

# *O Problema da Fraqueza da Vontade na Filosofia Prática Kantiana*

*Antonio Frederico Saturnino*

UFRJ

O objetivo do presente artigo é analisar o problema da fraqueza da vontade no quadro conceitual da filosofia moral kantiana. De início, a fim de esclarecer de que modo o problema deve ser posicionado na teoria da ação kantiana, servimo-nos do contraste com as análises de Donald Davidson sobre o fenômeno da fraqueza da vontade. O contraste com Davidson permite evidenciar que o modelo silogístico do raciocínio prático condiciona o modo como o fenômeno da fraqueza pode ser admitido e explicado no quadro conceitual kantiano. Isto significa que qualquer explicação kantiana do fenômeno tem de enfrentar uma dificuldade imposta pelo modelo silogístico, a dificuldade, a saber, de que nenhuma máxima grosso modo má pode ser silogisticamente derivada da Lei Moral, como único princípio geral que se justifica por si mesmo, de modo absoluto e incondicionado. Levando em conta esta dificuldade, e levando em conta que, no quadro conceitual kantiano, a fraqueza da vontade tem a ver com máximas más em sentido amplo, discutimos duas maneiras pelas quais o mal em sentido amplo poderia ser explicado dentro do sistema kantiano. A primeira maneira é problemática, por depender de uma noção que Kant em princípio rejeita, a de uma escolha indiferente à exigência de justificação. Como alternativa a ela, procuramos desenvolver a hipótese de que uma derivação desatenta e negligente a partir da Lei Moral poderia tornar compreensíveis os diversos estágios da maldade em sentido amplo do ser humano. E procuramos sustentar ainda que os dois primeiros estágios da propensão ao mal podem ser considerados correspondentes ao fenômeno usualmente chamado de fraqueza da vontade.

\*\*\*

Na *Introdução* a seu livro *Fraqueza da Vontade*, William Charlton explica que se fala de “fraqueza da vontade” em dois sentidos, às vezes como um fenômeno a ser explicado, às vezes como um tipo de explicação para este fenômeno (CHARLTON, 1988, p.1). A exposição de Charlton deixa claro que as explicações apresentadas para o fenômeno da fraqueza da vontade dependem da descrição que se faz do mesmo, na medida em que, dependendo da descrição, o fenômeno apresenta-se como algo de paradoxal e extraordinariamente problemático - e até mesmo impossível -, ou, ao contrário, como algo que pode ser tranquilamente explicado por uma análise mais nuançada do raciocínio prático e da ação livre (*IBIDEM*, p.3-4). Assim, o primeiro problema que o tema suscita diz respeito à descrição que se vai adotar do fenômeno da fraqueza da vontade -conforme a descrição adotada, a explicação vai se configurar mais como resolução de um paradoxo ou, antes, como esclarecimento do equívoco que transformou algo facilmente explicável em um grande problema. Cabe destacar desde já que a segunda espécie de explicação não é menos relevante e fecunda do que a primeira; afinal, como afirma Davidson em influente artigo sobre o tema, em resposta ao comentário de E. Lemmon de que o problema relativo à existência e à possibilidade da fraqueza da vontade seria um “pseudoproblema”, - como afirma Davidson, “como você pode saber que cometeu um equívoco sem saber qual é o equívoco, seu problema pode ser real” (DAVIDSON, 1970, p.23).

De um modo geral, as descrições do referido fenômeno envolvem três elementos essenciais, estreitamente relacionados entre si. Em primeiro lugar, o agente faz algo que ele julga ser mau, ou deixa de fazer algo que ele julga ser bom; em segundo lugar, ao fazê-lo o agente age intencionalmente, ele de algum modo quer ou quis praticar a ação; em terceiro lugar, teria sido possível ele fazer outra coisa, mais adequada ao seu juízo sobre o que é bom ou mau. Ora, cada um desses elementos pode receber uma interpretação que suaviza o caráter paradoxal do fenômeno da fraqueza da vontade. Não é o caso de analisarmos aqui as diversas possibilidades de interpretação suavizadora. Entretanto, tendo em vista nosso propósito de utilizar o contraste com Davidson para analisar o problema da fraqueza da vontade no quadro da filosofia prática kantiana, gostaríamos de mencionar uma determinada possibilidade de suavização, que entra em jogo tanto nas análises de Davidson quanto na teoria kantiana.

