiden wird, aber keiner kann die Entscheidung des andern kennen, bevor er selbst schon entschieden hat. Die Entscheidung wird daher für jeden zu einer Entscheidung unter Unsicherheit.

Diese Einsicht in den Zusammenhang zwischen Entscheidungen in Konfliktsituationen und Entscheidungen unter Unsicherheit hilft unserer

[15] Das Prinzip der Maximierung von Nutzenerwartung ist erstmals 1738 von D. Bernoulli formuliert worden. Vgl. dazu H. Schneeweiss: *Entscheidungskriterien bei Risiko* (Berlin/Heidelberg/New York 1967).

Untersuchung ein wesentliches Stück voran. Falls dem Prinzip der Nutzenmaximierung auch für Entscheidungen in Konfliktsituationen eine sinnvolle Interpretation gegeben werden könnte, hätten wir in ihm ein erstes konfliktethisches Prinzip gefunden, ein Nullprinzip gleichsam; denn es verlangt von uns, das zu tun, was wir immer schon tun. Mit ihm zu beginnen, hat einen guten systematischen Grund. Falls es nicht genügt, lassen sich, durch zunehmende Verschärfung der moralischen Restriktionen, schrittweise die weiteren möglichen Prinzipien gewinnen. Die Hoffnung, dass Nutzenmaximierung zum Grundgesetz einer rational geordneten Gesellschaft gemacht werden könnte, ist überdies auch historisch gesehen nicht völlig abwegig. Mandeville kleidete sie in die ironische These von den privaten Lastern als den öffentlichen Tugenden[16], und A. Smith brachte sie in die ökonomische Theorie ein[17]. Als eines der Grundelemente der sogenannten freien Marktwirtschaft beherrscht sie noch heute weitgehend das politische und ökonomische Denken[18]. Der erste erfolgversprechende Versuch, die in dem Prinzip der Nutzenmaximierung enthaltenen Kriterien für Entscheidungen in Konfliktsituationen genau zu formulieren, wurde jedoch erst in jüngster Zeit unternommen: mit der mathematischen Spieltheorie[19]. Für uns stellt sich daher die Frage: Genügt das spieltheoretisch interpretierte Nutzenmaximierungsprinzip den Bedingungen eines universalen Konfliktlösungsverfahrens[20]?

Nach der Voraussetzung der Spieltheorie kann jeder Konflikt mit Hilfe der Präferenzordnungen der beiden Akteure für ihre möglichen Handlungskombinationen ausgedrückt werden. Zur übersichtlichen Darstellung wird dabei meist die Form einer Matrix gewählt, deren Zeilen die

[16] B. Mandeville: *The Fable of the Bees* (1724), hg. P. Harth (London 1970) 67, 167, 371.

[17] A. Smith: *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776). Cannan-Ausgabe (New York 1937) 14.

[18] Die moderne Nationalökonomie hat sich allerdings von dem ursprünglich rein egoistisch verstandenen Eigennutz-Axiom weitgehend freigemacht. Vgl. etwa: G. Gäfgen: *Theorie der wirtschaftlichen Entscheidung* (Tübingen ²1968) 1. Teil.

[19] Vgl. die in Anm. 3 genannten Werke. – Zur Anwendung der Spieltheorie auf sozialwissenschaftliche Probleme vgl. etwa: M. Shubik (Hg.): *Game Theory and Related Approaches to Social Behavior* (New York u.a. 1964); dtsch.: *Spieltheorie und Sozialwissenschaften* (Hamburg 1965); R. Selten: Anwendungen der Spieltheorie auf die politische Wissenschaft. In: H. Maier u.a. (Hg.): *Politik und Wissenschaft*, a.O. (Anm. 13) 287-320.

[20] Ich will mich dabei auf die sogenannten Zwei-Personen-Spiele beschränken. Mehr-Personen-Spiele lassen sich grundsätzlich, wenn man von dem schwierigen Problem der Koalitionsbildung absieht, wie Zwei-Personen-Spiele behandeln.

möglichen Handlungen des ersten Akteurs und deren Kolonnen die möglichen Handlungen des zweiten Akteurs repräsentieren. In die Felder der Matrix werden die kardinal oder ordinal ausgedrückten Präferenzen der beiden Akteure eingetragen: die des ersten in der linken untern, die des zweiten in der rechten obern Ecke eines Feldes. Der Konflikt im Beispiel der beiden Tomatenproduzenten liesse sich demnach wie folgt darstellen (Tafel 1):

Tafel 1

A \ B	produziert weniger b_1	produziert gleich viel b_2
produziert weniger a_1	2 / 2	−1 / 3
produziert gleich viel a_2	3 / −1	1 / 1

Für A ist der idealste Fall zweifellos (a_2, b_1): Er behält seine bisherige Produktion bei und B produziert weniger; für B dagegen (a_1, b_2): Er produziert wie bisher und A senkt seine Produktion. Für den Einzelnen weniger ideal, aber für beide gleich erstrebenswert ist (a_1, b_1): Die Produktion wird von beiden eingeschränkt. Bedeutend ungünstiger für beide ist es, wenn keiner die Produktion einschränkt (a_2, b_2). Für den Einzelnen am schlimmsten aber sieht es aus, wenn er – wie A im Fall (a_1, b_2) und B im Fall (a_2, b_1) – als einziger die Produktion senkt. Wie soll der einzelne Produzent entscheiden? Um der überraschenden Antwort der Spieltheorie ein wenig von ihrer Paradoxie zu nehmen, wollen wir uns das spieltheoretische Entscheidungsprinzip vorerst an einem einfachern und besser durchschaubaren Konfliktfall vergegenwärtigen.

