

Zeitschrift:	Freiburger Zeitschrift für Philosophie und Theologie = Revue philosophique et théologique de Fribourg = Rivista filosofica e teologica di Friburgo = Review of philosophy and theology of Fribourg
Band:	41 (1994)
Heft:	1-2
Artikel:	Une approche déflationniste de la liberté de la volonté : un autre visage de Kant
Autor:	Schroeter, François
DOI:	https://doi.org/10.5169/seals-761156

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: 24.01.2026

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

FRANÇOIS SCHROETER

Une approche déflationniste de la liberté de la volonté: un autre visage de Kant

Pour beaucoup, le concept de liberté contient une promesse de profondeur métaphysique. Kant s'impose certainement à cet égard comme l'un des exemples les plus typiques, et la *Critique de la raison pratique* rappelle avec emphase le rôle privilégié que joue le concept de liberté dans l'accès à la sphère métaphysique. «... de toutes les idées de la raison pure spéculative, le concept de la liberté est proprement le seul qui donne à la connaissance, même si ce n'est qu'à la connaissance pratique, une si grande extension dans le champ du suprasensible ...» (KpV, V 103/II 734)¹. Sans s'engager dans des investigations cosmologiques, l'être humain peut trouver en lui-même une voie conduisant à l'être véritable. On comprend dès lors qu'une notion aussi féconde soit appelée à former «la clé de voûte de tout l'édifice d'un système de la raison pure, y compris de la raison spéculative» (KpV, V 3–4/II 610).

¹ Pour les références au texte kantien, le tome et la page de l'*Akademieausgabe* sont d'abord indiqués; suivent le tome et la page de la traduction française dans les trois volumes des *Oeuvres philosophiques*, édition publiée sous la direction de Ferdinand ALQUIÉ, Paris: Editions de la Pléiade 1980–1985–1986. Pour la *Critique de la raison pure*, la référence à l'*Akademieausgabe* est remplacée par un renvoi à la première (A) et à la deuxième (B) édition. «GMS» signifie *Grundlegung zur Metaphysik Sitten* (*Fondements de la métaphysique des moeurs*), «KpV» signifie *Kritik der praktischen Vernunft* (*Critique de la raison pratique*).

Un tel enthousiasme métaphysique risque de susciter aujourd’hui méfiance et suspicion. Et l’on trouve de fait, en particulier dans la discussion anglo-américaine, une approche tendant à contester à la liberté toute profondeur métaphysique. La liberté se réduirait à un simple phénomène de surface. C’est cette approche que j’aimerais interroger ici afin d’en mesurer la plausibilité. Kant, le champion de la profondeur, servira de point de départ. La stratégie est certes peu orthodoxe, mais elle permettra, je l’espère, de mieux enracer le débat dans l’histoire des idées et de mettre en cause certains préjugés tenaces. Ma lecture du texte kantien sera sélective: elle prendra pour point d’appui l’approche de l’autonomie de la volonté que proposent les *Fondements* et la *Critique de la raison pratique*. En marge de sa théorie officielle de l’autonomie, Kant suggère une conception du phénomène d’autorégulation qui anticipe clairement certaines des théories actuelles de la liberté de la volonté.

1. *La théorie officielle de l’autonomie de la volonté*

L’autonomie de la volonté et l’idée d’autorégulation [*Selbstgesetzgebung, eigene Gesetzgebung*] qui en constitue le noyau forment le «sens positif» (KpV, V 33/II 647) de la liberté². De façon formelle, l’autonomie peut être caractérisée par ses relations à un réseau de notions voisines, telles que l’autolégislation, l’autorégulation et l’autocontrôle. Ce serait dans ce réseau qu’il faudrait, selon Kant, chercher le sens véritable de la liberté.

Tant dans les *Fondements* que dans la deuxième *Critique*, Kant s’attache à tracer une ligne de partage précise entre l’autonomie et l’hétéronomie. Pour ce faire, il propose une interprétation bien particulière du concept d’autolégislation. La volonté ne se donne à elle-même sa loi que dans des circonstances très spécifiques. Ma première tâche sera en conséquence herméneutique et consistera à délimiter avec précision les critères qui permettent à Kant de distinguer l’autolégislation de l’hétérolégislation.

Deux caractérisations de l’autonomie proposées par les *Fondements* peuvent servir de point de départ.

² Selon les *Fondements*, la liberté et la législation propre [*eigene Gesetzgebung*] de la volonté sont des «concepts réciproques [*Wechselbegriffe*]» (GMS, IV 450/II 320).

«L'autonomie de la volonté est la propriété de la volonté d'être à elle-même sa loi (indépendamment de toute propriété des objets du vouloir)» (GMS, IV 440/II 308).

«On voyait l'homme lié à son devoir par des lois, mais on ne réfléchissait pas qu'il n'est soumis qu'à sa propre législation, encore que cette législation soit universelle, et qu'il n'est obligé d'agir que conformément à sa volonté propre, mais à sa volonté établissant par destination de la nature [*dem Naturzwecke nach*] une législation universelle» (GMS, IV 432/II 299).

C'est la parenthèse de la première définition qui me paraît apporter l'élément décisif. Il semble que, pour fonctionner de façon autonome, la volonté doive faire abstraction de tout ce qu'elle pourrait tirer de ses objets. Par «objet» il faut entendre toute fin que l'on pourrait désirer ou, selon les termes de la deuxième *Critique*, toute détermination matérielle de la volonté (KpV §2, V 21/II 630). Ces dernières semblent donc conduire à une hétéronomie de la volonté. De façon plus positive, la deuxième caractérisation propose de lier «la volonté propre» à la loi universelle, c'est-à-dire à la loi morale. La volonté propre, qui se donne à elle-même sa loi, n'est pas ouverte à toutes les fins possibles, elle entretient un lien privilégié avec la loi formelle de la moralité.

L'opposition qui se dessine ici, entre la loi universelle et le désir d'un objet, et qui semble structurer le couple conceptuel autonomie – hétéronomie soutient l'éthique inaugurée par les *Fondements* et la deuxième *Critique*. La conception kantienne de l'autonomie ne peut s'éclairer que dans la mesure où les contours de cette opposition ont été rendus explicites. Un détour est ici inévitable.