A possibilidade a que estamos nos referindo é a seguinte. O juízo contra o qual o agente age pode ser um juízo de algum modo distinto daquele juízo que, na esfera das ações intencionais, determina o que um agente intenciona ou quer fazer no momento mesmo da ação.

Dependendo do tipo de distinção entre os dois juízos, o problema da fraqueza da vontade perde a feição de paradoxo e ganha o aspecto de problema relativamente corriqueiro. Por exemplo, é relativamente comum o agente contrariar certos juízos sobre o bem que ele toma como princípios corretos e valiosos, princípios que ele preza e gostaria sinceramente de seguir, princípios que em certo sentido ele quer cumprir, “tem vontade” de cumprir. Este fenômeno pode sem dúvida ser chamado de “fraqueza da vontade”. Na teoria kantiana da ação, como veremos abaixo, este tipo de fraqueza ocupa um lugar importante, constituindo-se na verdade em um problema teoricamente mais complexo do que os conhecidos conflitos entre desejos de primeira e de segunda ordem. Mais precisamente, na teoria kantiana este fenômeno é uma das expressões da “maldade” (em sentido amplo) do ser humano, que Kant chama de “fraqueza” ou “fragilidade” do coração humano (KANT 1793, B 21-22, Ak 6:29-30). No quadro conceitual da filosofia prática kantiana, não é fácil explicar não apenas esta, mas todas as demais expressões da maldade do ser humano. De qualquer modo, entretanto, é inegável que este tipo de fraqueza da vontade é bem distinto do fenômeno paradoxal que Davidson primordialmente focaliza, ou seja, é bem distinto do fenômeno que alegadamente ocorre quando a pessoa contraria aquele juízo singular e situado que, na esfera das ações intencionais, costuma ser identificado àquilo que o agente intenciona ou quer fazer no momento mesmo da ação. Este último fenômeno em princípio é mais paradoxal do que a fraqueza mencionada por Kant na *Religião nos Limites da Simples Razão*. Será que a fraqueza ou “incontinência” focalizada por Davidson encontra espaço na teoria da ação kantiana?

Na *Introdução à Metafísica dos Costumes*, Kant explicita uma distinção que, segundo Allison (ALLISON, 1990, p. 129), “já é operativa na *Crítica da Razão Prática* e desempenha um papel primordial na *Religião nos Limites da Simples Razão*”, a saber, a distinção entre *Wille* (Vontade) e *Willkür* (Arbítrio). De acordo com Allison, “Kant usa os termos *Wille* e *Willkür* para caracterizar respectivamente as funções legislativa e executiva de uma faculdade unificada de volição, que ele igualmente designa como *Wille*” (IBIDEM, p.129). No âmbito da explicação desta distinção, Kant afirma que “As leis procedem da vontade; as máximas, do arbítrio” (KANT 1797, AB 26, Ak 6:226). Para Kant, como é sabido, dizer que as leis práticas procedem da vontade é dizer não apenas que tais princípios são prescritos pela vontade, mas também que eles são queridos pela vontade como faculdade de desejar superior. No âmbito da distinção entre vontade e arbítrio, portanto, pode sem dúvida haver divergência entre o (bom) princípio prescrito e querido pela vontade e, por outro lado, as máximas e ações (más) adotadas pelo livre-arbítrio. Esta

divergência, como já foi dito, não apenas pode ser chamada de fraqueza da vontade, como constitui-se em problema de difícil explicação no quadro do sistema kantiano, como veremos abaixo. Entretanto, no âmbito da teoria kantiana, parece não ser possível o fenômeno que Davidson primordialmente focaliza como fraqueza da vontade, o qual consiste, como dito acima, no fenômeno que alegadamente ocorre quando a pessoa contraria aquele juízo singular e situado que, na esfera das ações intencionais, costuma ser identificado à intenção do agente, ou seja, àquilo que o agente tenciona ou quer fazer no momento mesmo da ação.

