

# Kants begreb om det radikale onde

*Ph.d.-stipendiat, cand.theol.*  
*Jacob P.B. Mortensen*

Abstract: The article is a reading of Book One in Kant's *Religion within the Limits of Reason Alone*. The main focus is the concept of radical evil, which Kant posits as a philosophical analogue to the Christian concept of Original Sin. The article unfolds the relations between the concepts that Kant uses to establish the concept of radical evil. The main point is that Kant ends up contradicting his own conceptual definition because he ascribes evil to the concept of freedom, which is fundamentally good. The article thus follows a peripheral and marginalized trajectory within Kantian scholarship by proposing a Kant who is inconsistent and paradoxical. Even though this 'contradictory Kant' ends up not explaining what he sets out to explain, the article appreciates his work for his effort to find a foothold in the question concerning the problem of evil.

Keywords: Kant – concept of evil – freedom – will (Wille/Willkür) – religion.

## Indledning

Fra et overordnet perspektiv involverer Kants filosofi et partnerskab mellem gud og mennesker i skabelsen og opretholdelsen af det moralske univers. Tilnærmelsen til 'dydens rige' (*Religion*, 753)<sup>1</sup> må for Kant både være mulig og begribelig, hvis ikke fornuften skal gå under i sin egen mangelfuldhed. Det er denne problematik Kant kæmper med i *Religion* og som skriftet kan ses som den endelige kamp om. Det radikale onde underminerer aktualiseringen af det moralske univers, og paradoksalt nok for Kants tænkning er det netop denne funktion, ondskab har – den underminerer og fordærver. Ondskab gør verden uforståelig og ubegribelig, fordi den udspringer af frihe-

---

1 Alle henvisninger til Kants værker er fra W. Weischedels udgave, *Immanuel Kant – Werke in Sechs Bänden* (Darmstadt: Wissenschaftliche Buchgesellschaft 1998). Værkerne forkortes således: *Kritik der Reinen Vernunft*: KRV (Kant 1998, Band II); *Kritik der Praktischen Vernunft*: KPV (Kant 1998, Band IV); *Grundlegung zur Metaphysik der Sitten: Grundlegung* (Kant 1998, Band IV); *Metaphysik der Sitten*: MS (Kant 1998, Band IV); *Die Religion innerhalb der Grenzen der blossen Vernunft*: *Religion* (Kant 1998, Band IV).

den. Frihedens ubetingede godhed drages i tvivl, og det onde bliver en fundamental trussel mod hele systemet. I forsøget på at inkorporere forestillingen om det onde i det filosofiske system ender Kant således med at modarbejde sig selv.

For Kant kulminerer religion i det øjeblik, hvor den moralske person indser sin manglende evne til at forstå relationen mellem det, der bør opnås af ham selv og det, der kun kan forventes af Gud – den moralske genrejsning. Med læren om det radikale onde befinder Kant sig således på grænsen af sin egen pelagianisme. I retrospektivt lys bliver det muligt for os at se, at Kant er *allermest* interessant, når han er *allermindst* pelagiansk – hvilket er en tragisk-ironisk modpol til resten af forfatterskabet. I fremstillingens forløb afsløres det gradvist, at Kant sår tvivl om den menneskelige autonomis muligheder i mødet med dets selvskabte problem: fremstillingen ender i det pinlige paradoks, at menneskets fald er selvforskyldt, da fornuften ikke er herre i eget hus. Inden for Kants system betyder det, at fornuften vender sig mod sig selv, hvorved friheden producerer sin egen sørgelige svaghed.

### Placering af *Religion* i forfatterskabet

Jeg vil i denne artikel analysere Kants fremstilling af begrebet om det radikale onde ud fra hans religionsfilosofiske skrift, *Religion innerhalb der Grenzen der blossen Vernunft*. *Religion* er Kants antropologiske pendant til kristendommens lære om arvesynden. Sigtet med analysen er at vise, hvordan Kant i forsøget på at inkorporere forestillingen om det onde med forestillingen om friheden ender med at modarbejde og modsige sig selv. I værket gennemspilles traditionelle kantianske tematikker som vilje, autonomi, ren praktisk fornuft, rationel moralitet, natur, frihed, nødvendighed, etc.<sup>2</sup> Det spændende ved værket er, at de traditionelle tematikker gennemspilles i transponerede tonearter. Den normale fokusering på friheden, viljens autonomi og de deraf afledte begreber er flyttet, fordi ondskaben forsøges forklaret på samme niveau som og udspringende af friheden. Traditionelle apologetiske fortolkere af Kant ser *ikke* denne accent-forskydning, men bliver inden for grænserne af *main-stream* fortolkninger.<sup>3</sup> Denne

2 Hoyt H. Hudson, "Wille, Wilkür, and the Imputability of Immoral Actions", *Kant-Studien* 82:2 (1991), 179-196 (179); Stephen R. Palmquist, "Does Kant Reduce Religion to Morality", *Kant-Studien* 83:2 (1992), 129-148 (134ff.); Larry Herrera, "Kant on the Moral *Triebfeder*", *Kant-Studien* 91 (2000), 395-410 (395ff.); Allen W. Wood, *Kant's Moral Religion* (Ithaca & London: Cornell UP 1970), 208.

3 Lewis W. Beck, *A Commentary on Kant's Critique of Practical Reason* (Chicago: The University of Chicago Press 1960), 176ff.; John R. Silber, "The Ethical Significance of Kant's *Religion*", *Religion within the Limits of Reason Alone*, red. Theodore

artikelens fokus på en selvmodsigende Kant er således marginaliseret og perifer inden for Kant-receptionen, der ikke opererer med en inkonsistent Kant. Stephen Palmquist betegner arbejdet hos den apologetiske og syntetiserende Kant-reception som 'kludetæppe-fortolkning' (Palmquist 1992, 132).

I Kants praktiske filosofi spiller menneskets autonomi en afgørende rolle. Som almindelige, fornuftige mennesker er vi selv-styrende i moralske anliggender og ligger dermed ikke under for ydre autoriteter i vore valg. Derudover besidder vi – i kraft af vores fornuft – magten til at kontrollere os selv. J.B. Schneewind udtrykker det således: "Although no one can lose the autonomy that is part of the nature of rational agents, social arrangements and the actions of others can encourage lapses into governance by our desires, or heteronomy. Kant ... found it difficult to explain just how this could happen".<sup>4</sup> Citatet sætter fingeren på *præcis* de to ting, der bliver problematisk i *Religion* i modsætning til den tidligere praktiske filosofi, og som min fremstilling fokuserer på: for det første bliver menneskets autonomi vanskelig at opretholde for Kant, fordi ondskaben er medfødt (selvom den er tilvalgt); og for det andet bliver det yderst vanskeligt at forklare, *hvorfor* det hænger således sammen, når ondskabens grund-løse logik sættes op imod rationaliteten i den tidligere praktiske filosofi.

Artiklen falder i to hoveddele, en analyserende og en fortolkende. Baggrunden for den omfattende fremstilling af begrebet om det radikale onde i den første og analyserende del er, at flere af de begreber, der uomgængeligt er knyttet til det radikale onde, præsenteres andre steder i forfatterskabet. Da Kant imidlertid selv tillagde fremstillingen i *Religion* stor værdi, og da progressionen i værket har en afgørende rolle, er det vigtigt at præsentere begreberne i deres sammenhæng i dette værk. Derudover afføder fremstillingen en forklaringsmæssig gevinst i forhold til fortolkningen ved først at redegøre for den sammenhæng, fortolkningen læses ind i. Mange fortolkere mister dette aspekt, da de eksklusivt fokuserer på enkelte begreber, der relateres

---

M Green & Hoyt H. Hudson (New York: Harper and Row 1960), lxxix-cxxxix (lxxixff.); Hudson (1991), 181ff; Clement C.J. Webb, *Kant's Philosophy of Religion* (Oxford: The Clarendon Press 1926), 62; Michael K. Green, "Kant and Moral Self-Deception", *Kant-Studien* 83:2 (1992), 149-169 (149); Daniel O'Connor "Good and Evil Disposition" *Kant-Studien* 76:3 (1985), 288-302 (288); Herrera (2000), 396f.; Wood (1970), 222. For kommentarer til denne tendens inden for forskningen, jf. Palmquist (1992), 142-3; Pablo Muchnik, "On the Alleged Vacuity of Kant's Concept of Evil", *Kant-Studien* 97 (2006), 430-451 (434 note 12). Med begrebet om *main-stream* fortolkninger refererer jeg til, hvilke artikler i *Kant-Studien*, der skriver om *Religion*.