Die extremsten und deshalb auch einfachsten Konflikte sind von der Art des folgenden Beispiels (Tafel 2):

Tafel 2

A \ B	b_1	b_2
a_1	-3 3	-2 2
a_2	-1 1	-4 4

Die beiden Akteure haben diametral entgegengesetzte Präferenzordnungen: Der Gewinn des einen ist der Verlust des andern. Es gibt, im Unterschied zum vorigen Konflikt, kein einziges gemeinsames Interesse. Konflikte dieser Art werden in der Spieltheorie ‹Nullsummenspiele› genannt (weil die Summe der Nutzenbeträge in jedem Feld der Matrix gleich Null ist). Ich will sie – zur Abgrenzung von jenen Konflikten, bei denen es nicht nur entgegengesetzte, sondern auch gemeinsame Interessen gibt – als *reine* Konflikte bezeichnen.

Wie würden zwei rationale Akteure im Fall eines reinen Konflikts entscheiden? Falls jeder, dem Prinzip der Nutzenmaximierung entsprechend, das für ihn günstigste Resultat herbeizuführen versucht, so wird keiner sein Ziel erreichen und keiner zu einer Entscheidung kommen. A müsste z. B. wie folgt überlegen: Das für mich beste Resultat käme dann zustande, wenn ich a_2 und B b_2 täte, aber B wird dieses Resultat um jeden Preis zu vermeiden suchen und deshalb b_1 wählen. In diesem Fall ist es für mich aber besser, wenn ich mich für a_1 entscheide. Doch halt, B wird sich dies natürlich auch überlegen und mit der Wahl von b_2 reagieren. Das beste für mich wäre also doch, a_2 zu tun ... Und so beginnt die gleiche Überlegung wieder von vorne. Der einzige Ausweg aus diesem Zirkel ist das sogenannte Maximin-Prinzip: man entscheide nicht im Hinblick auf den höchsten möglichen, sondern im Hinblick auf den höchsten sichern Nutzen, d. h. man wähle jene Handlung, mit der man – gleichgültig, was der andere auch immer tut – am besten fährt.

A müsste sich also fragen: Welche Handlung sichert mir den höchsten minimalen Gewinn? Und dies ist offensichtlich a_1, denn mit a_1 gewinnt er im Minimum zwei Nutzeneinheiten, während a_2 nur eine Einheit garantiert. B anderseits müsste sich fragen: Durch welche Handlung werde ich am wenigsten verlieren? Dies führt ihn zur Wahl von b_1; denn mit b_1 verliert er im schlimmsten Fall drei, mit b_2 hingegen vier Einheiten. Entscheiden sich

die Akteure auf diese Weise für die Strategie (a_1, b_1), so wird sich jeder, wenn er die Entscheidung des andern zu wissen bekommt, sagen können, dass er, selbst wenn er sie schon im voraus gekannt hätte, nicht besser hätte entscheiden können. Ihr Konflikt hat eine Lösung gefunden.

Die Spieltheorie nennt ein Strategienpaar, von dem gilt, dass keiner der beiden Spieler angesichts der Entscheidung des andern anders gewählt zu haben wünscht, einen *Gleichgewichts-* oder *Sattelpunkt.* Die Frage, ob jeder Konflikt rational gelöst werden könne, ist daher gleichbedeutend mit der Frage, ob es für jeden Konflikt einen Gleichgewichtspunkt gebe.

Eines der erstaunlichen Ergebnisse der Spieltheorie ist der Nachweis, dass es für die meisten reinen Konflikte (Nullsummenspiele) einen Gleichgewichtspunkt gibt und dass man, falls auch sogenannte gemischte Strategien als Lösungen zugelassen werden (diese bestehen darin, dass nach einer bestimmten Häufigkeit zwischen zwei Strategien alterniert wird), alle reinen Konflikte als gelöst betrachten kann. Grund zum Jubeln besteht allerdings wenig, denn reine Konflikte kommen in der Wirklichkeit, das heisst, wenn man von Schach und ähnlichen Spielen absieht, kaum je vor und sind selbst in Kriegssituationen nicht die Regel. Gegner ohne gemeinsame Interessen sind ebenso selten wie Freunde ohne divergierende. Die meisten Konflikte bewegen sich deshalb irgendwo zwischen den Extremen des reinen Konflikts und der spontanen Kooperation. Sie gehören zur Klasse der Nicht-Nullsummenspiele oder der gemischten Konflikte, bei denen die Akteure sowohl gemeinsame wie entgegengesetzte Präferenzen haben. Gibt es auch für alle gemischten Konflikte eine Lösung?

Bei reinen Konflikten oder spontaner Kooperation kann jeder Akteur seine Entscheidung unabhängig vom andern treffen: Eine gegenseitige Verständigung ist unnötig oder unerwünscht. Bei gemischten Konflikten dagegen könnten die Akteure durch Kooperation und gegenseitige Verständigung stets mehr herausholen als im Alleingang. Die entscheidende Frage ist daher, ob es zur Kooperation kommen und bei ihr bleiben wird, technisch ausgedrückt, ob es jeweils einen sogenannten kooperativen Gleichgewichtspunkt gibt: eine gemeinsame Strategie, bei der es im Interesse jedes einzelnen liegt, dass nicht nur er selbst, sondern auch jeder andere nicht von ihr abweicht. Betrachten wir das folgende Beispiel (Tafel 3):

Tafel 3

A \ B	b_1	b_2
a_1	1 3	−1 −1
a_2	−3 −2	2 1

Wählt jeder unabhängig vom andern, so wird sich – nach dem Maximinprinzip – A für a_1 entscheiden müssen (Verlust von einer statt von zwei Einheiten) und B für b_2 (Verlust von einer statt von drei Einheiten) mit dem Resultat, dass jeder eine Nutzeneinheit verliert. Kooperation drängt sich auf, da beide die Konsequenzen von (a_2, b_2) oder (a_1, b_1) diesem Resultat vorziehen würden. Sobald sie sich aber zur Kooperation entschliessen, so stehen sie vor einem neuen Konflikt; denn A möchte die Strategie (a_1, b_1) befolgt sehen und B (a_2, b_2).