C'est le critère d'universalisation, la pure forme universelle de la loi, qui constitue selon Kant la loi morale. Cette caractéristique formelle permet à Kant de construire une opposition entre la loi morale et l'ensemble des désirs d'objets, l'ensemble des déterminations matérielles de la volonté. On peut à cet égard parler d'une distinction fondamentale³, opposant deux types de principes déterminants de la volonté [*Bestimmungsgründe des Willens*]. En ce sens, le couple matière et forme établit une distinction entre deux façons de se déterminer à l'action, ce que l'on pourrait également considérer comme deux types de «mécanismes» d'action. La volonté peut être déterminée (a) soit immédiatement par la pure forme universelle de la loi, indépendamment de tout

³ Cf. SCHROETER, François, *La Critique kantienne de l'eudémonisme*, Berne: Peter Lang 1992, chap. 2.

objet, son principe de détermination étant alors formel, (b) soit par le désir d'un objet, son principe de détermination étant alors matériel. Lorsque la volonté est déterminée par la loi morale, aucun désir d'objet n'intervient à titre d'intermédiaire. La volonté sort donc du cercle des désirs et des préférences, pour être déterminée de façon purement formelle. Dans tous les autres cas, c'est un désir d'objet qui fonctionne comme principe de détermination: la distinction est présumée exhaustive.

La première section des *Fondements* vise à démontrer la pertinence morale de la distinction fondamentale. Le chemin est souvent sinueux, mais il est possible, je crois, de dégager les points forts suivants. Lorsque la volonté obéit aux exigences morales, lorsque, autrement dit, elle agit par devoir, elle est déterminée de façon purement formelle, sans que n'intervienne aucun désir⁴. Si, par contre, l'on refuse d'exclure l'influence de l'inclination, la volonté agit par intérêt et vise un certain plaisir, obtenu grâce à la réalisation de son objet, grâce aux «effets attendus» (GMS, IV 400/II 259).

C'est dire que deux voies, et deux voies seulement, semblent s'offrir à la volonté.

«En effet, la volonté placée juste au milieu entre son principe *a priori*, qui est formel, et son mobile [*Triebfeder*] *a posteriori*, qui est matériel, est comme à la bifurcation de deux routes [*gleichsam auf einem Scheidewege*]; et puisqu'il faut pourtant qu'elle soit déterminée par quelque chose, elle devra être déterminée par le principe formel du vouloir en général [*überhaupt*], du moment qu'une action a lieu par devoir; car alors tout principe matériel lui est enlevé» (GMS, IV 400/II 259).

Les contours de l'alternative me paraissent relativement clairs. Si la volonté entre dans le circuit de l'inclination et du plaisir (si son principe de détermination est le désir d'un objet, s'il est autrement dit matériel), l'action produite ne peut satisfaire aux requêtes de la moralité. Il faut alors postuler un autre type de mécanisme d'action: la loi morale elle-même, par le pur respect qu'elle produit, doit déterminer immédiatement la volonté, sans emprunter le circuit de l'inclination et du désir d'un objet.

Se fondant sur sa distinction fondamentale, Kant peut alors proposer une *dédiction* du principe de la raison morale commune (le critère

⁴ «... une action accomplie par devoir doit exclure [*absondern*] complètement l'influence de l'inclination et avec elle tout objet de la volonté ...» (GMS, IV 400/II 259).

d'universalisation). On peut parler de déduction en référence à la locution «il ne reste plus que [so bleibt nichts als]» (GMS, IV 402/II 261) qui introduit le critère d'universalisation avant que l'exemple de la fausse promesse n'en expose le fonctionnement (GMS, IV 402–3/261–2). Puisqu'une volonté déterminée matériellement ne peut agir par devoir, puisque toute matière a donc été exclue, seule reste en lice une détermination formelle de la volonté. En conséquence, le principe moral ne peut être autre que la forme universelle. Cette déduction me semble assez capitale malgré le peu d'emphase rhétorique qui l'entoure. Elle seule permet en effet d'accomplir la tâche que Kant se fixe dans le titre de son ouvrage, c'est-à-dire fonder un principe permettant de déduire, suivant la voie tracée par les quatre exemples de la deuxième section, la totalité des devoirs exposés dans la Métaphysique des mœurs. L'on pourrait montrer comment le couple matière et forme structure également l'opposition entre les impératifs catégorique et hypothétique, comment il conduit là aussi à une déduction, mais le détour risquerait de trop se prolonger.

Le décor est planté. Les contours de l'opposition que Kant établit entre l'autolégislation et l'hétérolégislation apparaissent maintenant, je l'espère, de façon plus claire. Le paragraphe intitulé «L'hétéronomie de la volonté comme source de tous les principes illégitimes de la moralité» apporte les confirmations souhaitées.

«Quand la volonté cherche la loi qui doit la déterminer *autre part* que dans l'aptitude de ses maximes à instituer une législation universelle qui vienne d'elle; quand en conséquence, passant par-dessus elle-même, elle cherche cette loi dans la propriété de quelqu'un de ses objets, il en résulte toujours une *hétéronomie*. Ce n'est pas alors la volonté qui se donne à elle-même la loi, c'est l'objet qui la lui donne par son rapport à elle» (GMS, IV 441/II 309).

Hétéronomie et détermination matérielle de la volonté se recouvrent parfaitement. La volonté n'est autonome que lorsqu'elle se détermine en circuit fermé, par la pure forme de l'universalisation, sans passer par la sphère des désirs d'objets. Si, par contre, la volonté sort de ce circuit fermé, son principe de détermination est matériel, c'est la propriété qu'a l'objet de produire du plaisir dans le sens interne qui la détermine. Pour Kant, un tel type de détermination est toujours synonyme d'hétéronomie. L'éthique kantienne refuse donc à la volonté désirant un objet l'accès tant à la sphère de la moralité qu'à la sphère de la liberté.

C'est sur cette double exclusion que peut se construire l'identité ou, plutôt, la réciprocité⁵ de la moralité et de la liberté. «... une volonté libre et une volonté soumise à des lois morales sont par conséquent une seule et même chose. Si donc la liberté de la volonté est supposée, il suffit d'en analyser le concept pour en déduire la moralité avec son principe.» (GMS, IV 447/II 316). La voie ouverte par la conception kantienne de l'autonomie est donc particulièrement étroite. Seule la volonté morale, pure de toute détermination matérielle, échappe à l'hétérolégislation.

2. *Les objections classiques*

Pourquoi l'autonomie requiert-elle une exclusion aussi massive des déterminations matérielles de la volonté et donc de la totalité des désirs? Que peut-on reprocher au juste à la volonté désirante, en quoi les déterminations matérielles sont-elles opposées à l'autolégislation de la «volonté propre»?

Les ébauches d'argument que l'on peut rencontrer dans le texte kantien ne vont jamais bien loin. Ainsi, dans la deuxième section des *Fondements*, lorsqu'il concentre son propos sur le couple autonomie – hétéronomie, Kant affirme-t-il que, pour une volonté déterminée matériellement, «... l'attrait [Antrieb] que la représentation d'un objet réalisable par nos forces doit exercer sur la volonté du sujet selon ses facultés naturelles fait partie de la nature du sujet..., [dans ce cas] ce serait donc proprement la nature qui donnerait la loi» (GMS, IV, 444/II, 313). Faut-il comprendre que la sphère des désirs est régie par les forces naturelles, par la constitution psycho-physiologique du sujet, et échappe de ce fait à un contrôle véritable? On cherche en vain une explication ou une défense plus précise de cette thèse programmatique.