<sup>4</sup> J.B. Schneewind, "Autonomy, obligation, and virtue: An overview of Kant's moral philosophy", *The Cambridge Companion to Kant*, red. Paul Guyer (Cambridge: Cambridge University Press, 1992), 309-341 (310).

til resten af forfatterskabet. Yderligere gør den pointe sig gældende, at min fortolkning er funderet i den forskydning, der opstår mellem de begreber, der er med til at etablere begrebet om det radikale onde. Ved at præsentere grundlaget for fortolkningen så omhyggeligt vindes dermed en klarhed i fremstillingen.

### Kants begreb om det radikale onde

I det følgende vil jeg analysere de vigtigste tematikker i første bog af *Religion* for at kunne fremstille begrebet om det radikale onde. Jeg vil forsøge at trække dem frem enkeltvis og udfolde, hvorfor de er vigtige. Jeg vil yderligere forsøge at relatere de enkelte tematikker til hinanden, hvorved sammenhængen og systematikken i Kants fremstilling understreges. Analysen sigter dog hele tiden på begrebet om det radikale onde, hvorfor spørgsmål, der relaterer til underbegreberne, ikke tages op.

#### *Handlinger vs. maksimer*

Formålet med første bog af *Religion* er at udrede, hvorfor mennesker handler ondt, og hvori det onde består. Kant indleder sin betragtning med påstanden om, at vi ikke kalder en person ond, fordi han udfører onde (böse, gesetzwidrige) handlinger, “sondern weil diese so beschaffen sind, dass sie auf böse Maximen in ihm schliessen lassen” (*Religion*, 666). Det onde er altså funderet i maksimer – ikke handlinger. Maksimer er subjektive grundsætninger for fornuften, dvs. grundsætninger der ikke kan tillægges almengyldighed som fx det kategoriske imperativ, der er et objektivt princip.<sup>5</sup> En maksime er et praktisk princip, der indikerer, at man ‘vil’ handlingen.<sup>6</sup> Problemet med at placere det onde i maksimerne er, at “die Maximen kann man

5 For diskussioner af maksimers rolle jf. Muchnik (2006), 432; Hudson (1991), 179; Matthew Caswell “Kant’s Conception of the Highest Good, The *Gesinnung*, and the Theory of Radical Evil” *Kant-Studien* 97 (2006), 184-209 (191ff.).

6 Ved hjælp af forestillingen om maksimen kan Kant indfange forskellen mellem måder, man *rent faktisk* handler på, og måder man *burde* handle på. Hvis subjektive, praktiske maksimer kan tænkes at være i overensstemmelse med det kategoriske imperativ, er der tale om ægte, moralske maksimer, og handlingen er en moralsk handling. Når mennesker handler efter maksimer, der kun er bestemt af tilbøjeligheder (en sanselig tilskyndelse), handler det ufrit. Mennesket er kun frit, når det handler ud fra ægte moralske maksimer, fordi kun moralen er fornuftig og dermed fri.

nicht beobachten” (ibid.), og derfor kan man ikke med sikkerhed vurdere, om mennesket vitterligt er ondt.<sup>7</sup>

Hvis det skal være muligt at kalde et menneske ondt, må det være muligt *a priori* at slutte fra en enkelt ond handling udført med bevidstheden om det onde i den til en underliggende ond maksime. Yderligere må det være muligt, ud fra denne maksime, at slutte sig frem til “einen in dem Subjekt allgemein liegenden Grund aller besondern moralisch-bösen Maximen, der selbst wiederum Maxime ist” (ibid.). Der findes altså en almen grund i subjektet for alle onde maksimer, og denne grund er selv en maksime. Denne underliggende maksime kvalificeres senere som menneskets sindelag (Gesinnung), og minder om noget, vi kan kalde menneskets natur.

### *Natur vs. frihed*

Normalt skal natur forstås som modsætning til frihed; naturen er nødvendig, friheden er fri. Når det drejer sig om ‘menneskets natur’, er denne sondring imidlertid ugyldig. Menneskets natur er “der subjektive Grund des Gebrauchs seiner Freiheit überhaupt” (*Religion*, 667). Denne subjektive grund er den nødvendige forløber for alle handlinger og må selv være et udtryk for friheden. Hvis den subjektive grund *ikke* selv var et udtryk for friheden, ville menneskets brug af vælgeevnen (Willkür) ikke kunne tilskrives mennesket, hvorved det gode eller onde ikke ville kunne vurderes moralsk.<sup>8</sup> Kun fordi menneskets subjektive grund eller natur *er* frihed, kan mennesket tilskrives gode og onde handlinger. Kilden til ondskab kan dermed hverken ligge i en genstand, der påvirker vælgeevnen (Willkür) gennem en tilbøjelighed (Neigung) eller i en naturlig impuls (Naturtrieb). Kilden til ondskab kan kun ligge i en regel, som vælgeevnen (Willkür) selv skaber sig til brug for sin frihed – det vil sige i en maksime. Hvis den subjektive grund *ikke* endegyldigt var en maksime, *så* kunne vi spore brugen af vores frihed til naturlige årsager, fx til en tilbøjelighed (Neigung). Men dette går direkte imod forståelsen af frihed. Når mennesket altså ‘af natur’ enten er god eller ond, betyder det, at mennesket “enthält einen (uns unerforschlichen) ersten Grund der Annehmung guter, oder ... böser (gesetzwidriger) Maximen” (*Religion*,

7 Den moralsk handlende person ved aldrig, om handlingen er moralsk rigtigt, eller om sindelaget er godt (jf. KRV 501 (A 552)), hvilket medfører et paradoksalt epistemologisk problem for en teori om moralen. David Sussman “Kantian Forgiveness” *Kant-Studien* 96 (2005), 85-107 (97ff.) berører dette i forbindelse med tildeling af Guds nåde, der også foregår paradoksalt og uden menneskelig indsigt.

8 Kant betegner Willkür som “der absoluten Spontaneität ... (der Freiheit)” (*Religion*, 670). Willkürs spontanitet skal forstås i sammenhæng med menneskets autonomi og frihed. I en hvis forstand *er* mennesket vælgeevnen (Willkür).

667-8). *Qua* menneske besidder ethvert menneske denne subjektive, uransagelige grund. Og på baggrund af denne er mennesket 'af natur' godt eller ondt. Det betyder, at naturen – i modsætning til friheden – ikke er skyldig, men at mennesket selv bestemmer sine handlinger, fordi menneskets natur er frihedens grund. Menneskets natur *er* altså frihed, og i frihed vælger mennesket ondskaben. Senere viser det sig, at denne frit valgte ondskab er naturlig og medfødt.