Wie soll dieser Konflikt entschieden werden? An dieser Frage zerbricht leider auch die schöne Uniformität der Spieltheorie. Es liegen zwar zahlreiche Lösungsvorschläge für derartige *Kooperationskonflikte*, wie ich sie nennen möchte, vor[21], aber diese Vorschläge beruhen alle auf ziemlich willkürlichen Annahmen. Schwerwiegender und nach meiner Ansicht entlarvend aber ist die Tatsache, dass stets von dem Postulat Gebrauch gemacht werden muss, die Lösung habe die den beiden Akteuren zur Verfügung stehenden Drohmöglichkeiten zu reflektieren.

Nach J. F. Nashs Vorschlag z. B.[22] müsste im obigen Konflikt B gegenüber A bevorteilt werden; denn ein allfälliger Verhandlungsabbruch – so argumentiert Nash – würde A härter treffen als B, weil B nicht mehr als drei, A dagegen vier Einheiten zu verlieren hätte. B könne daher A mit der Drohung des Verhandlungsabbruchs unter Druck setzen.

Das Drohpostulat enthüllt das Dilemma, in dem sich die beiden Akteure befinden, wenn jeder rational im Sinne des Nutzenmaximierungsprinzips zu handeln versucht: Das Prinzip fordert von ihnen, dass sie kooperieren, aber verhindert zugleich, dass sie kooperieren können. Wenn

[21] Vgl. etwa H. Raiffa: Arbitration schemes for generalized two-person games. In: H. W. Kuhn/A. W. Tucker (Hg.): *Contributions to the Theory of Games II* (Princeton 1953) 361-387; R. D. Luce/H. Raiffa: *Games and Decisions*, a. O. (Anm. 13) 114-154; L. S. Shapley: A Value for n-Person Games. In: Kuhn/Tucker, a. O. 307-317; J. F. Nash: Two-Person Cooperative Games. In: *Econometrica* 21 (1953) 128-140; R. B. Braithwaite: *Theory of Games as a Tool for the Moral Philosopher* (Cambridge 1955).

[22] Two-Person Cooperative Games a. O. (Anm. 21).

eine Kooperation doch noch zustande kommt, dann nicht darum, weil sie ein kooperatives Gleichgewicht gefunden hätten, sondern weil der eine sich den Drohungen des andern gebeugt und klein beigegeben hat. Dies bedeutet nicht nur eine Rückkehr zu den von unserer Voraussetzung her verbotenen Konflikthandlungen, sondern ist auch eine klare Verletzung der Neutralitätsbedingung. Die Konfliktlösung wird durch den Akteur in der stärkeren Verhandlungsposition diktiert. Vertauschte man dessen Präferenzen mit denen des schwächeren Akteurs, so würde der Konflikt entsprechend anders entschieden[23].

Die Schwierigkeit bei Kooperationskonflikten besteht darin, eine für beide Akteure annehmbare Lösung zu finden. Haben sie sich aber einmal – auf die eine oder andere Weise – auf eine kooperative Strategie geeinigt, so ist ihre Lösung ein echter kooperativer Gleichgewichtspunkt: Wer von ihr abweicht, schadet nicht nur dem andern, sondern auch sich selbst. In ein noch übleres Dilemma geraten die Akteure jedoch im umgekehrten Fall: Eine kooperative Lösung lässt sich zwar leicht finden, aber für jeden gibt es eine mögliche Strategie, mit der er noch besser fahren würde. Eine gute Illustration dafür ist unser noch immer ungelöstes Ausgangsbeispiel: der Konflikt der beiden Tomatenproduzenten (vgl. Tafel 1).

Handeln beide im Alleingang, so wird sich A für a_2 und B für b_2 entschliessen müssen, um zumindest den bisherigen Gewinn von einer Einheit halten zu können. Es gibt jedoch eine Strategie, die beide vorziehen würden: die Wahl von (a_1, b_1). Kooperation, so scheint es, wäre im Interesse beider, und die Einigung auf eine gemeinsame Strategie problemlos. Aber dieser Schein trügt. Die von ihnen bevorzugte Strategie ist kein kooperativer Gleichgewichtspunkt: Jeder kann, falls der andere sich getreu an die Abmachung hält, durch Vertragsbruch mehr für sich herausholen als durch Vertragstreue: A kann, falls B bei b_1 bleibt und die Produktion einschränkt, durch die Wahl von a_2, d.h. durch uneingeschränkte Produktion, einen dreifachen statt bloss einen doppelten Gewinn erzielen, und das Umgekehrte gilt für B. Beide werden daher, sei es aus Gewinnsucht oder aus gegenseitigem Misstrauen auf die einzig sichere, aber ihnen im gleichen Masse verhasste Strategie (a_2, b_2) zurückfallen: (a_2, b_2) ist, von der Spieltheorie her gesehen, die einzige rationale Lösung ihres Konflikts.

Was hier als rational gilt, ist ein klarer Verstoss gegen unsere fünfte Rationalitätsbedingung: die Bedingung der Effizienz. Obwohl es für beide Akteure eine eindeutig bessere Lösung gäbe, müssen sie sich mit der schlechtern zufrieden geben.

[23] Auf die ethische Fragwürdigkeit der spieltheoretischen Verhandlungsmodelle wurde schon gelegentlich hingewiesen: vgl. etwa J. R. Lucas: Moralists and Gamesmen. In: *Philosophy* 34 (1959) 1-11; A. K. Sen: *Collective Choice and Social Welfare* (San Francisco 1970) 118ff.