La deuxième *Critique*, dans le premier scolie du paragraphe 3, permet peut-être d'avancer d'un pas de plus. Une volonté déterminée matériellement recherche le plaisir que peut procurer son objet. Elle détermine donc son agir en fonction de sa réceptivité au plaisir, autrement dit, en fonction de la constitution naturelle du sens interne.

⁵ L'expression est due à Allison (cf. ALLISON, Henry, «Morality and Freedom: Kant's Reciprocity Thesis», in: *The Philosophical Review* 95 [1986] 393–425).

«Si une représentation, quand bien même elle aurait son siège et son origine dans l'entendement, ne peut déterminer l'*arbitre* [Willkür] qu'en présupposant dans le sujet un sentiment de plaisir, il dépend entièrement de la nature [Beschaffenheit] du sens interne qu'elle soit un principe déterminant de l'*arbitre*, c'est-à-dire que celui-ci en puisse être affecté agréablement» (KpV, V, 23/II, 632).

Le sens interne, la partie sensible de l'homme, prennent alors l'initiative. Ce ne serait pas la volonté pure, mais son «autre», à savoir la nature, «un jeu mécanique de penchants» (KpV, V, 39/II, 655), qui dicterait le cours des événements. La simple médiation du sens interne, la présence d'un sentiment de plaisir, suffiraient donc à enlever à la volonté le contrôle de son agir. Le domaine des désirs est-il si radicalement extérieur à la maîtrise d'une volonté se déterminant elle-même?

En fait, il n'est pas difficile de deviner les raisons qui ont pu conduire Kant à estimer que la simple médiation du sens interne était suffisante pour disqualifier la volonté déterminée matériellement. Le sens interne introduit en effet l'ordre temporel et par là, selon Kant, «la nécessité naturelle», c'est-à-dire un domaine de «causes étrangères» (GMS, IV 446/II 315).

Pour Kant, la suite des phénomènes dans le temps est gouvernée par un ordre causal immuable. Chaque événement dérive immédiatement, d'après les lois fixes de la nature, de celui qui le précède. Dans l'ordre de la temporalité, tout est donc soumis au mécanisme de la nature. La fin de l'*Analytique de la deuxième Critique* («*Examen Critique*») est très explicite à cet égard. Dans toute séquence de phénomènes, ce qui arrive dans un moment du temps dérive, par nécessité naturelle, de l'événement précédent. Or, comme «le temps passé n'est plus en mon pouvoir [*in meiner Gewalt*]» (KpV, V 94/II 723), je suis donc soumis à une nécessité extérieure qui échappe à mon pouvoir et à mon contrôle⁶. Ainsi tous les événements présents dans le sens interne (et par là assujettis à la tem-

⁶ Un tel type d'argument appartient à la famille de ce que van Inwagen appelle «*the consequence argument*». «If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us.» (VAN INWAGEN, Peter, *An Essay on Free Will*, Oxford: Oxford University Press 1983, 16.). Pour un aperçu des difficultés rencontrées par un tel type d'argument, Cf. WATSON, Gary, «Free Action and Free Will», in: *Mind* 96 (1987) 155–7.

poralité) sont-ils déterminés par la nécessité naturelle. On peut certes, à ce niveau, avoir l'impression psychologique d'être libre, mais cette liberté est dérisoire; «elle ne vaudrait au fond guère mieux que celle d'un tourne-broche, qui, une fois monté, exécute de lui-même tous ses mouvements»⁷ (KpV, V 97/II 726).

On comprend aisément que, face à un déterminisme naturel aussi envahissant, le souci de Kant ait été avant tout de tenter de sauver un espace qui puisse échapper à la nécessité naturelle et où la liberté puisse élire domicile, où la volonté puisse retrouver son pouvoir sur les choses. Seule une volonté soustraite à toute emprise des phénomènes, une volonté se déterminant par la pure forme *a priori* de la loi et fonctionnant comme cause nouménale, peut occuper, selon Kant, cet espace de liberté. Inversément, il n'est pas étonnant que Kant se contente de qualificatifs très généraux («empirique», «sensible», «phénoménal»⁸) pour soumettre au déterminisme naturel une volonté déterminée matériellement et l'exclure ainsi du domaine de la liberté et de l'autonomie. L'interprétation kantienne du déterminisme naturel semble donc permettre la condamnation de toute détermination matérielle de la volonté: une telle détermination conduit immanquablement à l'hétéronomie.

On a opposé à cette ligne d'argumentation des objections particulièrement fortes, devenues désormais classiques. Peut-on vraiment utiliser la thèse du déterminisme naturel pour exclure les seules déterminations matérielles de la volonté du domaine de l'autonomie, domaine réservé alors aux déterminations formelles et où s'exerce, semble-t-il, la causalité nouménale du monde *intelligible*? On peut sérieusement en douter. Il faut en effet remarquer, avec Ortwein⁹, que le dualisme des mondes *intelligible* et *sensible* s'enchaîne mal au dualisme formé par l'autonomie et l'hétéronomie, le premier dualisme étant de nature conjonctive alors que le second repose sur une structure disjonctive.

⁷ Kant utilise également la métaphore de la pierre lancée, de la montre, de la marionnette et de l'automate de Vaucanson.

⁸ Cf. par exemple KpV, V 29/II 641–2. Si une volonté se plonge dans le cercle des désirs, «ses principes déterminants eux-mêmes doivent être des phénomènes». Elle subit ainsi «la loi des phénomènes, donc ce mécanisme de la nature qui est précisément le contraire de la liberté».

⁹ ORTWEIN, Birger, *Kants problematische Freiheitslehre*, Bonn: Bouvier 1983, en particulier 115.

Après avoir rendu le déterminisme naturel aussi radical et envahissant – tout le domaine temporel lui est irrémédiablement soumis – Kant ne peut sauver la liberté qu'en la situant à un autre niveau, celui d'un monde intelligible strictement distinct du niveau des phénomènes. Ces deux niveaux se superposent et ne sauraient entrer en concurrence. C'est pourtant d'une concurrence qu'a besoin Kant afin de marquer une ligne de partage entre deux modes de détermination de la volonté qui s'excluent radicalement l'un l'autre.

La difficulté à laquelle je fais allusion est bien visible dans la troisième section des *Fondements*. Dès la seconde partie du paragraphe intitulé «De l'intérêt qui s'attache aux idées de la moralité» (GMS, IV 450/II 320), Kant reprend explicitement le modèle des deux points de vue superposés (le point de vue du monde intelligible et le point de vue du monde sensible) que la *Critique de la raison pure* proposait, dans la solution de la troisième antinomie¹⁰, afin de réconcilier liberté et déterminisme naturel. Un tel modèle conduit à une co-présence du déterminisme naturel ainsi que du monde intelligible dans toutes les actions humaines.