### *Anlæg, hang, tilbøjelighed og Triebfeder*

Ifølge Kant består menneskets natur af både gode og onde anlæg. Anlægget til ondskab kvalificeres senere som en hang, men af gode anlæg har mennesket tre. Det første er et anlæg til menneskets dyriskhed *qua* noget levende. Det er en fysisk og rent mekanisk selvkærlighed, der ikke kræver nogen fornuft. Det andet er et anlæg til menneskets menneskelighed *qua* noget levende og samtidig fornuftigt. Det er en selvkærlighed, der er fysisk, men dog sammenligner, hvorfor der er brug for fornuft. Det tredje er et anlæg til personlighed "als eines vernünftigen, und zugleich der *Zurechnung* fähigen Wesens." (*Religion*, 673, alle kursiveringer her og senere er Kants egne.) Det tredje anlæg angår altså menneskets duelighed for tilregnelighed. Om alle tre anlæg gælder, at de ikke blot i sig selv er gode, men de er også anlæg *for* eller *til* det gode. Udover at være gode er de også "*ursprünglich*, denn sie gehören zur Möglichkeit der menschlichen Natur" (*Religion*, 675). De to første anlæg kan misbruges til noget ondt, men de kan ikke fjernes. Det tredje kan ikke misbruges, da det er selve fornuften, hvori viljen (*Wille*) har sæde.

Ved hang skal indledningsvis forstås "den subjektiven Grund der Möglichkeit einer Neigung" (*ibid.*). En hang kan være erhvervet, når den er god, og være nedbragt over mennesket, når den er ond. Men Kant vil primært beskæftige sig med hangen til moralsk ondskab, fordi ondskab "nur als Bestimmung der freien Willkür möglich ist" (*Religion*, 676). Og da vælgtevnen (*Willkür*) kun kan være god eller ond i relation til sine maksimer, så må hangen til ondskab i kvalificeret forstand bestå i "dem subjektiven Grunde der Möglichkeit der Abweichung der Maximen vom moralischen Gesetze" (*ibid.*). I det omfang denne hang er udbredt blandt mennesker, kan den altså kaldes en naturlig hang til ondskab i mennesket.

Kant opererer med tre niveauer af ondskab. Den første kalder han hjertets svaghed eller skrøbelighed. Den anden er hangen til at blande moralske og amoralske motiver eller drivfjedre (*Triebfedern*), hvilket kaldes urenhed. Den tredje er hangen til at indoptage onde maksimer, hvilket er en decideret ondskabsfuldhed og det egentlige onde.

Om denne gælder, at den er “der Hang der Willkür zu Maximen, die Triebfeder aus dem moralischen Gesetz ändern (nicht moralischen) nachzusetzen” (*Religion*, 677). Dette kaldes for hjertets bagvendthed eller perversitet, fordi det omvender den etiske orden i henseende til drivfjedrene (Triebfedern) for en fri vælgeevne (Willkür). Når dette sker “wird doch die Denkgungsart dadurch in ihrer Wurzel (was die moralische Gesinnung trifft) verderbt, und der Mensch darum als böse bezeichnet” (ibid.). Når mennesket altså omvender den etiske orden ved at producere amoralske maksimer, så fordærves det i sin rod og kan betegnes som ondt.

I forbindelse med påstanden om, at menneskets tænkemåde fordærves i sin rod, henviser Kant til Romerbrevet 14,23: Alt, hvad der ikke er af tro, er synd. Citatet understreger, hvilken betydning motivet eller drivfjederen (Triebfeder) har for bestemmelsen af maksimen. Når andre drivfjedre (fx ambitioner, selvkærlighed, sympati) end moralloven selv – som den praktiske fornuft (Wille) forestiller for vælgeevnen (Willkür) – er nødvendige for at bestemme vælgeevnen til lovmæssige handlinger, så er det en tilfældighed, at de stemmer overens med loven, da de ligeså vel kunne føre til overtrædelsen af den. Dette hænger sammen med Kants bestemmelse af drivfjederen som en motiverende kraft.<sup>9</sup> Motivation er en kraft, hvorfor der er tale om kontradiktoriske og *ikke* kontrære principper. Det betyder, at hvis det gode = a, så er fraværet af det gode (det ikke-gode) *ikke* = 0, men derimod = -a (*Religion*, 669 noten). Dette hænger sammen med, at moralloven er en motiverende kraft (= a), og at fraværet af vælgeevnens overensstemmelse med denne lov kun er mulig som konsekvensen af en virkelig og modsat bestemmelse af vælgeevnen. Det betyder, at modsætningen til loven (= -a), er en ond vælgeevne (Willkür). Der er altså ingen middelgrund mellem et godt og et ondt sindelag (Gesinnung), hvorfor mennesket nødvendigvis må bestemmes som enten godt eller ondt som en helhed.

Inden Kant kommer til begrebet om det radikale onde, foretager han en vigtig distinktion i begrebet om hangen. En hang kan enten angå noget fysisk eller noget moralsk. Når hangen angår noget fysisk, kan der ikke være tale om moralsk ondskab, fordi moralsk ondskab nødvendigvis må udspringe af friheden. “[E]in physischer Hang ... zu irgend einem Gebrauche der Freiheit ... ist ein Widerspruch.” (*Religion*, 678) Altså kan en hang til det onde kun angå vælgeevnens moralske formåen. Og kun det er moralsk ondt “d.i. zurechnungs-

<sup>9</sup> Muchnik (2006), 437f. diskuterer dette ud fra *Grundlegung*. Herrera diskuterer radikaliteten i Kants udsagn i *Religion* om moralske følelser vs. sanselige følelser, og hvorfor det er afgørende, at det må være en bevægende kraft (Herrera (2000), 403). Man kunne ønske, han havde fortsat argumentationen og koblet den til sondringen god vs. ond, frem for blot at blive ved (moralisk) rigtig vs. forkert.

fähig”, der er vores egen handling eller valg.<sup>10</sup> Hvis en hang på den anden side angår noget moralsk, så gælder, at “dem Begriffe eines Hanges einen subjektiven Bestimmungsgrund der Willkür, der *vor jeder Tat vorhergeht*, mithin selbst noch nicht *Tat* ist” (*Religion*, 679). Men dette betyder, at handling eller valg (Tat) kan forstås på to måder, hvis ikke begrebet om hangen skal være selvmodsigende. For det første gælder, at handling eller valg kan forstås som den udøvelse af frihed, hvorved den øverste eller første maksime indoptages i vælgeevnen. Derudover gælder, at handling eller valg også kan forstås som den udøvelse af frihed, hvor handlingen selv (materielt forstået) udføres i overensstemmelse med maksimen. Der er altså tale om to niveauer af valg.

Når det drejer sig om det onde, gælder: “Der Hang zum Bösen ist nun *Tat* in der ersten Bedeutung (*peccatum originarium*), und zugleich der formale Grund aller Gesetzwidrigen *Tat* im zweiten Sinne genommen, welche der Materie nach demselben wiederstreitet, und Laster (*peccatum derivativum*) genannt wird” (ibid.).<sup>11</sup> Om den første handling gælder, at den forbliver aktualiseret i menneskets natur, selvom den anden handling gentagne gange undgås. Den første er nemlig “intelligible *Tat*, bloss durch Vernunft ohne alle Zeitbedingung erkennbar”, hvorimod den anden er sanselig, empirisk og tidsligt givet (ibid.). Den første er en simpel hang og er medfødt, fordi den ikke kan udslettes. Det må den være, fordi vi ikke kan tilskrive nogen yderligere årsag for korrupsionen i os af ondskab af den øverste maksime – selvom det er vores egen handling – end vi kan tilskrive nogen anden årsag for nogen af vores andre egenskaber. Baggrunden for dette er, at det tidlige ikke kan indvirke på friheden, hvorfor vi ikke kan angive en anden årsag. Og vi kan ikke forklare, hvorfor vi handler, som vi gør, da vi jo netop handler frit – *i.e.* vi handler spontant eller vilkårligt og friheden er os “unerforschlich”.