Das Kooperationsdilemma[24] enthüllt auf eine geradezu klassische Weise die Grenze des Nutzenmaximierungsprinzips: Wenn die Akteure nur das für sie (individuell) Beste im Auge haben, sind sie nicht in der Lage, das für sie (kollektiv) Beste zu tun[25]. Um ihrem Dilemma zu entkommen, müssten sie auf ein mittelbares Verfahren zurückgreifen; denn da keiner der Akteure einen hinreichenden Grund hat, den Kooperationsvertrag zu halten, müssen sie sich erst einen Grund geben, sei es, dass sie sich gegenseitig für den Fall des Vertragsbruchs irgendwelche Sanktionen androhen oder dass sie einen Agenten, man ist versucht zu sagen, einen Souverän[26], einsetzen, der für die Einhaltung des Vertrags sorgt.

Die Frage, mit der wir an die Spieltheorie herangegangen sind – kann das Prinzip der Nutzenmaximierung zum universalen Konfliktlösungsverfahren werden? – hat eine klare Antwort gefunden: Dieses Prinzip verletzt mindestens zwei der minimalen Rationalitätsbedingungen eines universalen Konfliktlösungsverfahrens und muss daher als unannehmbar verworfen werden. Der eigentliche Gewinn unseres kurzen Ausflugs in die Spieltheorie liegt jedoch weniger in der nicht eben überraschenden Feststellung, dass das Prinzip der Nutzenmaximierung als konfliktethisches Prinzip nicht taugt, sondern vielmehr in der Einsicht, warum es nicht taugt. Die Schwierigkeiten mit dem Maximierungsprinzip stammen offensichtlich nicht daher, dass der Nutzen maximiert werden soll, sondern dass nur der *eigene* Nutzen maximiert werden soll. Jeder, so scheint es, müsste entweder weniger für sich oder mehr für andere wollen. Die Frage ist nur, wie wenig oder wieviel.

[24] Dieses von A. W. Tucker erstmals entdeckte Dilemma ist, aufgrund seiner Standardinterpretation, unter dem Namen «Gefangenen-Dilemma» bekannt geworden.

[25] Diese Defizienz der Spieltheorie wird von den meisten Kritikern zum Ausgangspunkt ihrer Kritik gemacht: vgl. etwa A. Rapoport/A. M. Chammah: *Prisoner's Dilemma* – A Study in Conflict and Cooperation (Ann Arbor 1965); A. Rapoport: *Two-Person Game Theory* – The Essential Ideas (Michigan/Ann Arbor 1966); H. Arnaszus: *Spieltheorie und Nutzenbegriff aus marxistischer Sicht* (Frankfurt a. M. 1974).

[26] Man kann sich an diesem Modell auf einzigartige Weise klarmachen, was Rousseau mit seiner Unterscheidung zwischen volonté générale und volonté de tous und der mysteriösen Formel, man müsse die Bürger zwingen, frei zu sein, gemeint haben könnte. Vgl. W. G. Runciman/A. K. Sen: Games, Justice and the General Will. In: *Mind* 74 (1965) 554-562. – Dass ich hier die Einsetzung eines Agenten als unberechtigt erkläre, steht nur scheinbar im Widerspruch mit der anfangs eingeräumten Möglichkeit (S. 133f.), unvollkommene moralische Gesinnung durch um so vollkommenere politisch-rechtliche Institutionen wettzumachen. Der geradezu paradoxe Unterschied liegt darin, dass dort – ähnlich wie in Rousseaus Modell – Kooperation erzwungen werden muss, um das Streben nach Eigennutz zu bekämpfen, hier aber, um Eigennutz überhaupt zu ermöglichen.

Gesetzt, die beiden Akteure wären beseelt von einem franziskanischen Altruismus, jeder hätte das Wohl des andern und nur das des andern im Sinn, wären sie dann gefeit gegen nicht-rationale Konfliktlösungen? Ein Blick auf die Matrix unseres Kooperationskonflikts (vgl. Tafel 3) zeigt, dass die beiden noch schlimmer dran sind als zuvor. B wählt b_1 (um A die Chance von drei Einheiten zu geben), A entscheidet sich für a_2 (um B zwei Einheiten ermöglichen zu können), und das Resultat ist das schlimmstmögliche: $(-2, -3)$.

Zu einer ähnlichen Verkehrung kommt es, wenn die beiden Akteure zu wenig für sich beanspruchen. Nehmen wir z. B. an, sie seien beide Stoiker, nur um ihre Ataraxie besorgt und deshalb bereit, lieber ihre Ziele aufzugeben, als mit andern in Konflikt zu geraten. Jeder will für sich nur die Minimallösung, A wählt daher a_2 und B b_1. Das Resultat ist wiederum das katastrophale $(-2, -3)$. – Das gesuchte Prinzip muss offenbar, wie die recta ratio der Antike, irgendwo in der Mitte liegen: jedem soviel Nutzen wie möglich und keinem weniger als nötig. Jeder billige dem andern zu, was er auch sich zubilligen würde; keiner nehme mehr als ein Auge für ein Auge, aber keiner halte auch noch die andere Wange hin. Mit einem Wort: Die nächstliegende Handlungsmaxime ist ein Prinzip, das die gleiche oder die unparteiische Berücksichtigung der Interessen aller zum Ziele hat. Ich will es das *Prinzip des unparteiischen Wohlwollens* nennen.

Dieses Prinzip ist in der Tat eines der fundamentalsten ethischen Prinzipien. Explizit zum Axiom erhoben wurde es zwar erst in der auf Hume, Hutchinson und A. Smith zurückgehenden «Idealen Beobachter Theorie» der Ethik[27], aber faktisch liegt es den meisten moralischen Prinzipien oder Theorien zugrunde: der neutestamentlichen «goldenen Regel» ebenso sehr wie dem modernen Utilitarismus oder der vor allem mit den Namen von Locke, Rousseau und Kant verbundenen Vertragstheorie, die mit der gegenwärtig weit diskutierten «Theorie der Gerechtigkeit» von John

[27] Die «Ideale Beobachter Theorie» vertritt die These, dass unser moralisches Urteil über das, was wir tun sollen, nur dann richtig sei, wenn es dem entspreche, was eine Person sagen würde, die alle Fakten kennt, absolut klar ist in ihrem Denken und wohlwollend und unparteiisch gegenüber den Personen, die von der in Frage stehenden Handlung betroffen sind. – Zur modernen Version dieser Theorie vgl. etwa W. Kneale: Objectivity in Morals. In: W. Sellars/J. Hospers (Hg.): *Readings in Ethical Theory* (New York 1952) 681-697; R. Firth: Ethical Absolutism and The Ideal Observer. In: *Philosophy and Phenomenological Research* XII (1951/52) 317-345; dazu die Diskussion zwischen R. B. Brandt und R. Firth, a.O. XV (1954/55) 407ff., 414ff., 422ff.; ferner J. Harrison, a.O. XVII (1956) 256ff.