«Il a donc deux points de vue d'où il [sc. l'homme] peut se considérer lui-même et connaître les lois de l'exercice de ses facultés, par suite de toutes ses actions; d'un côté, en tant qu'il appartient au monde sensible, il est soumis à des lois de la nature (hétéronomie); de l'autre côté, en tant qu'il appartient au monde intelligible, il est soumis à des lois qui sont indépendantes de la nature, qui ne sont pas empiriques, mais fondées uniquement dans la raison.» (GMS, IV 452/II 323, je souligne «toutes ses actions»). La liberté et l'autonomie, comme causa noumenon, peuvent donc coexister avec le déterminisme implacable de la nature, si on les situe hors du temps, à un niveau totalement différent.

Certes, une telle vision du problème permet de dépasser les contradictions qui menaçaient de s'établir entre le déterminisme de la nature et la liberté. Mais, selon ce modèle, il n'y a pas d'antagonisme possible entre les lois de la nature et celles de la liberté: chacun des deux systèmes de lois décrit un niveau qui lui est propre et qui est si radicalement différent de l'autre qu'aucun conflit ne peut éclater. Comme le dit Bennett, chaque action humaine «leads a double life»¹¹, chacune des vies

¹⁰ Cf. en particulier A 538 = B 566/I, 1171–2.

¹¹ BENNETT, Jonathan, Kant's Dialectic, Cambridge: Cambridge University Press 1974 190.

étant soumise à sa nécessité propre et soustraite à toute interférence en provenance de l'autre niveau. De sorte que l'argument de la temporalité ne peut permettre la partition de l'agir humain requise par le couple autonomie – hétéronomie. Si le déterminisme constitue une menace, cette menace touche toute action, du fait de sa composante phénoménale, et ne saurait se limiter aux seules déterminations matérielles de la volonté.

C'est ici que s'enchaînent des objections bien connues, qui visent cette fois la cohérence interne du modèle kantien et non plus sa seule utilisation dans le but de disqualifier les déterminations matérielles de la volonté. On peut douter en effet que le modèle kantien réussisse à rétablir, face au déterminisme naturel, le pouvoir de la volonté libre sur les événements du monde phénoménal¹². Si le déterminisme naturel constitue véritablement une menace pour la liberté et s'il tient en outre sous son emprise toutes les actions, il est bien difficile de comprendre comment la superposition d'un second niveau, incapable de changer le cours naturel, peut être en mesure de sauver la liberté. En s'expatriant, la liberté abdique.

J'en conclus que, telle qu'elle a été présentée ici, l'exclusion des déterminations matérielles de la volonté hors du domaine de l'autonomie n'est pas légitime. La conception nouménale de l'autolégislation ouvre une porte trop étroite. On pourrait alors tenter d'effacer l'emprise du couple conceptuel matière et forme et de soumettre le texte kantien à une série d'amendements successifs afin de colmater les brèches: l'autonomie et l'hétéronomie seraient, par exemple, considérées comme des «*models of choice*»¹³ et la référence aux «mondes» intelligible et sensible

¹² La liberté nouménale est en effet située hors du temps et hors de la sphère des phénomènes. Elle double éternellement le monde des phénomènes qui suit, quant à lui, imperturbablement son cours. On peut se demander, avec Bennett, comment une telle liberté peut «leave its mark on the world of events, making a difference to what occurs in that world» (BENNETT, Kant's Dialectic 200). Cf. également les remarques de Röttges: «Wenn Kant also die Freiheit dadurch gegen die Verabsolutierung des Naturmechanismus zu retten meint, daß er die intelligible Ursache als zumindest außer der Zeitreihe (A 531, 552) stehend bestimmt, mithin nicht nur als ohne zeitlich Vorhergehendes, sondern auch als ohne eine es wiederum Bewirkendes darstellt, so nimmt er ihr damit zugleich die Möglichkeit, mit einem zeitlich Nachfolgenden im Ursache–Wirkung Verhältnis zu stehen... Dasselbe Argument, das die Denkmöglichkeit von Freiheit begründen soll, hebt also ihre reale Möglichkeit auf» (RÖTTGES, Heinz, «Kants Auflösung der Freiheitsantinomie», in: *Kant-Studien* 65 [1974] 46–7).

¹³ Solution proposée par REATH, Andrew, «Hedonism, Heteronomy and Kant's Principle of Autonomy», in: *Pacific Philosophical Quarterly* 70 (1989) 42–72.

deviendrait une métaphore indiquant différents points de vue déchargés de toute portée ontologique¹⁴. Il est peut-être possible, en suivant cette stratégie, de reconstruire certaines parties de l'édifice kantien. Le sens de l'ensemble du projet, ainsi que le sens de l'exclusion des déterminations matérielles de la volonté, s'en trouveraient bien sûr largement modifiés. Je ne vais pas m'engager dans cette voie, mais proposer un nouveau point de départ, tiré d'une conception plus modeste de l'autolégislation, conception suggérée par le texte kantien lui-même.

3. *L'autolégislation des maximes*

Si l'idée d'un contrôle de la volonté sur son agir n'est pas dénuée de sens, si la distinction entre une auto- et une hétérodétermination ne relève pas exclusivement de l'illusion, il faut trouver une interprétation plus souple et nuancée, c'est-à-dire non-nouménale, du «se donner à soi-même sa loi» qui constitue le cœur de l'autonomie. J'aimerais montrer que le texte kantien fournit, en contre-point à la «théorie officielle» de l'autonomie, une telle interprétation, ou du moins l'esquisse de la conceptualité qui lui est nécessaire. Dans cette perspective, une identification massive entre hétéronomie et détermination matérielle de la volonté ne saurait être maintenue: désirer un objet ne signifie plus alors la perte automatique du pouvoir de contrôle et la chute dans l'hétéronomie.

Je pense ici au concept kantien de maxime, dont Bittner a été l'un des premiers à relever l'originalité et la fécondité¹⁵. Kant ne se réfère pas aux maximes lorsqu'il définit la notion d'autonomie. Pourtant, Kant ne considère pas les maximes comme des règles imposées de l'extérieur par la nature, mais comme des règles «que l'on se donne»¹⁶, que la volonté se fixe à elle-même dans un acte d'autodétermination¹⁷. Dans

¹⁴ Cf., pour un exemple récent, ROSEN, Michael, «Kant's Anti-Determinism», in: *Proceedings of the Aristotelian Society* 89 (1989) 125–41.

¹⁵ BITTNER, Rüdiger, «Maximen», in: G. FUNKE (Hg.), *Akten des 4. Kant-Kongresses*, Berlin: de Gruyter 1974, 485–98. Cf. également, en français, HÖFFE, Otfried, *Introduction à la philosophie pratique de Kant*, Albeuve: Castella 1985, 82–94.