Expliquemos a afirmação que acaba de ser feita. Para tanto, cabe lembrar que Kant não analisa a ação intencional como ação imediatamente determinada por um juízo singular e situado, mas, sim, como ação derivada de uma máxima, isto é, um juízo de caráter mais geral. A este respeito pode-se citar como exemplar a seguinte passagem da *Fundamentação da Metafísica dos Costumes*: “Toda coisa da natureza atua segundo leis. Só um ser racional tem a faculdade de agir segundo a representação das leis, isto é, segundo princípios, ou uma vontade (*Wille*). Visto que se exige a razão para derivar de leis as ações, a vontade (*Wille*) nada mais é do que razão prática” (KANT 1785, BA 36, Ak 4:412). Dois esclarecimentos são importantes para a compreensão desta passagem. Em primeiro lugar, como argumenta persuasivamente Allison (ALLISON, 1990, p.86), as “leis” cuja representação determina as ações do ser racional são as máximas, que Kant em algumas ocasiões caracteriza como “leis subjetivas”, e que se distinguem das leis objetivas pelo fato de serem os princípios segundo os quais os sujeitos racionais agem, e não os princípios segundo os quais eles devem agir. Em segundo lugar, na *Fundamentação* Kant não recorre explicitamente à distinção entre vontade e arbítrio, de tal modo que na nossa passagem o termo *Wille* pode referir-se a qualquer das duas capacidades que serão posteriormente distinguidas, a vontade em sentido estrito e o livre-arbítrio.

Em relação a esta questão, cabe argumentar que, embora a distinção entre *Wille* e *Willkür* não seja estritamente operativa na *Fundamentação*, ela corresponde a uma distinção que é latente nesta obra, a distinção, justamente, entre as funções legislativa (*Wille*) e executiva (*Willkür*) de uma faculdade de volição unificada. Ora, admitindo-se que a distinção entre as funções legislativa e executiva da vontade em sentido amplo possa ser retrospectivamente aplicada ao texto da *Fundamentação* acima citado, parece razoável afirmar que o termo *Wille* se refere nele à faculdade executiva da vontade, isto é, ao arbítrio. Com efeito, “agir segundo a representação das leis” (que seriam as leis ou princípios segundo os quais o sujeito age, as máximas) parece

claramente corresponder à função executiva específica do arbítrio. Assim, poderíamos propor uma paráfrase da última frase da citação da *Fundamentação* feita acima (ou seja, a frase “Visto que se exige a razão para derivar [ableiten, Ableitung] de leis as ações, a vontade (*Wille*) nada mais é do que razão prática”); a paráfrase seria: “Visto que se exige a razão para derivar de máximas as ações, o arbítrio nada mais é do que razão prática”. Esta paráfrase corresponde perfeitamente à concepção da relação entre ações e máximas implícita na seguinte passagem da *Religião*: “Chama-se, porém, mau a um homem, não porque pratique ações que são más (contrárias à lei), mas porque estas são tais que deixam inferir (*schliessen*) nele máximas más” (KANT 1793, BA 5, Ak 6:20). Nos dois casos, a relação entre as máximas e as ações, que corresponde à função do ajuizamento prático no momento da ação, é compreendida como uma relação mais lógica do que perceptiva ou intuitiva, uma relação mais de derivação ou inferência do que de “visão certa” de todas as considerações relevantes na situação da ação.

Tendo em vista o contraste com Davidson que será estabelecido logo a seguir, o argumento que estamos querendo defender é o seguinte. A teoria da ação kantiana está fundada em uma exigência de explicabilidade semelhante àquela que caracteriza sua teoria da experiência objetiva. Esta última, com efeito, é marcada pela exigência de que todo evento possa ser explicado como resultado de um princípio (lei) geral em conjunção com uma circunstância singular subsumida a este princípio. De modo análogo, a teoria da ação é marcada pela exigência de que toda ação possa ser explicada como resultado de um princípio geral - uma máxima - em conjunção com uma circunstância singular identificada como caso desta máxima. A diferença essencial consiste no fato de que, enquanto o conteúdo da lei da natureza é compreendido como oriundo de uma natureza externa e independente do sujeito, o conteúdo da máxima é compreendido como oriundo de uma capacidade de escolha do próprio sujeito, seu livre-arbítrio. Em ambos os casos, porém, a explicação (do evento ou da ação) remete ao modelo do silogismo, ou seja, explicar consiste em apresentar o *explanandum* como conclusão de um silogismo que tem como premissa maior uma lei da natureza ou uma máxima, e como premissa menor uma circunstância singular identificada como caso do princípio geral.