---

10 I forskningen kaldes dette tilregnelighedshypotesen (Hudson (1991), 179). Jf. yderligere Beck (1960), 176ff; Muchnik (2006), 436; O’Connor (1985), 291; Caswell (2006), 194-6; Herbert J. Paton, *The Categorical Imperative* (Philadelphia: University of Pennsylvania Press 1948), 213f; Roger J. Sullivan, *Immanuel Kant’s Moral Theory* (Cambridge: Cambridge UP 1989), 279ff.; (*Religion*, 673; 691).

11 Selvom Kant specifikt arbejder sig frem imod en *filosofisk* bestemmelse af ondskab *i modsætning* til en teologisk, er det bemærkelsesværdigt, at hangen til ondskab som oprindelig og intelligibel handling betegnes *peccatum originarium*, da denne betegnelse i den teologiske tradition er knyttet til begrebet om arvesynden (Søren Kierkegaard *Begrebet Angest*, SKS bind 4 (København: G.E.C. Gad 1997), 332ff.).



*Mennesket er af natur ondt*

Konsekvensen af det ovenfor udviklede er, at sætningen: “mennesket er af natur ondt”, betyder “er ist sich des moralischen Gesetzes bewusst, und hat doch die ... Abweichung von demselben in seine Maxime aufgenommen. Er is *von Natur böse*” (*Religion*, 680). Dette gælder prædikativt om mennesket som art. Selvom en sådan kvalitet ikke kan afledes ud fra genrebestemmelsen af mennesket – da den i så fald ville være nødvendig – kan vi ikke dømme anderledes ud fra, hvad vi kender i erfaringen. Og da hangen til ondskab i sig selv må forstås som moralsk ond gælder, at hangen “folglich in gesetzwidrigen Maximen der Willkür bestehen muss” (*ibid.*). Hvis friheden stadig skal gøres gældende for mennesket, må påstanden dog kvalificeres. For det første må maksimerne være kontingente, og for det andet må den øverste subjektive grund for alle maksimer være forbundet med og funderet i menneskeheden selv, da ondskaben ellers ikke kan være universel. Kun således kan det forstås som en naturlig hang til ondskab i mennesket. Og da mennesket er selvforskyldt i denne – da det er frit og derfor *kunne* have handlet anderledes – kan vi kalde det “ein *radikales*, angebornes ... *Böse* in der menschlichen Natur” (*ibid.*). Dermed er vi nået frem til begrebet om det radikale onde gennem vælgeevnens indoptagelse af en maksime, der bunder i en drivfjeder, der udtrykker hangen til ondskab, der er placeret i menneskets natur – selvom det er frit valgt.

Det radikale onde i den menneskelige natur, der er medfødt, kan *ikke* placeres i menneskets sanselige natur og de deraf følgende tilbøjeligheder (*Neigung*), fordi vi ikke er ansvarlige for de sanselige tilbøjeligheders tilstedeværelse i os. Årsagen til det radikale onde kan *heller ikke* placeres i en fordærvelse af den moralsk-lovgivende fornuft (der er “*Wille*”), da fornuften ikke kan ødelægge autoriteten i den lov, der er dens egen. Da moralen *er* fornuftig, kan den ikke gå imod sig selv, da det ville være en selvmodsigelse. Når det gælder om at finde grunden for det moralsk onde i mennesket, kan vi altså konkludere: 1) menneskets sanselige natur omfatter for lidt, fordi når drivfjederne, der motiverer friheden, fjernes, så reduceres mennesket til et dyrisk væsen, der kun handler ud af sanselige tilbøjeligheder; 2) en fornuft, der er fritaget fra moralloven, omfatter for meget, da modstanden mod moralloven i så fald ville kunne være en drivfjeder i sig selv, hvorved subjektet ville blive et djævelsk væsen. Hvis fokus altså ligger på menneskets sanselige natur, tildeles mennesket for lidt skyld i ondskab, og mennesket bliver blot et ondt dyr. Hvis fokus ligger på fornuften (en ond fornuft (*Wille*)), så tildeles mennesket for meget skyld, hvorved det gøres djævelsk. Da ingen af disse veje er farbar, må årsagen til det onde søges andetsteds.

### *Maksimernes orden*

Moralloven tvinger sig ind på mennesket i kraft af, at mennesket er fornuftigt og i kraft af, at mennesket har et naturligt anlæg for det gode. Moralloven er i sig selv en drivfjeder (Triebfeder), og hvis ingen anden drivfjeder modarbejdede moralloven, ville mennesket uden videre handle godt. Da mennesket imidlertid også er et sanseligt væsen, får det også drivfjedre fra sanserne, som det inkorporerer i maksimerne. Hvis mennesket kun indoptog de sanselige drivfjedre i maksimerne som tilstrækkelige for bestemmelsen af vælgeevnen (Willkür) – uden at konsultere moralloven – ville mennesket være moralsk ondt. Og hvis mennesket på samme tid indoptog både sanselige og fornuftige drivfjedre i maksimerne, ville mennesket på samme tid være både godt og ondt. Dette er imidlertid en selvmodsigelse. Sondringen kan altså ikke drages ud fra drivfjedrenes indoptagelse i maksimerne, da begge drivfjedre hele tiden må være til stede, da mennesket på samme tid og hele tiden både *er* sine sanser og sin fornuft.<sup>12</sup> Derfor må forskellen på, om mennesket er godt eller ondt, afhænge af maksimernes orden, altså “welche von beiden er zur Bedingung der andern macht” (*Religion*, 685). Hvis den sanselige drivfjeder gøres til betingelse for den fornuftige, er mennesket ondt og *vice versa*. Dermed er “der Mensch ... nur dadurch Böse, dass er die sittliche Ordnung der Triebfedern, in der Aufnehmung derselben in seine Maximen, umkehrt” (ibid.). Når mennesket underordner den moralske drivfjeder til den sanselige, produceres en ond maksime og ondskaben bliver aktivt tilvalgt.

### *Det radikale onde*

Hvis der ligger en hang til netop denne underordning i den menneskelige natur, så har mennesket en hang til det onde. Og da denne hang fordærver mennesket i sin rod (Wurzel, jf. *Religion*, 677), er det en radikal ondskab “weil es den Grund aller Maximen verdirbt” (*Religion*, 686).<sup>13</sup> Som en naturlig og medfødt hang er den u-udrydelig med menneskelige kræfter, fordi udryddelsen kun kan ske gennem gode maksimer, og dette kan ikke finde sted, når den ultimativt subjektive grund for alle maksimer er fordærvet. Alligevel må det være mennesket selv, der har gjort sig til det, det er, i moralsk forstand,

12 Muchnik (2006), 436 diskuterer dette ud fra *Grundlegung* og betegner viljen som ‘strukturelt heterogen’.

13 ‘Radikal’ trækker på forbindelsen til det latinske *radix* (rod). Radikal angår *ikke* et maksimum af kvantitativ art på en gradueringsskala, selvom Hannah Arendt og mange efter hende har tolket Kant i denne retning. Radikal betegner heller ikke ondskabens effekt men dens *sted*, nemlig i sindelaget; det er således en rumlig metafor.

da det ellers ikke ville være skyldigt i ondskaben. Om mennesket er godt eller ondt er altså en virkning af den frie vælgeevnen (*Willkür*), da man ellers ikke kan holdes ansvarlig, og derfor hverken kan være moralsk god eller ond. På dette punkt, hvor Kant har fundet frem til ondskabens rod, og hvorfor vi handler ondt, synes han altså at have ført sig selv ind i en position, der umuligt kan opretholdes uden at være selvmodsigende. Dermed er hele hans system i fare for at bryde sammen. Mennesket er af natur ondt, fordi det har en medfødt hang til ondskab – men mennesket vælger frit ondskaben, da ondskaben ellers ville være nødvendig. Ikke desto mindre vælger *alle* mennesker det onde, hvorved de afskærer sig fra at vælge det gode, fordi ondskaben fordærver mennesket i sin rod. Alligevel lyder fordringen universelt: du *skal* handle godt, fordi du *kan*. Denne paradoksale position er udgangspunktet for fortolkningen.