¹⁶ Cf. entre autres GMS, IV 422/II 286; KpV, V 19, 27/II 628, 639.

¹⁷ Cf. BITTNER, Maximen, 486–487: «Kants bevorzugte Wendung «jemand macht sich etwas zur Maxime», verglichen etwa mit «jemand hat die Maxime», unterstreicht ein wesentliches Moment: Maxime ist nur, was einer sich zu seiner macht. Man hat sie nicht einfach, man setzt sie sich als eine solche ..., als ob von einer Maxime nur nach der Formel des sich einer Maxime Setzens gesprochen werden könnte».

un passage de la deuxième section des *Fondements*, Kant utilise même le concept de maxime pour caractériser le monde des êtres raisonnables (*mundus intelligibilis*), considéré comme un règne des fins, et l'opposer au règne de la nature, considéré quant à lui comme une machine.

«Un règne des fins n'est donc possible que par analogie avec un règne de la nature; mais le premier ne se constitue que d'après des maximes, c'est-à-dire des règles que l'on s'impose à soi-même [*sich selbst auferlegten Regeln*], tandis que le dernier ne se constitue que selon des lois de causes efficientes soumises à une contrainte extérieure» (GMS, IV 438/II 306).

Les indices semblent suffisants pour reconnaître dans l'acte de se fixer une maxime une conception alternative de l'autonomie¹⁸, une interprétation non-nouménale de l'idée de «se donner à soi-même sa loi».

En quoi consiste l'autorégulation exprimée dans les maximes? Il n'est pas d'emblée aisés de le déterminer avec précision, car Kant ne fournit au sujet des maximes que des indications lacunaires et qui ne peuvent que difficilement s'intégrer dans un tout cohérent¹⁹. Les maximes sont des principes pratiques subjectifs²⁰ et s'opposent en conséquence aux lois pratiques objectives. Leur validité dépend des conditions subjectives de l'agent et en particulier de la réceptivité au plaisir de son sens interne. Si l'on en croit les *Fondements*, l'étude des maximes concernerait avant tout la psychologie empirique et ressortirait ainsi à la philosophie de la nature. On comprend dès lors que l'éthique kantienne, soucieuse de fonder la loi morale dans une raison indépendante des

¹⁸ Ce que Bittner appelle «*natürliche Autonomie*» (BITTNER, Maximen, 494). Cf. également CARNOIS, Bernard, La cohérence de la doctrine kantienne de la liberté, Paris: Seuil 1973, 139: «La maxime est une règle qui pour être subjective n'en est pas moins rationnelle; et là où il y a raison, il y a aussi liberté». Il faut noter que l'autolégislation des maximes n'est pas un simple épiphénomène de peu d'importance. Si l'on en croit *La Religion dans les limites de la simple raison* («la liberté de l'arbitre est d'une structure bien particulière, puisqu'elle ne peut être déterminée à une action par aucun mobile, à moins que l'homme ne l'ait admis dans sa maxime...» [VI 24/III 34–5]), elle semble jouer un rôle central dans la caractérisation de la liberté humaine.

¹⁹ Pour un relevé des exemples et définitions proposés par Kant, Cf. BITTNER, Maximen, 485. Allison présente plusieurs des problèmes auxquels se heurte toute interprétation du concept de maxime (ALLISON, Henry, Kant's Theory of Freedom, Cambridge: Cambridge University Press 1990, 86–94).

²⁰ GMS, IV 400, 421/II 260, 284.

contingences de la nature humaine, n'ait consacré que peu de détails à une analyse des maximes²¹.

Se fondant sur le premier paragraphe de la *Critique de la raison pratique*, où Kant considère les maximes comme des propositions générales «dont dépendent plusieurs règles» (KpV, V 19/II 628), Bittner définit les maximes comme des principes de vie fondamentaux, déterminant le mode et la direction d'une vie dans sa totalité, c'est-à-dire quelle sorte d'homme l'on veut être²². Ne pourraient donc être considérées comme maximes que les règles les plus générales. Cette voie d'interprétation a trouvé un accueil favorable²³ et a considérablement influencé l'analyse des maximes; elle s'accorde avec l'étymologie du terme²⁴ et permet surtout d'écartier certains contre-exemples bien connus du test d'universalisation proposé par l'impératif catégorique. Des règles telles que «je jouerai tous les soirs au tennis» ne passent pas, en effet, le test d'universalisation, et ce pour une raison triviale: le manque de courts de tennis. De telles règles ne semblent pourtant pas contreviendre aux exigences de la moralité. En réduisant les maximes aux principes les plus généraux, on peut contrer avec élégance l'objection: le type de règle en question ne peut être considéré comme une maxime, il n'est par conséquent pas soumis au test d'universalisation de la loi morale²⁵.

²¹ «Dans une philosophie pratique, où il s'agit de poser, non des raisons de ce qui *arrive*, mais des lois de ce qui *doit arriver*, quand même cela n'arriverait jamais, c'est-à-dire des lois objectives pratiques, nous n'avons pas par là-même ... à nous demander sur quoi repose le sentiment du plaisir et de la peine, comment de ce sentiment naissent les désirs et inclinations, comment des désirs et inclinations naissent, par la coopération de la raison, des maximes: car tout cela fait partie d'une doctrine empirique de l'âme qui devrait constituer la seconde partie d'une doctrine de la nature, si l'on considère celle-ci comme *philosophie de la nature*, en tant qu'elle est fondée sur des *lois empiriques*.» (GMS, IV 427/II 292).

²² BITTNER, *Maximen*, 488–9.

²³ Cf. HÖFFE, *Introduction*, 85–7; O'NEILL, *Onora, Constructions of Reason. Explorations of Kant's Practical Philosophy*, Cambridge: Cambridge University Press 1990, 129. O'Neill interprète les maximes comme des intentions fondamentales, distinctes des intentions plus spécifiques.

²⁴ Selon Beck, la maxime peut être comprise comme la *sententia maxima*, première prémissse, la plus élevée et la plus globale, d'un polysyllogisme pratique (BECK, Lewis White, *A Commentary of Kant's Critique of Practical Reason*, Chicago: University of Chicago Press 1960, 61).

²⁵ Il faut rappeler que le test ne concerne que les maximes (GMS, IV 421/II 285).

Pourtant, une telle stratégie ne va pas sans problèmes. L'assise textuelle de cette ligne d'interprétation est faible; dans les *Fondements*, par exemple, on ne trouve aucun indice qui permette de déterminer le degré de généralité des maximes. La définition du premier paragraphe de la *Critique de la raison pratique* est ici bien seule²⁶ et des considérations d'ordre architectonique ne sont peut-être pas étrangères à la manière dont les maximes y sont définies. Afin de respecter les grandes divisions de la première *Critique*²⁷, Kant conçoit les maximes et les lois pratiques en analogie avec les principes [*Grundsätze*] théoriques qui constituent les règles les plus générales de la connaissance. Un souci de parallélisme a peut-être infléchi la définition des maximes dans la deuxième *Critique*. Cette inflexion paraît provisoire, puisque bon nombre des maximes que Kant cite en exemple ne peuvent être considérées comme des règles fondamentales de vie²⁸.