Retornemos agora à ideia apresentada acima, de que Kant não analisa a ação intencional como ação imediatamente determinada por um juízo singular e situado, mas, sim, como ação derivada de uma máxima, isto é, um juízo de caráter geral. Ora, apesar deste foco na máxima como princípio geral, parece plausível afirmar que Kant até poderia admitir um juízo singular e

situado em sua teoria da ação intencional, mas não poderia admitir a possibilidade de qualquer divergência entre este juízo e a ação praticada, na medida em que a relação entre o (presumido) juízo situado e a ação praticada teria de ser uma relação de derivação ou inferência, nos mesmos moldes da relação entre a máxima e a ação, cujo caráter silogístico parece ser essencial à teoria da ação kantiana. Ora, uma relação de derivação ou inferência não abre espaço para aquele tipo de divergência que explica o que Davidson chama de fraqueza da vontade.

O contraste com Davidson é fecundo para a continuidade das reflexões sobre o problema da fraqueza da vontade na filosofia kantiana. Passemos então a uma breve exposição do influente argumento de Davidson. O cerne de seu argumento consiste na análise da diferença entre o modelo silogístico do raciocínio prático, e, por outro lado, um modelo por assim dizer intuitivo ou perceptivo, que abre espaço para outra diferença, interna a ele, a diferença entre um juízo intuitivo do tipo “considerando todas as coisas”, que, apesar de ligado à situação de escolha, tem um caráter relacional e condicionado, e, em segundo lugar, um juízo intuitivo de caráter categórico e incondicional, que exprime a intenção do agente no momento mesmo da ação. Podemos apresentar o argumento (e a diferença essencial nele operativa) partindo de um exemplo do modelo silogístico. Se um agente adota o princípio geral de que “se um prato é gorduroso, ele deve ser evitado” e identifica certo prato como gorduroso, ele endossa o juízo (a conclusão) de que “é mau comer este prato de comida”. Suponhamos agora que mesmo assim ele coma o referido prato. Ainda que isto possa ser chamado de fraqueza da vontade, o comportamento poderia ser explicado pelo fato de que, na mesma situação, o agente desenvolve outro silogismo prático, paralelo ao acima indicado. Com efeito, na mesma circunstância ele pode também adotar o princípio geral de que “se uma situação é prazerosa, ela deve ser aproveitada”, e identificar a circunstância do prato estar à sua frente como um caso desta regra geral, o que o levaria a endossar o juízo (conclusão) de que “é bom aproveitar este prato de comida”.

Tal como apresentada até aqui, a situação tem o aspecto de um conflito entre “razão (prudencial)” e “impulso”, e de certo modo aproxima-se do conflito entre desejos de segunda e de primeira ordem rapidamente mencionado acima, na medida em que o fenômeno poderia ser explicado como um daqueles casos de conflito entre o princípio/desejo “racional” que o agente preza e gostaria de seguir e o princípio/desejo “impulsivo” que ele de fato segue em determinada situação. Mas o interesse de Davidson é argumentar que o modelo silogístico da deliberação é totalmente inadequado para dar conta deste tipo de conflito, que só pode ser adequadamente

analisado por meio de outro modelo, centrado na variedade das perspectivas perceptuais sob as quais uma mesma situação pode ser considerada. Para acentuar o contraste, insistamos por um momento no modelo silogístico, supondo que o agente viesse a desenvolver um terceiro silogismo prático na mesma situação, como, por exemplo, o seguinte. Premissa maior: “se um comportamento é prazeroso para minha avó, ele é bom”; premissa menor: “aproveitar este prato preparado por minha avó dará grande prazer a ela”; conclusão: “é bom e justo aproveitar este prato de comida”.