Jeg vil i det følgende samle de vigtigste tematikker og relatere dem til hinanden i et systematisk perspektiv. Uden at producere en decideret læsning af værket, vil jeg forsøge at give en fortolkning af de centrale tematikker, der har relevans for forståelsen af det radikale onde.

## Fortolkning

Som ledemotiv for min fortolkning af begrebet om det radikale onde står påstanden, at Kant har forvildet sig ind i en forklaringsmæssig impotens. Hvor han tror, han har tilført begreberne et forklaringsmæssigt potentiale, ligger han i realiteten under for en dialektisk illusion, der resulterer i et forklaringsmæssigt vakuum. Min læsning ligger i forlængelse af en perifer og marginaliseret tendens inden for Kant-forskningen.<sup>14</sup> Størstedelen af Kant-receptionen anser fremstillingen i *Religion* for at være i kontinuitet med resten af det kritiske forfatterskab og blot en forventelig systematisk fremstilling af bestemte følger af teorien om viljen.<sup>15</sup> Selvom en sådan tilgang lokaliserer en indiskutabel kontinuitet, så undervurderer den en faktisk forandring i *Religion*. Et afgørende træk ved arbejdet med *Religion* er

14 O'Connor (1985) og Gordon E. Michaelson, *Fallen Freedom* (Cambridge: Cambridge UP 1990), 30-72

15 Wood (1970); Allen Wood "Kant and the Intelligibility of Evil", *Kant's Anatomy of Evil*, red. Sharon Anderson-Gold & Pablo Muchnik (Cambridge: Cambridge UP 2010), 144-172; Pablo Muchnik, *Kant's Theory of Evil* (Plymouth: Lexington Books 2009); Pablo Muchnik "An Alternative Proof of the Universal Propensity to Evil", Anderson-Gold & Muchnik (2010); Muchnik 2006, 437ff.; Palmquist (1992); Stephen Palmquist, *Kant's Critical Religion* (Aldershot: Ashgate 2000); Adina Davidovich "How to Read *Religion within the Limits of Reason Alone*" *Kant-Studien* 85:1 (1994).

således, om man ser det som et nødvendigt – men mindre – område af moralteorien, som *Grundlegung* og KPV havde ladet ufuldstændigt tilbage, eller om man ser *Religion* og teorien om det radikale onde som en afgørende nyudvikling. Er begrebet om det radikale onde en korrektiv og supplementsagtig forståelse til det allerede udfoldede i moralteorien, eller er det et område *sui generis*? Min læsning ser fremstillingen i *Religion* som et brud med det tidligere kritiske forfatterskab, da Kant forsøger at indlemme begrebet om det radikale onde i sit frihedsbegreb – der grundlæggende anser friheden for at være god – og begrebet ikke kan rumme både at være godt og ondt. Konsekvensen er, at systemet underminerer sig selv og forklaringen bliver paradoksal.

### *Anlæg, hang, tilbøjelighed og Triebfeder*

Forud for menneskets anvendelse af friheden besidder vi et ‘anlæg til det gode’. Det betyder, at vi ikke rigtig ‘er’, hvad vi er, før vi udøver friheden. Det ‘noget’, vi er før udøvelsen af friheden, er anlæggene til dyriskhed, menneskelighed og personlighed. Personlighed er den mest interessante, da den er ensbetydende med ‘den moralske følelse’ (*Religion*, 674), og er et produkt af Wille. Det er en krydsning af Willes rationalitet og pligtfølelsen. Den moralske følelse bliver altså en drivfjeder (Triebfeder) for Willkür, idet moralloven påvirker mennesket i form af en moralsk følelse. Den moralske følelse er dermed synonym med respekt for moralloven, og personligheden er evnen til at respektere moralloven som en tilstrækkelig drivfjeder til bestemmelse af Willkür. Dermed er den moralske følelse den eneste ‘praktiske’ eller ‘rationelle’ følelse, da alle andre følelser er sanselige.<sup>16</sup> Men de to andre anlæg med dertilhørende tilbøjeligheder (Neigung) er også i sig selv gode (*Religion*, 710). Krop og sanselighed må ikke forstås som noget ondt eller moralsk forkert. Det moralsk onde må udelukkende karakteriseres ved den frie handling, der underordner moralloven til følelsesmæssige tilbøjeligheder: “Das Böse hat nur aus dem Moralisch-Bösen ... entspringen können” (*Religion*, 693). Godhed og ondskab afhænger altså af, hvad vi gør med vores anlæg. Og selvom anlæggene i sig selv er relativt fastlagte – da vi ikke kan gøres ansvarlige for dem – så er vores anvendelse af dem ikke fastlagt.

16 Herrera omtaler en “intellectualist” og en “affectivist” forståelse af motivationen af den moralske handling inden for forskningen. Han fastholder selv, at den moralske handling er “affectivist”, altså at moralloven afstedkommer en følelse, der forstås som god og moralsk rigtig (Herrera (2000), 395ff.). Jf. yderligere Michaelson (1990), 37ff.

Anlæggene muliggør dermed afstanden mellem det, at vi er 'skabt til det gode' og det at vi 'kan blive onde'.

I afstanden mellem 'skabt til det gode' og muligheden for at 'blive god' ligger hangen, der er den subjektive grund for muligheden af en tilbøjelighed. Pointen med, at grunden er 'subjektiv', er elementet af valgfrihed. Dette er tydeligvis fraværende i begrebet om anlægget. Hvor vi *ikke* har noget valg i forbindelse med vores anlæg, har vi et valg – og dermed en potentiel tilregnelig skyld – i forbindelse med hangen. Pointeringen af det subjektive i forholdet mellem frihed og hang bliver dog problematisk, når vi tænker på tilbøjeligheden (*Neigung*). Ifølge Kant må tilbøjeligheden forstås som noget, der opstår ud af vores sanselige natur i en form for sanselighedens logik. Men hvordan kan tilbøjeligheden forbindes med forestillingen om frihed, der knytter an til 'en subjektiv grund' i definitionen af hangen?<sup>17</sup> Hvis vi frit kunne vælge bestemte tilbøjeligheder, kunne vi frit vælge, hvilke aspekter af anlægget til dyriskhed vi ville påvirkes af. Men tilbøjelighederne vælger vi ikke – de påvirker os på baggrund af anlæg og hang. Derfor går Kants eksempel om vilde folk, der besidder en hang "zu berauschenden Dingen" (*Religion* 676 noten), paradoksalt nok imod intentionen i argumentet, der sigter på moralen.

I definitionen af hangen kan Kant ikke tænke på fysiske, sanselige tilbøjeligheder. Det må være drivfjederens forbindelse til tilbøjeligheden, der er afgørende, hvis hangen til ondskab skal give mening. Og det er disse begrebers gensidige relation, der udgør Kants påstand om, at vi *kan* holdes ansvarlige for en drivfjeder-struktur, der er funderet i noget, vi *ikke* kan holdes ansvarlige for – *i.e.* vores sanselighed.<sup>18</sup> Fokus ændrer sig dermed fra en overordnet forståelse af hangen til selve *anvendelsen* af hangen. Og det er denne *ene* specifikke subjektive grund for muligheden af en tilbøjelighed, der er væsentlig, fordi netop den ligger bag tendensen til at afvige fra moralloven. Vores valg af onde maksimer kommer til udtryk i et uundgåeligt mønster, og det er dette mønster, der afføder indsigten om det radikale onde. Hangen bliver således Kants transcendentale skitsering af de onde maksimers uundgåelige mønster samtidig med, at mønsteret ikke defineres som en nødvendighed. Derfor udbygges den første definition af hangen også med den anden definition, hvor hangen betegnes som den subjektive grund for muligheden af maksimernes afvigelse fra moralloven (*Religion*, 676). Hvad Kant forsøger at fremlægge, er altså

17 For yderligere diskussioner jf. Wood (1970), 215ff; Wood (2010), 158ff; Michaelson (1990), 38ff; O'Connor (1985), 296.