Il semble donc plus prudent de ne pas imposer de restrictions trop strictes à l'interprétation du concept de maxime et d'accepter que le degré de généralité de ces dernières puisse être variable²⁹. Ce d'autant plus qu'une vision plus nuancée des maximes se révèle plus propice à une interprétation du processus d'autolégislation qu'elles contiennent. A vouloir en effet considérer les maximes comme des orientations radicales et fondamentales, on risque de les isoler des facteurs qui interviennent explicitement dans les processus de délibération et de décision. Dans cette perspective, les maximes désigneraient plutôt des principes implicites que des règles que l'on a présentes à l'esprit. Mais on peut alors sérieusement se demander dans quelle mesure de telles orientations de vie, fondamentales et implicites³⁰, peuvent encore être considérées sans ambiguïté comme des «règles que l'on se donne à soi-même». Comment, en effet, des orientations si difficilement identifiables pourraient-elles être soumises à la maîtrise d'un choix explici-

²⁶ La Religion se réfère certes à une maxime fondamentale, mais semble admettre en outre une hiérarchie de maximes de différents degrés de généralité.

²⁷ Cf. KpV, Introduction, V 16/II 625.

²⁸ Cf. par exemple: «La maxime de son action signifierait ceci: quand je crois être à cours d'argent, j'emprunte et je promets de rendre, bien que je sache que je n'en ferai jamais rien» (GMS, IV 422/II 286).

²⁹ Je me rallie en cela à la proposition de ALLISON, Kant's Theory of Freedom, 93.

³⁰ «On ne peut pas observer les maximes, et cela même pas toujours en soi-même» ajoute même la *Religion* (VI 20/III 31).

te³¹? Par contre, une interprétation moins restrictive du concept de maxime – qui considérerait par exemple la maxime d’arrêter de fumer comme archétype – rapproche les maximes des processus de décision et en fait d’authentiques «règles que l’on se donne à soi-même». Pour ces raisons, je ne considérerai pas les principes définissant une orientation existentielle, fondamentale et radicale, comme les modèles de l’auto-détermination caractéristique des maximes.

Allison propose, de façon je crois pertinente, de formaliser ainsi la structure de surface des maximes: lorsque tu te trouves dans une situation de type S, accomplis une action de type A³². Ce qui donne, pour l’un des exemples des *Fondements*: lorsque tu as besoin d’argent, emprunte tout en promettant trompeusement de payer en retour. On pourrait dès lors penser que le lien d’un comportement déterminé à un type de situation spécifique constitue la propriété marquante des maximes, l’agent se donnant ainsi une ligne de conduite uniforme. D’où un certain arrière-goût peu recommandable, car les maximes semblent renvoyer alors à la psychologie somme toute très spéciale de «l’homme à principes», qui suit des règles fixes auxquelles il tente de plier une réalité souvent trop complexe. Faut-il voir dans l’autorégulation des maximes une planification outrancière, une structure de contrôle rigide et en fait peu prometteuse?

Certains éléments de la théorie des maximes ébauchée par Kant vont peut-être dans ce sens, même si une revalorisation du jugement pratique permet certainement d’enlever à l’image ainsi esquissée une partie de sa rigidité. Mais l’élément décisif me semble résider ailleurs, dans ce qui est présupposé par le schéma d’Allison et qui n’est formulé explicitement que dans certains exemples de Kant. Le lien que la maxime établit entre la situation et l’action, en l’occurrence entre le besoin d’argent et la fausse promesse, repose sur une certaine structure motivationnelle. Il est le résultat du choix d’un ordre de priorité dans les préférences. Pour ce qui est de la maxime citée en exemple, ce lien repose sur la priorité accordée à l’intérêt personnel sur les exigences de la moralité. En se donnant pour règle une telle maxime, l’agent intro-

³¹ La Religion introduit ici un choix intemporel de la maxime suprême. Cette voie me paraît trop liée à une conception nouménale et intemporelle de la liberté pour donner naissance à une interprétation prometteuse de l’autonomie.

³² ALLISON, Kant’s Theory of Freedom, 90–1.

duit une structure dans ses préférences et choisit de conformer son comportement à cette structure. Cette référence à une priorité de préférences est souvent implicite, parce que triviale; elle apparaît explicitement dans certains des exemples des *Fondements*: «par amour de moi-même» (GMS, IV 422/II 285–6), je me pose comme maxime de me suicider si la vie me procure plus de maux que de satisfactions. C'est à nouveau l'intérêt personnel qui a priorité sur l'ensemble des autres considérations. La maxime d'arrêter de fumer établit également un ordre de priorité: la santé ou le désir de ne pas incommoder autrui prennent le dessus sur le besoin créé par l'habitude.

En se donnant une maxime, l'agent prend position face à ses préférences et les organise en un ordre hiérarchique. Il s'identifie à certains de ses désirs et en rejette d'autres, qualifiés de moins importants. Grâce à cette structure hiérarchique, l'agent peut exercer un autocontrôle sur son agir, en le soumettant aux désirs qu'il considère comme centraux et authentiques, autrement dit, aux désirs constitutifs de son identité personnelle. C'est l'ordre de priorité dans les préférences qui forme la structure profonde d'une maxime et définit le sens de l'autorégulation qui est ainsi produite. La structure de surface et la référence à un certain type de situation apparaissent alors comme secondaires, elles servent à rendre opérationnelle et efficace la hiérarchie de préférences qui a été fixée. En ce sens, les maximes me paraissent suggérer une interprétation prometteuse de l'autolégislation constitutive de l'autonomie. Reste à préciser les détails de cette interprétation et à se demander si elle peut suffire à éclairer le sens du concept de liberté de la volonté.

4. *Les développements du modèle hiérarchique*

Ce sont peut-être les travaux de Harry Frankfurt qui fournissent le meilleur point de départ pour développer le type d'interprétation hiérarchique de l'autonomie qui est suggérée par les maximes. Certaines intuitions simples peuvent fournir un bon accès à ce modèle. Qui, tel le prisonnier dans ses chaînes, ne peut traduire ses désirs dans des actions, n'est bien évidemment pas considéré comme libre. Il y a désaccord entre ce que veut l'agent et son action, la cause du désaccord se trouvant par exemple dans un obstacle extérieur ou dans une contrainte exercée par autrui. Ce type de paradigme définit un premier niveau de liberté, que l'on pourrait appeler liberté d'action.