O que agora emerge com clareza é o fato de que este modelo silogístico da deliberação só pode ser adotado caso se acredite na possibilidade de resolver os inevitáveis conflitos entre princípios gerais (“premissas maiores”) por meio de um princípio supremo e único - como Kant em princípio acredita. Caso não se acredite nesta possibilidade (como é o caso do próprio Davidson), a única alternativa seria reformular cada um dos silogismos acima, transformando-os em argumentos sobre o que é *prima facie* bom na situação em que o agente se encontra. Por exemplo, “À luz do fato de se tratar de um prato gorduroso, evitar o prato é *prima facie* melhor do que comê-lo”; e “À luz do fato de se tratar de um prato saboroso e prazeroso, aproveitar o prato é *prima facie* melhor do que evitá-lo”; e ainda “À luz do fato de que aproveitar o prato daria grande prazer à minha avó, saborear o prato é *prima facie* melhor do que evitá-lo”. A conclusão geral deste conjunto de argumentos sobre o que seria *prima facie* bom ou melhor seria um enunciado do tipo “considerando todas as coisas” - digamos “Nesta situação, considerando todas as coisas, aproveitar o prato revela-se como melhor do que evitá-lo”.

Davidson enfatiza que, à diferença das conclusões silogísticas, esta última conclusão não deriva logicamente das diversas considerações parciais que a embasam, ela não pode ser extraída destas considerações por nenhum procedimento lógico ou formal. E é justamente por isto que este último tipo de conclusão tem um caráter condicional ou relacional, e não categórico e incondicional (DAVIDSON, 1970, p.38-39). Em outras palavras, uma conclusão do tipo “todas as coisas consideradas” estabelece uma relação entre, por um lado, a totalidade das considerações relevantes na situação de escolha e, por outro lado, a afirmação de que certo comportamento aparece como melhor do que outro nesta situação; mas tal conclusão não afirma categórica e incondicionalmente que «X é melhor do que Y». Ora, o juízo que, no âmbito das ações intencionais, costuma ser identificado à intenção do agente no momento da ação é, justamente, o juízo categórico e incondicional de que “X é melhor do que Y” - virtualmente

idêntico a “escolho fazer X, e não Y”. Para Davidson, embora de fato não exista a possibilidade de um agente livre e intencionalmente contrariar o juízo categórico e incondicional de que “nesta circunstância, X é melhor do que Y” (isto é, não existe a possibilidade de ele efetuar este juízo e concomitantemente fazer Y em vez de X), - embora esta possibilidade de fato não exista, existe por outro lado a possibilidade de o agente livre e intencionalmente contrariar o juízo relacional e condicional de que “Nesta situação, considerando todas as coisas, X revela-se como melhor do que Y”. O comportamento do agente evidenciaria neste caso que no momento mesmo da ação ele passou deste juízo condicional e relacional para o juízo categórico e incondicional de que “Y é melhor do que X”.

Para Davidson, embora este comportamento seja possível, ele claramente é irracional, não apenas no sentido de violar o princípio da racionalidade prática que ele chama de “princípio da continência” - “realize a ação julgada melhor com base em todas as razões relevantes disponíveis” (DAVIDSON, 1970, p.41), mas também, ou concomitantemente, no sentido de que o agente não tem nenhuma razão para fazer o que ele faz, ou melhor, não tem nenhuma razão para no momento da ação substituir o juízo relacional que ele endossa como sua melhor razão pelo juízo categórico que acaba determinando sua intenção, sua escolha e sua ação (IBIDEM, p.42). Para Davidson, um juízo do tipo “considerando todas as coisas” é um juízo que, de modo perceptual ou intuitivo, incorpora e processa a efetiva relevância deliberativa de cada uma das razões em jogo, tanto daquelas razões que exprimem mais a identidade que o agente preza e gostaria de desenvolver quanto daquelas que exprimem mais a identidade que o agente de fato tem ou está tendo. A passagem de um determinado juízo relacional para um juízo categórico contrário não remete a inconsistências deliberativas que atuam no plano da ponderação das razões que se tem, mas remete a uma influência causal surda destas mesmas razões. Para Davidson, em outras palavras, a passagem de um juízo a outro deve ser explicada pelo fato de que as razões operantes em certa pessoa muitas vezes “causam” juízos e comportamentos distintos daqueles que a pessoa acredita serem justificados pelas razões nela operantes. Em suas palavras, “Não há paradoxo em supor que uma pessoa algumas vezes sustente que tudo aquilo que ela acredita e valoriza apoia determinado curso de ação, quando ao mesmo tempo estas mesmas crenças e valores causam nesta pessoa uma rejeição deste curso de ação” (IBIDEM, p.41).