18 Sussman (2005), 95 har misforstået kompleksiteten i Kants redegørelse, hvorfor han sidestiller en ond vilje med en pseudo-vilje, der så egentlig bare er en sammenhobning af følelser.

en teoretisk forbindelse mellem vores sanselige natur og moralsk ansvarlighed.

Den forbindelse, Kant forsøger at knytte, fremstiller sanseligheden som i sig selv uskyldig, mens der stadig kan henvises til den, når det drejer sig om moralsk ondskab. Dette er væsentligt, eftersom ondskab består i underordnelsen inden for en maksime af morallovens drivfjeder til en sanselig tilbøjelighed. Hangen hjælper altså til at forklare, hvordan den sanselige tilbøjelighed får overtaget i maksimeproduktionen. Problemet er bare, at hangen *ikke* forklarer, hvordan eller hvorfor underordnelsen af moralen til sanseligheden finder sted. Det, hangen til ondskab forklarer, er den rene mulighed, der er nødvendig for, at noget moralsk ondt kan fremkomme. Dermed opstår en væsentlig mangel, fordi fremstillingen af anlæggene *ikke* forklarer denne mulighed, da den pointerer, at alle anlæg i sig selv er gode. Teorien om anlæggene forklarer altså alle de muligheder, vi har i livet – *undtagen* vores mulighed for at blive moralsk onde, selvom dette netop er Kants intention.<sup>19</sup>

Ud af begrebet om hangen vokser et andet problem, der er knyttet til de tre grader af ondskab (*Religion*, 676-7). De tre grader står nærmest alene uden nogen form for forklaring. Spørgsmålet er, om ondskab *består* af disse tre, eller om det kun er ondskab *qua* ondskabsfuldhed, der udtrykker hangen til ondskab. Problemet er, at Kant hævder, at både moralske og sanselige drivfjedre er til stede i maksimeproduktionen. Dermed sløres ideen om 'urenhed' (anden grad af ondskab, der er hangen til at blande moralsk og amoralsk motiverende årsager), så den nærmest forsvinder. Hvis urenhed består i at underordne moralen til sanseligheden, så forvandles forestillingen om urenhed til det, der kaldes ondskabsfuldhed.<sup>20</sup> Men hvis omvendt kampen mellem de to drivfjedre udgør urenheden, så bliver den normale situation for et menneske, der både er rationelt og sanseligt, til ondt i kraft af naturlige tilbøjeligheder og drivfjedre (*Religion*, 685). Hvis fremstillingen udelukkende sigter på at omhandle ondskabsfuldheden, så sløres de to andre grader af ondskab, og det bliver utydeligt, om de overhovedet er eksempler på moralsk ondskab.<sup>21</sup>

Trods usikkerhed i fremstillingen er det dog klart, at hangen til ondskab ikke i sig selv må forstås som ond. Ligeledes er tilbøjeligheden ikke i sig selv ond, da moralsk ondskab i så fald ville have et

19 Jf. Michaelson (1990), 42.

20 Muchnik (2006), 448-50 ender med at hævde dette: "Heteronomy is... the source... of evil in strict sense..." Sussman forstår det på samme måde som jeg (Sussman (2005), 96).

21 Jf. Michaelson (1990), 45. For yderligere diskussioner af de tre graders relation se Palmquist (2000), 152ff.; Muchnik (2010), 129; Wood (1970), 217ff.; O'Connor (1985), 300.

fundament i en naturlig egenskab. Hangen er blot *tendensen* til at underordne moralloven til sanselige tilbøjeligheder i maksime-produktionen. Hangen er et anlæg, og er dermed *kun* en mulighed. I hvilket omfang vi *faktisk* afviger fra moralloven, afgøres i en specifik anvendelse af friheden. Det betyder ikke ifølge Kant, at nogle mennesker kan undgå dette; *alle* bukker under for hangen til ondskab. Det problematiske i Kants fremstilling af denne påstand er bare, at han ikke argumenterer for det. På et så afgørende punkt er Kants fremgangsmåde empirisk: "Er ist von Natur Böse ... dieses gilt von ihm in seiner Gattung betrachtet; nicht als ob solche Qualität aus seinem Gattungsbegriffe ... könne gefolgert werden (denn alsdann wäre sie notwendig), sondern er kann nach dem, wie mann ihn durch Erfahrung kennt, nicht anders beurteilt werden, oder man kann es, als subjektiv notwendig, in jedem ... Menschen voraussetzen" (*Religion*, 680). Konsekvensen af dette er, at evnen til at handle uafværgeligt – men *ikke* nødvendigt – bliver til produktionen af faktisk ondskab. Problemet for Kant bliver altså, at fordi hangen til ondskab er universel, ender han med det tilsyneladende paradoks, at ondskab både er frit valgt og medfødt. Påstanden ender således med enten at fratage friheden eller det medfødte sin egenskab.

### *Gesinnung*

Det 'radikale' i det radikale onde består i, at man *altid* underordner moralen til sanseligheden i maksime-produktionen. Pointen er dog ikke kvantitativ men kvalitativ. *På grund af* det radikale onde bliver ens tænkemåde fordærvet i sin rod, hvorved "den faulen Fleck" i vor race produceres (*Religion*, 687). Denne forestilling udvikler Kant ved hjælp af begrebet om sindelaget (*Gesinnung*). Ved hjælp af begrebet om sindelaget muliggøres forklaringen af, hvordan tænkemåden går i forfald, hvilket medfører produktionen af onde maksimer uden afskaffelsen eller udelukkelsen af frihedens rolle. På et overordnet niveau kan teorien om sindelaget (sammen med hangen) således siges at udgøre Kants filosofiske modstykke til den kristne lære om arvesynden.

Fokuseringen på sindelaget berører naturligvis anvendelsen af 'radikal' i begrebet om det onde, men det peger også frem på de spørgsmål, der opstår i relationen mellem det onde og maksime-produktionen. Hvis vi accepterer, at det er formen og ikke indholdet i en maksime, der bestemmer dens moralske gyldighed, så kan det i sig selv godt være rigtigt. Men det forklarer ikke, hvorfor og hvordan vi bliver moralsk onde. Dermed problematiseres relationen mellem frihed og en given maksimes form. Og hvis vi ser bort fra maksimens indhold,

så flyttes opmærksomheden til det autonome valg, hvorved den moralsk handlende person underordner én slags drivfjedre til en anden.<sup>22</sup> Underordningsprocessen *producerer* altså ondskaben, og ondskaben bliver et karakteriserende træk ved udførelsen af det frie valg. Men hvori består denne proces, og hvordan bestemmes formen på en given maksime?

I forbindelse med påstanden om, at moralsk ondskab sidder i formen og ikke i indholdet af en maksime, gælder, at en maksimes form bestemmes af “en fælles underliggende grund, der selv er maksime” (*Religion*, 666). Denne fælles grund for *alle* maksimer er det subjektive sindelag. Sindelaget er “den første grund for antagelsen af maksimer” (*Religion*, 671-2), og da det er en enkelt persons sindelag, må det være enhedsligt (ikke fragmenteret eller kaotisk) og gælde for hele personens udøvelse af frihed (ibid.). Men selvom sindelaget selv er en form for maksime, er det ikke produceret eller funderet i noget andet end friheden selv, da en maksime netop er et praktisk princip, som den frie vilje skaber og udøver. Formålet med at pointere det subjektive element i definitionen af sindelaget er at understrege det påkrævede element af frihed. Sindelaget bliver således en form for megamaksime, eller “die oberste Maxime” (*Religion*, 679), der opstår ud af et frit valg.<sup>23</sup> Sindelagets funktion er at tilvejebringe karakteristiske tendenser og mønstre i maksime-produktionen.