On peut alors enchaîner un second niveau de liberté, plus radical que le premier, car celui qui agit conformément à ses désirs n'est pas toujours libre. Les cas paradigmatisques font ici référence à des situations où la volonté trouve des obstacles à l'intérieur d'elle-même, comme par exemple dans les phobies, la cleptomanie, la toxicomanie, etc. L'action a sa source dans les désirs de l'agent et n'est pas entravée par des circonstances qui feraient obstacle à sa réalisation. C'est ici le vouloir lui-même qui n'est pas libre, parce que confronté à des obstacles intérieurs. En ce sens, ce second niveau concerne le problème spécifique de la liberté de la volonté, la liberté du vouloir et de ses processus de choix et de décision, non la simple liberté d'accomplir ce que l'on veut. C'est ce type de liberté que s'attache à définir avec plus de précision le modèle hiérarchique et qui, selon Frankfurt, marque la spécificité de la personne humaine.

L'analyse de Frankfurt se fonde sur une distinction hiérarchique entre des désirs de premier ordre et des désirs de deuxième ordre. Les désirs de premier ordre ont pour objet les actions de l'agent et sont à leur tour l'objet de désirs de deuxième ordre. Grâce à ces derniers, l'agent peut contrôler son système de motivation: il peut prendre position face à ses désirs de premier ordre, les évaluer, leur donner ou leur refuser priorité. C'est ce type d'autorégulation qui constitue selon Frankfurt la liberté de la volonté. La volonté est libre lorsqu'elle se conforme à ses évaluations de deuxième ordre, car l'agent possède alors la volonté qu'il veut avoir. «... *[T]he statement that a person enjoys freedom of the will means... that he is free to will what he wants to will, or to have the will he wants... [I]n securing the conformity of his will to his second order volitions... a person exercises freedom of the will.*»³³ On peut appeler «*wantons*» les créatures qui, telles les animaux, ne prennent pas position face à leurs propres désirs: de telles créatures n'ont pas la possibilité de rendre leur volonté libre et, en conséquence, ne sont pas considérées comme des personnes. De même, le cleptomane qui, sans évaluation critique, obéit à ses penchants pour le vol se comporte comme un «*wanton*». Si, par contre, il forme un désir de deuxième ordre dans lequel, par exemple, il valorise négativement son penchant, il met en œuvre la structure caractéristique de la personne. Mais il est également vrai, dans un autre sens cette fois, que sa volonté n'est pas libre: faute de se conformer à son désir

³³ FRANKFURT, Harry, *The Importance of What We Care About*, Cambridge: Cambridge University Press 1988, 20.

de deuxième ordre, l'agent n'a pas la volonté qu'il désire avoir. La présence de désirs de deuxième degré ainsi que leur efficacité seraient donc les éléments fondamentaux de la liberté de la volonté.

Watson a adressé à Frankfurt une importante objection: «*Since second order volitions are themselves simply desires, to add them to the context of conflict is just to increase the number of contenders; it is not to give a special place to any of those in contention.*»³⁴ La simple présence d'une structure de contrôle hiérarchique ne semble pas suffire à assurer la liberté de la volonté. Cette structure, ainsi que les désirs de deuxième degré qui la commandent, pourraient très bien être dictés par des phénomènes tels que l'endoctrinement, le conformisme social ou l'autosuggestion. Dans de telles circonstances, le contrôle hiérarchique serait oppressif et inadapté, il ne pourrait donc rendre la volonté libre. Cette difficulté est communément nommée «problème de l'identification». Ce n'est que si les désirs de deuxième degré permettent à l'agent de s'identifier à ses vraies préférences qu'ils peuvent revendiquer une «place spéciale» et garantir que la volonté soit bien celle que l'agent veut avoir. Il devient dès lors impératif de préciser les contraintes que l'on doit imposer à ce processus d'identification afin que la liberté de la volonté soit sauvegardée. Frankfurt semble proposer un critère purement subjectif: lorsque la volonté, grâce à ses désirs de deuxième ordre, s'identifie de façon décisive³⁵ à ses désirs de premier ordre, les conditions sont alors suffisantes pour que la volonté soit libre. Mais un tel critère purement subjectif répond mal aux objections que l'on vient de formuler (phénomènes d'endoctrinement et d'autosuggestion): c'est en toute bonne foi que la personne endoctrinée croit obéir à ses vraies préférences et, pourtant, sa volonté n'est pas libre. Le processus d'identification doit être soumis à des restrictions plus sévères.

Le problème est ici complexe et il n'est pas question de prétendre trouver une solution simple et précise. On ne peut dévoiler les vraies préférences d'une personne, son identité personnelle, à l'aide d'une formule toute prête. La meilleure stratégie consiste certainement à dresser la liste des dysfonctionnements qui peuvent entraver le processus d'identification et à formuler en contrepartie un ensemble d'exigences qui, une fois satisfaites, permettent d'exclure de telles défaillan-

³⁴ WATSON, Gary, «Free Agency», in: *Journal of Philosophy* 72 (1975) 94.

³⁵ FRANKFURT, The Importance, 21, 164–7.

ces. Les travaux de Young et Double vont dans ce sens³⁶. Comme il existe de multiples défaillances possibles, les restrictions imposées au processus d'identification ne sont pas unifiées. Citons, à titre d'exemple, l'exigence d'auto-connaissance. Pour que les processus d'évaluation et d'identification fonctionnent correctement, un agent doit posséder une connaissance suffisante de ses propres états psychologiques. En cas d'ignorance ou d'auto-tromperie (*self-deception*), on s'expose à des dysfonctionnements du système de contrôle. De même, un seuil minimum de cohérence interne doit être présent dans l'ensemble des préférences et intérêts afin que l'identification puisse se dérouler de façon acceptable. Si l'unité d'un agent a été désintégrée et si différents sous-systèmes s'affrontent trop violemment, l'identification des vraies préférences et, par là, l'autocontrôle de la volonté sont menacés. On peut allonger la liste et mentionner l'exigence d'un minimum de réflexion critique et de rationalité au sens large. Si chacun des facteurs mentionnés est présent de façon suffisante, le processus d'identification se passe de façon adéquate et la structure hiérarchique exerce son contrôle en donnant priorité aux vrais désirs de l'agent.