Rawls[28] erneut zu philosophischem Ansehen gekommen ist[29]. Die Tatsache, dass sich das Prinzip des unparteiischen Wohlwollens einer ungebrochenen Tradition erfreut, macht die für uns entscheidende Frage um so dringlicher: Genügt dieses Prinzip, ich will es abkürzend als das U-Prinzip bezeichnen, den Rationalitätsbedingungen eines universalen Konfliktlösungsverfahrens? Ermöglicht es – um dies an unsern Testfällen zu prüfen –, die Kooperationskonflikte und Kooperationsdilemmas, an denen das Prinzip der Nutzenmaximierung gescheitert ist, auf rationale Weise zu lösen?

III

Betrachten wir zunächst das Kooperationsdilemma. Die Lösung dieses Dilemmas durch das U-Prinzip ist so offensichtlich, dass es auf eben diesen Fall zugeschnitten zu sein scheint: Die einzige Strategie, die den Interessen beider Konfliktparteien im gleichen Masse gerecht wird, ist die Strategie, die ihrem höchsten gemeinsamen Interesse entspricht. In unserem Beispiel (Tafel 1) ist dies (a_1, b_1). Falls sich beide Akteure das U-Prinzip zu eigen gemacht haben, werden daher beide (a_1, b_1) wählen und an dieser Wahl auch festhalten.

[28] *A Theory of Justice* (London/Oxford/New York 1972); dtsch.: *Eine Theorie der Gerechtigkeit* (Frankfurt a. M. 1975).

[29] Die Vertragstheorie in der modernen Spielart behauptet, moralisch richtig sei, was uns jene Prinzipien vorschreiben, die von rationalen und wesentlich selbst interessierten Personen unter den folgenden Bedingungen gewählt würden: Sie haben sich, vor Eintritt in ihre Gesellschaft, auf die Handlungsprinzipien zu verpflichten, denen sie in ihrem künftigen gesellschaftlichen Leben folgen wollen. Sie kennen dabei nur die allgemeinsten Gegebenheiten ihrer künftigen Umwelt und ihrer Gesellschaft, und vor allem wissen sie eines nicht: welche besondere Rolle sie in dieser Gesellschaft spielen werden. – Es ist leicht zu sehen, dass in einer solchen Vertragsposition eben jene Prinzipien gewählt werden, die auch ein «idealer Beobachter» wählen würde: (1) Die Bedingung, dass die Vertragsparteien nicht wissen dürfen, welche Rolle sie im aktuellen Leben zu spielen haben, zwingt sie zu eben jener Unparteilichkeit, die von einem idealen Beobachter verlangt wird; denn unparteiisch sein heisst unter anderem, Individuen nicht als Individuen zur Kenntnis nehmen, und es spielt keine Rolle, ob man dies darum tut, weil man so urteilen *soll* oder weil man – aufgrund seiner Unkenntnis – nicht anders urteilen *kann*. (2) Der Forderung des Wohlwollens ist in der Vertragssituation durch die Bedingung Rechnung getragen, dass die Vertragspartner selbstinteressiert sein sollen; denn wenn keiner weiss, wer er selbst sein wird, liegt es im Interesse eines jeden, das Beste für jeden zu wollen.

Doch was heisst, den Interessen beider Parteien gerecht werden in Kooperationskonflikten, in denen die Interessen des einen gegen die Interessen des andern stehen? Kann man, um dies an unserem Beispiel (Tafel 3) zu demonstrieren, bei einer Wahl zwischen (3, 1) und (1, 2) sich eindeutig für die eine oder die andere Strategie entscheiden und diese Entscheidung mit dem Hinweis auf das U-Prinzip rechtfertigen?

Diese Frage deckt eine Zweideutigkeit des U-Prinzips auf, die es klar zu sehen gilt. Das U-Prinzip verlangt zwei Dinge: Es sagt, *was* getan werden soll, und zugleich, *wie* es getan werden soll. Das Was ist geboten durch die Forderung des Wohlwollens: Es soll jene Strategie gewählt werden, die im maximalen Interesse beider Parteien liegt, und «beider» ist hier kollektiv zu verstehen: beider zusammengenommen. Das Wie, d.h. die Art und Weise, in der der kollektive Gesamtnutzen verteilt werden soll, ist durch das Postulat der Unparteilichkeit geregelt: Die Interessen einer jeden Partei (distributiv gesprochen) sind gleich zu berücksichtigen. Doch in diesem «gleich» liegt die Zweideutigkeit. In der einen Lesart könnte das Postulat etwa wie folgt formuliert werden:

1. Jeder hat das Recht auf gleiche Befriedigung seiner Interessen (d.h. man soll dafür sorgen, dass die Interessen eines jeden gleich befriedigt werden).