Da sindelaget er enhedsligt og udgør fundamentet for al maksimeproduktion, er det herfra, det radikale onde henter sin kraft. Hvis individuelle maksimer *de facto* er stedet for moralsk ondskab, og hvis alle maksimer har deres udspring i en enhedsmæssig grund, der er blevet ond, så udgør ondskaben personens væsen i et radikalt omfang.<sup>24</sup> “Dieses Böse ist radikal, weil es den Grund aller Maximen verdirbt” (*Religion*, 686). Derudover er sindelaget det begrebsmæssige redskab, der tillader en bestemmelse af den moralske person ud fra en enkelt og entydig moralsk vurdering til trods for en potentiel pluralitet i personens faktiske handlinger.<sup>25</sup> Som “die oberste Maxime” er sindelaget ansvarligt for enhver tendens, der karakteriserer en given persons faktisk realiserede handlinger i maksime-produktionen – inklusiv mønsteret af amoralske maksimer. Kant siger selv, at vi almindeligvis vurderer en persons moral ud fra hans sindelag, men Kant fastholder, at en sådan vurdering bunder i en almen betragtning og *ikke* i en teoretisk erkendelse – selv ikke når det gælder os selv. “[K]ein Beispiel

22 Jf. Michaelson (1990), 53.

23 Caswell (2006), 191ff. redegør udmærket for *Gesinnung*. Jf. yderligere Michaelson (1990), 53ff.; O’Connor (1985), 290.

24 Jf. yderligere Muchnik (2006), 434 særlig note 12.

25 For konsekvenserne af en tilgang, der undervurderer dette aspekt, jf. Muchnik (2006), 430ff.



in der äussern Erfahrung [ist] adäquat... als welche das Innere der Gesinnung nicht aufdeckt ... obzwar nicht mit strenger Gewissheit" (*Religion*, 716). Vores handlingers virkelige moralske status er altså fuldkommen skjult – selv for os selv.<sup>26</sup> Begrebet om sindelaget kan dermed siges at *rejse* lige så mange spørgsmål, som det *løser*.

Sindelaget er en enhed. Dette er en konsekvens af sondringen mellem fænomener og noumena. Da vores moralske karakter *er* sindelaget, kan vi ikke beskrive det med begreber hentet fra den fænomenale sfære, da det fænomenale og det noumenale ikke kan forenes. Konsekvensen af dette er, at en persons sindelag ikke kan beskrives ud fra en tidslig orden eller som en forandring over tid, fordi det er noumenalt.<sup>27</sup> Principielt afværger sindelaget dermed ethvert forsøg på periodisering og kan kun beskrives som en enhed.<sup>28</sup> "Den das moralisch subjektive Prinzip der Gesinnung... ist... nicht von der Art, dass sein Dasein in Zeitabschnitte teilbar, sondern nur als absolute Einheit gedacht werden kann, und da wir auf die Gesinnung nur aus den Handlungen (als Erscheinungen derselben) schliessen können, so wird das Leben zum Behuf dieser Schätzung nur als Zeiteinheit, d.i. als ein Ganzes, in Betrachtung kommen" (*Religion*, 725 noten). Kant opererer altså med to rækker af begreber, der på den ene side omhandler et enhedsmæssigt, tidløst sindelag og på den anden side manifestationen af sindelaget i almindelige handlingsudslag af maksime-produktionen. På den ene side står det første og frie valg, der bestemmer sindelaget, og på den anden side står den specifikke frie handling, der er resultatet af maksime-produktionens handlinger. Overordnet kan man altså tale om valg af primær og sekundær orden, hvor den primære orden udgøres af valget af den øverste maksime og den øverste maksimes valg. Den sekundære orden omfatter valget af enhver partikulær maksime i den daglige maksime-produktion.

Problemet med de to niveauer af maksimer er, at hvis sondringen ikke kan opretholdes, er det ikke tydeligt, hvordan Kant kan deducere opfattelsen af sindelaget som noget distinkt andet ud fra andre maksimer af lavere orden. Forklaringen på dette skal findes i sondringen mellem de to former for handling eller valg (Tat) (*Religion*, 679). Den første handling angår udøvelsen af frihed i indoptagelsen eller konstitueringen af sindelaget, og den anden handling angår udøvelsen af frihed, hvorved specifikke maksimer vælges ud fra den

26 Dette slås allerede fast i KRV (501 (A 552)), jf. yderligere Sussman (2005), 100; O'Connor (1985), 290ff.

27 Caswell (2006), 195f. forsøger at argumentere for den forbindelse mellem en fænomenal og noumenal sfære, som Kant aldrig selv angiver. Caswell ender dermed – i en ellers sympatisk fremstilling – med at modsige sig selv og Kant.

28 Sussman (2005), 99 diskuterer dette og ser – som en af de få – det paradoksale i sindelagets tidslighed.

praksis, der er etableret på baggrund af sindelaget.<sup>29</sup> På dette punkt opstår dog et fundamentalt problem for Kant: sondringen mellem handling og handling resulterer i, at selvet fordobles, hvis ikke det ligefrem mangfoldiggøres i det uendelige. Selvet eller subjektiviteten må nemlig *både* være det, der handler i indoptagelsen af sindelaget *og* i de specifikke handlinger, der er følger af sindelagets moralske struktur. Konsekvensen af denne opsplittning af selvet er selve udtrykket for det paradoksale i Kants fremstilling, da sondringen mellem de to niveauer er et forsøg på at forsvare den moralske subjektivitets enhed. Kant kan således sige at svæve i et forklaringsmæssigt vakuum, da han fastholder internt modstridende påstande.<sup>30</sup>

For Kant bliver sondringen mellem handling og handling på et dybere niveau udtryk for, at subjektivitet i langt højere grad kan beskrives som en form for handling end som essens, da sindelaget selv er en maksime.<sup>31</sup> I kraft af at det er en maksime, er det moralske selv *allerede* en form for vilje,<sup>32</sup> der logisk går forud den initierende viljehandling. Problemet opstår i det omfang, man forsøger at redegøre for denne oprindelige handling således, at den *forbliver* en handling og ikke opløses i et formløst kaos uden struktur, hvor den moralske bedømmelse ikke kan indfange den. Det, Kant forsøger at indfange med de to niveauer af handling, er et fast ståsted i viljens struktur, der er placeret 'bagved' selvet. Problemet er bare, at et sådant ståsted ikke kan opretholdes, fordi ideen om en moralsk subjektivitet, der tilhører den noumenale verden, præsenteres med begreber fra handlingernes univers, der tilhører den fænomenale verden.<sup>33</sup> Selvom sondringen således kan bruges til at forklare processen omkring valget af den øverste maksime og sindelagets konstituering, så griber den forkert, når det drejer sig om at belyse relationen mellem sindelaget og specifikke handlinger i maksime-produktionen. Og da Kant ikke tilbyder noget princip for integreringen af de to sfærer, afløses det forklaringsmæssige potentiale af en underdetermination og et vakuum. Dette er en afgørende mangel ved Kants fremstilling, da det angår selve kernen i det problem, Kant forsøger at komme overens med. Begrebet om

---

29 Muchnik (2006), 430-42 diskuterer forudsætningerne for sondringen mellem de to handlinger (Tat) i *Grundlegung*.

30 Problemet er også til stede som et erkendelsesteoretisk paradoks; det første bud blandt alle pligter over for en selv er, at man skal kende sig selv i moralsk forstand (MS, 576 (A 104)). Men ifølge *Religion* er det ondes rod, at man bedrager sig selv i troen på, at man kender sine motiver og handler korrekt ud fra dem (*Religion*, 687).