Não há oportunidade aqui para discutir a explicação proposta por Davidson. Importa apenas tentar situá-la na perspectiva kantiana. De um ponto de vista kantiano, a explicação



parece sugerir que a pessoa de fato não seria livre para fazer outra coisa. Mais precisamente, de um ponto de vista kantiano, a explicação de Davidson teria de ser colocada no plano das explicações dos eventos naturais, distinto do plano da compreensão (e justificação) das ações livres - o que significa que o fenômeno explicado por Davidson teria de ser colocado no plano dos eventos naturais, e não poderia ser classificado como fraqueza da vontade, uma vez que fraqueza da vontade pertence ao âmbito da liberdade humana. Com efeito, embora não trabalhe com um dualismo entre “substância pensante” e “substância extensa”, Kant trabalha com uma distinção essencial entre “sujeito passivo” e “sujeito ativo”, à qual corresponde a distinção entre uma explicação por causas, na qual o conceito de causa limita-se a ligar uma diversidade de aparições que são dadas a um sujeito essencialmente passivo, e, por outro lado, compreensão e justificação por razões, em que a diversidade do que aparece emerge a partir de conceitos e princípios (razões) adotados pela pura atividade do sujeito.

Do ponto de vista kantiano, portanto, tomada como fenômeno a ser explicado - ou melhor, compreendido -, a fraqueza da vontade consiste não simplesmente em uma falha na justificabilidade das ações, mas, mais precisamente, em uma falha que, primeiro, não pode ser explicada por qualquer influência meramente causal, e, segundo, remete a uma falha na justificabilidade do princípio (máxima) a partir do qual tais ações podem ser compreendidas. Cabe destacar a diferença entre compreensão e justificação que está em jogo aqui. Uma máxima má permite compreender uma ação má, como ação que diverge do princípio bom prescrito e querido pela vontade; ela permite compreendê-la na medida em que a ação má pode ser derivada da máxima como princípio geral. Mas uma máxima má não pode em última instância ser justificada (ela só pode ser justificada parcialmente, como produto derivado de outra máxima má, superior, a qual, entretanto, não poderia ser em última instância justificada), o que significa que as ações dela derivadas tampouco podem ser justificadas, apenas compreendidas. Ora, é à falha na adoção da máxima suprema que Kant reserva o termo “maldade” (em sentido amplo) do ser humano. Esta falha não pode ser justificada, embora possa, talvez, ser compreendida - e este é um ponto difícil e obscuro da filosofia kantiana, como veremos logo a seguir. De qualquer modo, uma eventual compreensão desta falha remete a uma falha primordial na relação do arbítrio com o único princípio que se justifica por si mesmo, de modo absoluto e incondicionado, a saber, a Lei Prática prescrita e querida pela Razão Pura Prática, a Vontade em sentido estrito. Uma máxima suprema má é injustificável pelo fato, precisamente, de que o único princípio

que se justifica por si mesmo consiste em uma Lei puramente racional, da qual não pode ser propriamente derivada nenhuma máxima má.

Analisemos as falhas recém mencionadas “de baixo para cima”, ou seja, começando pela falha na justificabilidade das ações, que remete a uma falha na justificabilidade da máxima da qual elas são derivadas. Mais precisamente, comecemos com as falhas mais graves na justificabilidade das ações. Para tanto, pensemos em uma ação claramente contrária à Lei Moral, por exemplo, uma ofensa motivada por pura arrogância. Em princípio, a máxima da qual este ato deriva exprime o grau mais elevado da maldade (em sentido amplo) do ser humano, que Kant chama de “malignidade” (*Bösartigkeit*) ou “perversidade” (*Verkehrtheit*) do coração humano (KANT 1793, B 23, Ak 6: 30). Esta espécie de máxima caracteriza-se por uma inversão da ordem moral dos móveis, uma inversão pela qual motivações nitidamente egoístas são postas na frente do móbil do respeito à Lei - embora este jamais possa ser totalmente suprimido, de acordo com Kant.