In der andern Lesart dagegen würde es lauten:

2. Jeder hat das gleiche Recht auf Befriedigung seiner Interessen.

Diese Formulierungen klingen zwar ähnlich, aber die Differenz könnte nicht grösser sein. Dies zeigt sich – wie so oft – am besten durch logische Formalisierung: x und z seien Variable über dem Bereich der Personen; y eine Variable über dem Bereich der Rechte zur Interessenbefriedigung; Fxy sei zu verstehen als: x kommt y zu. Die Postulate lauten dann:

(1) $(\exists y)\,(x)\,(Fxy \cdot (z)\,(x \neq z \rightarrow Fzy))$

(2) $(x)\,(\exists y)\,(Fxy \rightarrow (z)\,(x \neq z \rightarrow Fzy))$.

Unparteilichkeit in der ersten Version verlangt kategorisch, dass jedem ein Recht, und zwar das gleiche, auf Interessenbefriedigung zukommen soll. Die zweite Version dagegen spricht hypothetisch: *falls* jemandem das Recht auf Berücksichtigung seiner Interessen zukommt, dann kommt dies Recht auch jedem andern zu. Doch ob es überhaupt jemandem zukommt, bleibt offen. Diese Version erlaubte es durchaus zu sagen, jeder habe auf

die Befriedigung seiner Interessen nur soweit Anrecht, als dies seinem Verdienst, seinem Rang oder seiner Hautfarbe entspreche. Das Postulat wäre erst dann verletzt, wenn Personen von gleichem Verdienst und Rang oder von gleicher Hautfarbe verschieden behandelt würden.

Unparteilichkeit in der zweiten Version ist also nichts anderes als das Prinzip der Universalisierbarkeit oder der formalen Gerechtigkeit, wie es C. Perelman genannt hat[30]. Das Prinzip, Gleiche gleich zu behandeln, sagt aber bekanntlich nichts darüber aus, worin die Gleichheit der Gleichen oder ihre gleiche Behandlung besteht. Unparteilichkeit in der ersten Version dagegen läuft auf ein strenges Gleichheitsprinzip hinaus: Wo es um die Befriedigung von Interessen geht, sind alle gleich und demnach gleich zu behandeln[31]. – Was geschieht nun, wenn diese zwei verschiedenen Interpretationen mit dem zweiten Teil des U-Prinzips, dem Postulat der (kollektiven) Maximierung des Nutzens aller, verbunden werden?

Wird die Unparteilichkeit im Sinne des Universalisierbarkeitsprinzips verstanden, so führt sie, in Verbindung mit der Forderung des Wohlwollens, zu einer wohlbekannten ethischen Theorie: dem Utilitarismus. Im grössten Interesse aller (kollektiv verstanden) ist notwendigerweise jene Strategie, deren Gesamtnutzen am grössten ist, und dies ist – falls die Zahl der in den Konflikt verwickelten Personen gleich bleibt – zugleich auch jene Strategie, deren Durchschnittsnutzen am grössten ist. Das gleiche Recht eines jeden auf Berücksichtigung seiner Interessen ist, im Rahmen dieser Theorie, das Recht auf jenen Nutzenanteil, der mit dem höchstmöglichen Nutzen für alle verträglich ist. In unserem Kooperationskonflikt z. B. um die Verteilung der Nutzenbeträge (3, 1) und (1, 2) müsste die Strategie, die zu (3, 1) führt, gewählt werden (da hier der Gesamtnutzen am grössten ist), und jede Partei hätte das Recht auf den ihr durch diese Strategie zufallenden Anteil, nicht auf weniger und nicht auf mehr. A ist zwar bei dieser Verteilung bevorzugt, aber wenn er zufälligerweise in der Position von B wäre, so müsste auch er sich mit einer Einheit begnügen. Anders gesagt, es ist gleichgültig, welche Partei den Vor- oder den Nachteil hat, solange man für keine Partei eine Ausnahme macht. Unparteilichkeit in diesem Sinn ist spätestens seit Benthams berühmter Devise «Jeder zählt

[30] *Über die Gerechtigkeit* (München 1967) 22ff.

[31] Diese Zweideutigkeit des Gleichheitsprinzips wurde klar herausgestellt von R. Wollheim: Equality. In: *Proceedings of the Aristotelian Society LVI* (1955/56) 281ff.

als einer und keiner mehr als einer» ein wesentlicher Bestandteil des Utilitarismus[32].

Das dem Utilitarismus eigene harmonische Verhältnis zwischen dem Anspruch des Wohlwollens und dem der Unparteilichkeit zerbricht, sobald Unparteilichkeit im Sinne des Gleichheitsprinzips gedeutet wird. Wenn jeder das Recht auf gleiche Befriedigung seiner Interessen hat, dann darf die Maximierung des Nutzens aller nur soweit betrieben werden, als das Recht eines jeden auf einen gleichen Nutzenanteil gewahrt bleibt.

Nur, was heisst das Recht eines jeden wahren? Diese Redeweise mag ihren Sinn haben, wenn es, wie im Kooperationsdilemma, ein gemeinsames Interesse an ein und derselben kooperativen Strategie gibt; aber sie scheint wenig sinnvoll in Kooperationskonflikten. Wenn die Interessen in bezug auf die anzustrebende kooperative Strategie auseinandergehen, kann ein Anhänger des Gleichheitsprinzips nur zwei Lösungen anbieten, und beide sind gleich unannehmbar. Entweder behauptet er, das Prinzip fordere bloss, Lösungen im Interesse aller anzustreben, aber es sage nichts aus für jene Fälle, wo das Recht der einen oder der andern Partei geopfert werden müsse. Oder er erklärt das Prinzip als für alle Konflikte zuständig und verbietet kurzerhand, eine Strategie zu wählen, bei der nicht beide gleichen Nutzen gewinnen können. Die erste Möglichkeit kommt der Bankrotterklärung gleich, eine moralische Ordnung auf der Grundlage des unparteiischen Wohlwollens errichten zu können, und die zweite Möglichkeit stürzt uns bei Kooperationskonflikten in ein neues Dilemma: Obwohl eine kooperative Lösung im Interesse aller liegt und daher geboten wäre, darf keine gewählt werden, weil dies gegen das Recht einer der Parteien ginge.