31 Caswell (2006), 191 har som en af de få set dette.

32 Sussman (2005), 95 har misforstået koblingen af sindelaget med det onde, da han hævder, at der ikke er en ond vilje.

33 O'Connor (1985), 292ff. identificerer problemet, og jeg følger ham.

sindelaget løser dermed ikke Kants problem om den moralsk handlende persons enhed.

Ovenfor blev det sagt, at sindelaget ikke måtte forstås som en statisk essens, selvom Kant gentagne gange anvender metaforen om en grund med reference til sindelaget. Sindelaget skal i stedet forstås som en form for handling, der kræver sin egen motiverende kilde – en strukturerende struktur. Sindelaget kaldes også eksplicit en maksime (*Religion*, 666). Baggrunden for dette er, at hvis maksime-produktionen finder sted som en naturlig impuls, forsvinder menneskets ansvarlighed. Ikke desto mindre må den endelige grund til indoptagelsen af maksimer ligge i vælgeevnen (*Willkür*). Men dermed bliver sindelaget både grunden for maksime-produktionen og selv en maksime, der dog ikke yderligere kan redegøres for. Maksime-produktionen konstrueres dermed som en kæde, hvor en fri handling hviler på en anden fri handling, der er produktionen af den øverste maksime. Og grundlaget for denne kæde forklarer Kant aldrig, hvorved jagten på et fast fodfæste i viljens struktur aldrig tilvejebringes. Dermed står en uendelig regres og truer kæden af maksimernes orden, for hvis en frit produceret maksime har en bestemt underliggende grund, der selv er frit produceret,<sup>34</sup> så afføder det nødvendigvis spørgsmålet om, hvad denne mest grundlæggende grund er – *un so weiter*. Hvis ikke netop dette spørgsmål stilles, forbliver det uvist, hvorfor Kant ikke selv går videre med analysen på dette punkt, hvor grundenes grund og den øverste maksime er på spil. Og at Kant påberåber sig en systematisk agnosticisme<sup>35</sup> i forsøget på at redegøre for den endelige grund for maksime-produktionen, afslører blot det paradoksale i konstruktionen af handlingernes kædereaktion. “[D]er Mensch ... enthält einen (uns unerforschlich) ersten Grund der Annehmung ... Maximen” (*Religion*, 667-8). Det problematiske og paradoksale i Kants fremstilling er, at der ikke kan siges mere om denne første grund. “Weil wir also diese Gesinnung, oder vielmehr ihren obersten Grund nicht von irgend einem ersten Zeit-Actus der Willkür ableiten können, so nennen wir sie eine Beschaffenheit der Willkür, die ihr ... von Natur zukömmt” (*Religion*, 672). Sindelaget er altså resultatet af en fri noumenal handling uden for tiden, hvorom intet kan siges. Dermed tilvejebringes *ikke* et forklaringspotentiale for sindelaget men derimod en underdetermination.

Her er vi ved selve hjertet i teorien om, hvorfor man bliver god eller ond. Dette punkt er det tætteste, vi kan komme på den grundlæggende handling, hvor den øverste og første maksime vælges, og hvor den

34 O'Connor ser problemet men konkluderer, at Kants redegørelse er fyldestgørende (O'Connor (1985), 291).

35 Begrebet er hentet hos Michaelson (1990), 59.

radikale ondskab hører hjemme. Det er i dette valg, Willkür initierer en drivfjeder-struktur, der resulterer i alle deraf følgende handlinger i maksime-produktionen. Og det er et frit valgt valg, da Kant netop vil undgå enhver determinisme eller arvesynd, der kan fjerne ansvarligheden fra mennesket. Det paradoksale i Kants fremstilling er det forsvindende punkt, der forbinder den oprindelige moralske handling med det efterfølgende kaos og ubestemmelighed. Det er således ikke muligt at lægge vægten på denne første og initierende handling af Willkür i forsøget på at forklare rollefordelingen mellem denne og de efterfølgende handlinger. Ikke desto mindre er det netop denne første og initierende handling, der understreger subjektets og handlingens enhed. Omvendt er det lige så umuligt at fokusere på fundamentet eller motivationen for denne første handling, da den er kontingent og ubestemmelig. Det, der drejer sig om på dette sted, er vælgeevnen, der giver sig selv en drivfjeder, *inden* den har mulighed for at give sig selv en drivfjeder.<sup>36</sup> Og i kraft af den drivfjeder, den giver sig selv, vil vælgeevnens generelle praksis angående maksime-produktionen blive etableret.<sup>37</sup> Problemet er bare, at Kant ikke vil angive, *hvad* den grund er, eller *hvorfor* den vælges, da han så forklarer friheden, hvilket man ikke kan; friheden er fri – *i.e.* vilkårlig og absolut spontan. Det, der skal forklares, hverken kan eller vil Kant altså forklare, hvorfor fremstillingen nødvendigvis må ende med at slå sig selv for munden.

## Konklusion

I forsøget på at redegøre for menneskets ansvarlighed og tilregnelighed i skyld ender Kant med at afsløre sig som mere paulinsk end pelagiansk – ondskaben er medfødt (selvom den er frit valgt). En af konsekvenserne af denne kejtede position er, at den frie vilje ikke blot bliver sit eget problem, men bliver decideret uforståelig og et mysterium for sig selv i mødet med det radikale onde. Uanset hvor omfangsrigt Kants system ender med at være, bliver han uforståelig for sig selv. Dette afspejler sig i den gentagne anvendelse af ordet “*unerforschlich*” og den argumentative agnosticisme. Og det er ikke kun småting, der er uransagelige; vi ved hverken, hvordan vi ‘falder’ i ondskab, eller hvordan vi kommer ud af den igen. Og uanset hvor

---

36 Herrera ender sin artikel med det yderst relevante spørgsmål på dette niveau om, hvordan moralsk handlen overhovedet finder sted (Herrera (2000), 409).

37 Det er oven for blevet hævdet, at mennesket antropologisk set *er* vælgeevnen i en kantiansk optik. Påstanden er funderet i indsigten om, at vælgeevnen (Willkür) er det, hvorfra alt begynder. Problemet opstår, når Kant skriver, at Willkür først handler, når en drivfjeder (Triebfeder) er blevet indoptaget i en maksime (*Religion* 670), og alligevel lader Willkürs initierende og første handling foregå uden en drivfjeder.

moralsk årvågen man er, ved man ikke med sikkerhed, om man er moralsk god eller ond, da både ens egen og andres "oberste subjektive Grund" forbliver dunkel og uigennemskuelig. En sådan begrænsning af den moralske selv-indsigt resulterer i et indgreb på friheden, da vi aldrig kan være sikre på, at vi gør, hvad vi burde. På trods af Kants forsøg på at sikre tilskrivelsen af ansvarlighed til den moralsk handlende person ud fra sindelaget ender han med det resultat, at fundamentet for moralen – fornuften – ikke er herre i eget hus.

Det spændende ved *Religion* er, at det først er, når vi forstår, *hvorfor* Kant modsiger sig selv og ender med ikke at have forklaret noget som helst, at vi begynder at kunne værdsætte hans betydning og relevans i forsøget på at finde fodfæste i spørgsmålet om det onde. Det er først, når vi indser det paradoksale i forsøget på at *forklare* ondskaben – forsøget på at lade det onde og det gode (fornuften) udgå fra samme kilde (friheden) – det bliver tydeligt, hvor radikalt og personligt spørgsmålet må stilles. Og samtidig med indsigten om, at det onde synes at betvinge den menneskelige fornuft, melder spørgsmålet sig igen og igen om, hvorfor vi bliver ved med at vende tilbage til påstanden om, at mennesket besidder en medfødt, radikal ondskab. Dette er et subjektivt spørgsmål, der angår den enkelte. Og om end vi ikke overordnet kan angive et svar, kan vi i det mindste værdsætte, at Kant har lokaliseret et uomgængeligt problem, der bliver ved med at melde sig i menneskehedens historie.