Développé de la sorte, le modèle hiérarchique renonce donc à faire de la liberté de la volonté une propriété simple et transparente: une volonté est libre quand elle obéit à un ensemble complexe de conditions (présence d'une structure de contrôle et d'un processus adéquat d'identification des désirs prioritaires). Une fois ces conditions remplies, la volonté possède tout le contrôle qu'elle peut espérer. C'est en ce sens que le modèle hiérarchique peut être considéré comme déflationniste. Aucune propriété magique de spontanéité absolue, aucune sortie héroïque du déterminisme naturel ne sont exigées. Le moi est certes formé par des influences multiples, en provenance tant du monde social que du monde biologique, mais les conditions fixées par le modèle hiérarchique sont compatibles avec de telles influences. De même, il semble possible de défendre le modèle hiérarchique sans recourir à la thèse de l'indéterminisme. Frankfurt déclare en effet, à juste titre me semble-t-il, que le modèle hiérarchique «*appears to be neutral with regard to the*

³⁶ YOUNG, Robert, Personal Autonomy. Beyond Positive and Negative Liberty, London: Croom Helm 1986, 52–62; DOUBLE, Richard, The Non-Reality of Free Will, Oxford: Oxford University Press 1991, 37–49. Benn dresse déjà une table des divers dysfonctionnements possibles (Benn, Stanley I., «Freedom, Autonomy and the Concept of a Person», in: *Proceedings of the Aristotelian Society* 75 (1975) 113–6).

*problem of determinism. It seems conceivable that it should be causally determined that a person is free to want what he wants to want. If this is conceivable, then it might be causally determined that a person enjoys free will.»*³⁷ La liberté peut être considérée comme une qualité interne³⁸ que possèdent les choix de la volonté, peu importe alors la façon dont le monde est constitué, peu importe la manière dont les séquences causales s'enchaînent. Lorsqu'il s'agit de savoir si la volonté satisfait aux exigences posées par le modèle hiérarchique, la question du déterminisme semble ne pas avoir de pertinence. Pour le dire en termes kantiens, l'autonomie des maximes n'a que faire d'une liberté de type transcendental.

Mais, en évitant ainsi les problèmes, n'a-t-on pas perdu l'essentiel de la liberté? N'y a-t-il pas plus d'épaisseur dans le concept de liberté de la volonté, faut-il se contenter de ces quelques conditions somme toute bien superficielles? L'inflation menace à nouveau. Une défense du modèle hiérarchique exigerait de livrer bataille sur de nombreux fronts³⁹. Je ne peux ici qu'indiquer les raisons auxquelles peuvent faire appel les défenseurs du modèle hiérarchique afin de mettre sérieusement en doute les motifs qui ont traditionnellement poussé à l'inflation.

³⁷ FRANKFURT, *The Importance*, 25.

³⁸ Selon DOUBLE, *The Non-Reality*, 56, le modèle hiérarchique considère la liberté de la volonté comme un concept internaliste. Cette position est analogue à celle de Descartes et Fodor à l'égard des états psychologiques: pour Descartes un état psychologique est le type d'état qu'il est du fait de ses caractéristiques internes et fonctionnelles, indépendamment de ses connections avec le monde externe. De même, les choix restent des choix et conservent leur liberté indépendamment de leur situation dans le monde physique.

³⁹ Le problème le plus épineux concerne certainement ce que Frankfurt nomme «*principle of alternate possibilities*» (FRANKFURT, *The Importance*, 1). Nos intuitions pré-philosophiques lient le concept de liberté de la volonté à la possibilité d'agir autrement («*could have chosen otherwise*»). Le modèle hiérarchique propose une interprétation faible du principe des possibilités alternatives: l'agent libre aurait pu agir différemment *si* les conditions avaient été différentes. Ses adversaires tentent de défendre la nécessité d'une interprétation plus forte, c'est-à-dire non-hypothétique: est libre qui peut agir différemment dans les mêmes circonstances. Il faudrait alors s'engager dans une investigation plus substantielle de la structure interne du sujet humain et de sa volonté. Mais il est particulièrement difficile de trouver une interprétation non-hypothétique qui puisse être satisfaisante. Pour un survol de la problématique, cf. WATSON, *Free Action*, 151–63. Face à la persistance des désaccords, d'aucuns ont suggéré que les concepts de liberté de la volonté et d'autonomie n'ont pas une cohérence suffisante pour générer un débat constructif (cf. CHRISTMAN, John, «*Constructing the Inner Citadel: Recent Work on the Concept of Autonomy*», in: *Ethics* 99 (1988) 109–12; DOUBLE, *The Non-Reality*, 3–10). Ces concepts seraient organisés autour de paradigmes hétérogènes.

Dans notre tradition issue de la modernité, le concept de liberté de la volonté est intimement lié à ce qui constitue la dignité humaine⁴⁰. En parlant de liberté, on souhaite toucher les éléments décisifs de la spécificité humaine et de sa valeur, et l'on caresse l'espoir d'isoler une propriété simple, suffisamment remarquable pour démarquer l'homme du reste des créatures. Face à ces attentes, le modèle hiérarchique et l'attention privilégiée – exclusive, diront peut-être certains – qu'il porte à un catalogue d'obstacles internes peuvent apparaître comme une démission. Certes, par exemple, la menace que semblait impliquer le déterminisme a perdu une grande partie de sa force⁴¹. Pourtant, on souhaiterait comprendre comment nos contenus mentaux peuvent avoir une efficacité causale et être liés à des processus neurologiques. Et l'on se dit qu'une approche satisfaisante de la liberté de la volonté devrait éclairer cette question radicale. En fait, on espère que le problème de la liberté de la volonté permette de mettre en lumière la nature profonde de la capacité humaine de délibération et de choix. De sorte que, pour répondre aux attentes, il faudrait en somme tenter de rendre transparentes la volonté et la structure interne du sujet humain.

Face à une telle ambition, la stratégie du modèle hiérarchique présente des avantages incontestables: en limitant ses prétentions, elle permet de distinguer avec beaucoup plus de clarté les différents problèmes. Les concepts évaluatifs, au nombre desquels il faut compter la liberté, ont une plasticité qui encourage les amalgames. De ce fait, la stratégie déflationniste, qui pourrait apparaître d'abord comme une démission, doit plutôt être considérée comme l'invitation à une précision conceptuelle bienvenue. Et c'est peut-être l'attente d'une profondeur métaphysique qui doit être tenue en suspicion.

⁴⁰ Cf., bien sûr, Kant (GMS, IV 436/II 303). Tant Watson (Free Action, 63) que Kane (KANE, Robert, Free Will and Values, Albany: SUNY Press 1985, 8) s'appuient de fait sur l'idée de dignité pour réintroduire une approche inflationniste de la liberté.

⁴¹ Une fois écartées les métaphores simplificatrices (que Dennett appelle pompes à intuitions: Cf. le modèle de la table de billard, où ce n'est pas tant le processus causal que la façon simpliste dont il est représenté qui peut apparaître comme une menace), il est difficile de voir en quel sens la présence d'une chaîne causale ou de ce que Williams appelle «*strong naturalistic explanations*» (WILLIAMS, Bernard, «How Free Does the Will Need to Be?», Lindley Lectures, The university of Kansas 1985, 4–5) peuvent mettre en péril l'autocontrôle de la volonté. Cf. DENNETT, Daniel, Elbow Room. The Varieties of Free Will Worth Wanting, Oxford: Oxford University Press 1984, 32–4, 135–7.