Muss das Gleichheitsprinzip demnach als absurd verworfen werden? Nicht notwendigerweise, falls wir erlauben, von einem Recht aller auf gleiche Interessenbefriedigung auch dann zu sprechen, wenn nicht allen zu ihrem Recht verholfen werden kann. Dass dieser Gebrauch des Begriffs «Recht» nicht völlig abwegig ist, beweist die juristische Unterscheidung zwischen den Rechten im engeren Sinn, die eine ihnen entsprechende Pflicht implizieren, und den Freiheiten und Privilegien, die dies nicht tun. Ein Landeigentümer z. B. hat das Recht, andern den Zutritt zu seinem Grundstück zu untersagen, und das Privileg, es nach Belieben betreten zu

[32] Vgl. J. St. Mill: *Utilitarianism* (London 1836) c. 5; dtsch.: *Gesammelte Werke*, übers. u. hg. Th. Gomperz, Bd. 1 (1869, ND 1968) 173-200, bes. 198.

dürfen: Seinem Recht entspricht die Pflicht der andern, dieses Recht zu respektieren, aber gegenüber seinem Privileg gibt es keine entsprechende Pflicht[33]. Diese juristische Analogie ist freilich in einer wesentlichen Hinsicht schief: Das einem jeden zugebilligte Recht auf gleiche Berücksichtigung seiner Interessen ist mehr als ein Privileg, von dem man Gebrauch machen kann, falls man es wünscht und falls es die Umstände erlauben. Es ist ein positives Recht, und als solches muss es für jeden gesichert und geschützt werden. Aber man kann es nur schützen, indem man es beschränkt – zum Schutz des gleichen Rechts der andern. Anders ausgedrückt: Jedem ist das höchstmögliche Recht zu sichern, das mit dem gleichen Recht aller verträglich ist. Aber was heisst dies? Oder, konkreter gefragt, wie weit darf die Ungleichheit zwischen dem Nutzenanteil des einen und dem des andern gehen, damit dieses Postulat noch als erfüllt betrachtet werden kann? Als Antwort auf diese Frage sind die verschiedensten Prinzipien vorgeschlagen worden: so etwa das bei den Ökonomen so populäre Pareto-Prinzip oder das in jüngster Zeit von Rawls propagierte Maximinprinzip.

Das Pareto-Prinzip in seiner stärkeren Version entspricht unserer Effizienzbedingung, in seiner schwächeren Version besagt es, eine kooperative Strategie sei einer andern nur dann vorzuziehen, wenn mindestens einer sie vorziehe und alle andern indifferent seien. Nach dem Maximinprinzip darf eine kooperative Strategie dann und nur dann gewählt werden, wenn sie von der Person in der schlechtesten Position vorgezogen wird. Bei einer Alternative zwischen der Nutzenverteilung von (3, 1) und (2, 1) müssten wir uns also gemäss dem Pareto-Prinzip für (3, 1) entscheiden, denn für B sind beide Alternativen gleich gut, und A zieht die erste vor; das Maximinprinzip aber verlangte von uns, indifferent zu bleiben, weil die Person in der schlechtesten Position keiner den Vorzug geben kann.

Das Pareto-Prinzip und das Maximinprinzip sind jedoch beileibe nicht die einzig möglichen Antworten auf das Problem, dem Gleichheitsprinzip auf sinnvolle Weise nachzukommen. Ebenso nahe läge eine Art von Versicherungsprinzip: Man sichere der Person in der schlechtesten Position ein bestimmtes Minimum und folge dann der Präferenz der Person in der besseren Position. Eine noch adäquatere Lösung wäre ein Prinzip der

[33] Vgl. etwa W. N. Hohfeld: *Fundamental Legal Conceptions* (New Haven 1923) 36ff.

minimalen Differenz: Man wähle jene Strategie, bei der die Differenz der Nutzenanteile der beiden Akteure am geringsten ist. Angesichts dieser verwirrenden Vielfalt von möglichen Prinzipien ist der Verdacht nicht von der Hand zu weisen, dass sich kein bestimmtes Prinzip als das einzig richtige auszeichnen lässt und dass darum das Gleichheitsprinzip wegen Verletzung der Einzigkeitsbedingung von vornherein als konfliktethisches Prinzip ausscheidet. Zu allem Überfluss steht ihm auch die Vollständigkeitsbedingung weiterhin im Weg; denn für alle diese möglichen Prinzipien lassen sich Konfliktfälle konstruieren, in denen sie keine Entscheidung erlauben.

Die Hoffnung, dem utilitaristischen Prinzip durch Elimination des Gleichheitsprinzips zu einem raschen Sieg verhelfen zu können, erweist sich freilich als trügerisch, denn für den Utilitaristen steht die Sache eher noch schlimmer. Der Utilitarismus lässt uns für die ganze Klasse der reinen Konflikte entweder ohne Antwort oder dann mit einer überaus vieldeutigen Antwort, denn was heisst hier, die Strategie mit dem höchsten Durchschnittsnutzen wählen, wenn der Gesamtnutzen für jede mögliche Strategie gleich Null ist? Dasselbe Problem stellt sich zudem auch bei bestimmten Nicht-Nullsummenspielen, nämlich überall dort, wo die zwei umstrittenen kooperativen Strategien denselben Durchschnitts- oder Gesamtnutzen aufweisen. Nicht nur das Gleichheitsprinzip, sondern auch der Utilitarismus muss daher, um als ernstzunehmender Anwärter für die Rolle eines konfliktethischen Prinzips überhaupt in Betracht zu kommen, vorerst von seinen Zwei- und Mehrdeutigkeiten befreit werden. Doch wie soll dies geschehen? Der einzig mögliche Ausweg liegt in der Kombination einer hinreichend grossen Zahl verschiedener Prinzipien. Und damit die Bedingung der Einzigkeit gewahrt bleibt, muss für die verschiedenen Prinzipien eine Prioritätenfolge – Rawls spricht von einer serialen oder lexikalen Ordnung[34] – festgelegt werden: auf das Prinzip zweiter Ordnung darf erst dann zurückgegriffen werden, wenn das Prinzip erster Ordnung keine Entscheidung herbeiführt, auf das Prinzip dritter Ordnung erst, wenn das der zweiten Ordnung nicht ausreicht, usw.