O primeiro problema que esta espécie de ação produz consiste em saber se este fenômeno poderia ser classificado como fraqueza da vontade. De acordo com Kant, como é sabido, mesmo o ser humano “maligno” em certo sentido ainda “quer” cumprir a Lei Moral, “tem vontade” de cumpri-la, como é atestado pelo fato de que ele tem o sentimento de respeito por ela. Isto pareceria indicar que o fenômeno da ação maligna poderia ser classificado como fraqueza da vontade. Mas parece mais plausível reservar esta classificação para aquelas ações e máximas compatíveis não simplesmente com esta “vontade” inextirpável de cumprir a Lei, mas compatíveis, mais especificamente, apenas com o comprometimento sincero de cumprir a Lei, que parece ter sido abandonado pelo coração maligno. É verdade que o agente maligno poderia ser apresentado como a vítima mais grave da autoilusão que afeta os seres humanos em geral, pela qual eles escondem suas verdadeiras máximas de si mesmos, disfarçando o quanto elas se desviam da Lei Moral querida por sua vontade. Em pioneiro artigo sobre o problema da fraqueza da vontade em Kant, Alexander Broadie e Elizabeth Pybus enfatizam o papel da autoilusão nas análises kantianas sobre o fenômeno que nos ocupa (BROADIE e PYBUS, 1982, p.410-411). E o próprio Kant caracteriza a propensão para o mal como uma propensão para a “falsidade interna, isto é, uma propensão a mentir para si próprio”, o que explica o fato de que “também a Bíblia (na sua parte cristã) chama ao autor do mal (que em nós próprios habita) o Mentiroso, e assim caracteriza o homem no tocante ao que nele parece ser o fundamento capital do mal” (KANT 1793, B45, Ak 6: 42, nota).

A autoilusão permitiria dizer, talvez, que também o agente maligno poderia arrogar-se um comprometimento sincero com a Lei Moral; - a autoilusão residiria precisamente no fato de que ele inconscientemente disfarça de si próprio o quanto seu comprometimento é fingido e falso. Mas, de novo, parece mais plausível reservar a expressão “comprometimento sincero” para aquele tipo de compromisso que minimamente se evidencia para os observadores das ações do agente. Em outras palavras, parece mais plausível reservar a expressão para aqueles agentes cujas ações na grande maioria das vezes conformam-se à Lei Moral de um modo mais ou menos evidente, e cujas eventuais falhas morais, por isso mesmo, podem ser atribuídas à fraqueza que de algum modo afeta seu comprometimento de resto sincero com a moralidade. Assim, no caso da perversidade ou malignidade do coração humano, dizer que ela é uma propensão para máximas que “pospõem (*nachsetzen*) o móbil procedente da Lei Moral a outros móveis (não morais)” (KANT 1793, B 23, Ak 6: 30), - dizer isto equivaleria a dizer que, embora o sujeito em certo sentido ainda queira cumprir a Lei Moral, trata-se de um querer longínquo e abafado, secundário em relação ao querer egoísta. As exigências da Lei encontram-se a tal ponto obscurecidas pelo “amor-de-si” que não é mais possível dizer que o sujeito valoriza um compromisso efetivo com a moralidade. Assim, parece razoável estabelecer que as ações cabalmente contrárias à Lei Moral, assim como as máximas malignas que podem ser inferidas delas, não correspondem ao fenômeno da fraqueza da vontade.

Entretanto, a análise desta espécie de ação e de máxima é proveitosa para a reflexão sobre as relações entre vontade e arbítrio em Kant, essencial por sua vez para a tentativa de fornecer algum esclarecimento sobre os fenômenos da fraqueza da vontade admissíveis na teoria kantiana. Passemos então à análise das máximas malignas. Considerando que a adoção desta espécie de máxima não é suscetível de explicação pela causalidade dos impulsos e forças da natureza, não é fácil compreendê-la. Mais precisamente, considerando que a máxima maligna não pode ser racionalmente derivada da Lei Moral como único princípio (razão) que se justifica por si mesmo, de modo absoluto e incondicionado; considerando, além disto, que qualquer princípio do qual esta máxima pudesse ser racionalmente derivada seria igualmente maligno, e, portanto, igualmente injustificável à luz do único princípio que se justifica por si mesmo; - considerando tudo isto, a única maneira de compreender a adoção desta máxima ou princípio maligno parece remeter à ideia de uma rejeição simples, consciente e cabal da Lei Moral, como princípio prescrito e querido pela vontade. O arbítrio escolheria rejeitar o princípio querido pela

vontade. Mas isto pareceria equivaler a uma escolha absolutamente arbitrária e indiferente, quer dizer, indiferente a toda exigência de justificação. A fraqueza da vontade equivaleria, portanto, a uma força do arbítrio diante da vontade.