Die nötige Zahl von Prinzipien zu finden, dürfte einem Anhänger des Gleichheitsprinzips nicht schwer fallen, ein Utilitarist dagegen wird in eine gewisse Verlegenheit geraten. Er kann entweder den Gesamtnutzen

[34] *A Theory of Justice*, a.O. (Anm. 28) 42ff.

oder den Durchschnittsnutzen maximieren, aber damit ist sein Repertoir an möglichen Prinzipien erschöpft. Wenn er sein Prinzip über die Runden bringen will, bleibt ihm nichts anderes übrig, als bei der Gegenseite Anleihen zu machen und das utilitaristische Prinzip mit irgendwelchen distributiven Prinzipien zu ergänzen. Die umgekehrte Möglichkeit muss dann aber auch dem Egalitaristen zugebilligt werden: Er kann seinerseits das utilitaristische Prinzip als untergeordnetes Prinzip zulassen. Die Differenz zwischen den beiden Positionen ist mit dieser Möglichkeit gegenseitiger Anleihe jedoch nicht im entferntesten behoben; denn der entscheidende Streitpunkt bleibt nach wie vor, welche Art von Prinzip zum höchsten Prinzip erklärt werden soll: ein utilitaristisches oder ein egalitäres. Ein Verfechter des Gleichheitsprinzips wird um jeden Preis, und das heisst, selbst um den Preis eines noch so grossen Verlusts an Gesamt- oder Durchschnittsnutzen eine Verteilung zu verhindern suchen, bei der das Recht des einen dem Recht des andern geopfert wird. Für den Utilitaristen dagegen zählt zunächst nur die Aggregation und nicht die Distribution; er blickt nur auf die Verluste und Gewinne, aber nicht auf die Personen, die verlieren oder gewinnen. Bei der Wahl zwischen (101,0) und (50,50) sieht er nur die zusätzliche Einheit, die mit Hilfe der ersten Strategie zu gewinnen ist, der Nicht-Utilitarist aber die Person, die um dieser einen Nutzeneinheit willen ihren Anteil zu opfern hat.

Gibt es eine Möglichkeit, diesen Konflikt zwischen Utilitaristen und Nicht-Utilitaristen zu lösen und eine Entscheidung zugunsten der einen oder der andern Interpretation des U-Prinzips herbeizuführen? Ein Konflikt über das zur Lösung von Konflikten anzuwendende Entscheidungsverfahren unterscheidet sich im Grunde nicht von einem gewöhnlichen Interessenkonflikt. Die Ziele, die hier aufeinanderprallen, mögen zwar nicht egoistischer und privater Natur sein, sondern ihren Ursprung in unpersönlichen Idealen und Wertvorstellungen haben. Man wird auf der einen Seite von der Würde des Menschen, dem Respekt vor Personen und unverbrüchlichen Rechten sprechen und dies auf der andern Seite unter Berufung auf das Ideal der Wohlfahrt aller als pure Romantik und Irrationalismus abtun. Aber nach unserer Annahme sind Interessenkonflikte nicht dadurch definiert, dass die Akteure egoistische, sondern dass sie lediglich unvereinbare Ziele verfolgen. Ein Prinzipienkonflikt muss daher grundsätzlich auf dieselbe Weise entschieden werden können wie ein privater Interessenkonflikt. Dies stürzt uns jedoch dann und gerade dann

in ein Dilemma, wenn der Streitpunkt die Auslegung des höchsten konfliktethischen Prinzips selbst betrifft; denn dieser Konflikt liesse sich nur lösen, wenn er – als bereits gelöst vorausgesetzt würde. Um zu einer Entscheidung zu kommen, bedürfte es eines höhern, noch umfassenderen Prinzips. Aber welches könnte noch umfassender sein als das Prinzip des unparteiischen Wohlwollens, das, seiner Intention nach, alle in gleichem Masse umfasst?

Ohne das Prinzip des unparteiischen Wohlwollens scheint eine Konfliktethik nicht möglich zu sein, mit ihm aber ebensowenig, denn anstelle des einen legitimen Prinzips hat man gleich zwei, die beide, wenn sie den Bedingungen eines konfliktethischen Prinzips genügen wollen, Einzigkeit beanspruchen müssen. Fazit: Es gibt zwar rationale Konfliktlösungen, nämlich für alle jene Konflikte, bei denen man, wie im Fall des Kooperationsdilemmas, von einem höchsten gemeinsamen Interesse sprechen kann und bei denen daher sowohl utilitaristische wie distributive Prinzipien zum selben Resultat führen, aber es bleibt der weite Bereich der Kooperationskonflikte, bei denen die zwei Auslegungen des U-Prinzips neue Konflikte heraufbeschwören, Konflikte, auf die Max Webers berühmte, aber nicht weiter begründete These zutrifft, die Frage, wie sie zu schlichten seien, sei «ganz und gar Sache der Wahl oder des Kompromisses», weil es keinerlei rationales Verfahren irgendwelcher Art gebe, «welches hier eine Entscheidung geben könnte»[35]. Falls meine Argumentation stichhaltig ist, muss der Gedanke einer moralischen Ordnung endgültig als ein logisch unhaltbarer Traum beiseite gelegt werden: Es gibt kein moralisches Prinzip, das die minimalsten Bedingungen eines universalen Konfliktlösungsverfahrens auf konsistente Weise erfüllt.

[35] *Methodologische Schriften* (Frankfurt a. M. 1968) 248.