Com base nas análises de Christoph Horn sobre a teoria da vontade de Agostinho (HORN 2006; HORN 2014), parece plausível afirmar que o conceito agostiniano de livre-arbítrio contém traços desta ideia de uma escolha arbitrária e indiferente, contrária às razões associadas à vontade em sentido estrito. Horn mostra como Agostinho procura alcançar um equilíbrio entre duas concepções de vontade nitidamente distintas, oriundas, respectivamente, da filosofia grega e das fontes judaico-cristãs. No âmbito da primeira, a vontade é um impulso ou desejo para o bem em geral, e está naturalmente associada ao intelecto como capacidade de reconhecer ou identificar o bem. Deste ponto de vista, a ação má é vista não como negação consciente da orientação para o bem e do bem reconhecido, mas como falha na identificação do verdadeiro bem, que se traduz no fato de que o impulso para o bem em geral se realiza como impulso para um bem meramente aparente. Trata-se de uma falha essencialmente cognitiva, ainda que causada, talvez, por falhas na educação dos desejos. Em contrapartida, do ponto de vista da concepção de vontade oriunda das fontes judaico-cristãs, a vontade aparece como uma capacidade de decisão que pode conscientemente desviar-se daquilo que é reconhecido como bom e melhor, pode conscientemente negar o bem claramente reconhecido. Segundo Horn, as duas concepções se apresentam na obra de Agostinho, e a partir do texto *Contra Fortunatum* (392) a distinção entre elas tende a aparecer como distinção entre *voluntas*, como impulso para o bem associado ao reconhecimento do bem pelo intelecto, e, por outro lado, *liberum arbitrium*, como capacidade de decisão que pode dirigir-se para o mal mesmo quando o sujeito vê claramente que é um erro (HORN, 2006, p.153).

Não é possível discutir aqui as relações entre as concepções agostiniana e kantiana de vontade e livre-arbítrio. Precisamos nos ater à obra de Kant. E na obra de Kant encontramos pelo menos uma passagem em que ele explicitamente rejeita a noção de uma escolha indiferente, quer dizer, indiferente às razões ativamente reconhecidas como as melhores razões. Trata-se de uma passagem da *Introdução à Metafísica dos Costumes*, que, embora longa, vale a pena citar (KANT 1797, AB 27-28, Ak 6: 226-227):

“A liberdade do arbítrio não pode ser definida pela capacidade de escolher agir a favor ou contra a Lei (*libertas indifferentiae*) - como alguns tentaram -, embora o arbítrio, como fenô-

meno, ofereça frequentes exemplos disto na experiência. (...) As únicas coisas que podemos discernir bem são as seguintes: que, embora o homem enquanto ser sensível mostre, segundo a experiência, uma capacidade de escolher não apenas conforme a Lei, mas também contra ela, a sua liberdade enquanto ser inteligível não pode ser definida por meio disto, porque os fenômenos não podem tornar compreensível nenhum objeto suprassensível (tal como o livre-arbítrio); e que a liberdade jamais pode ser colocada no fato de o sujeito racional poder chegar também a uma escolha conflitante com sua razão (legisladora), ainda que a experiência demonstre com frequência suficiente que isso ocorra (do que, porém, não podemos conceber [*begreiften*] a possibilidade).(...) Apenas em relação à legislação interna da razão a liberdade é, propriamente, uma capacidade; a possibilidade de afastar-se dela é uma incapacidade. Como pode aquela, então, ser esclarecida ou definida (*erklärt*) a partir desta?"

A rejeição da *libertas indifferentiae* apresentada nesta passagem não chega a ser um argumento absolutamente decisivo, até porque ela parece entrar em conflito com uma passagem da *Religião* na qual Kant defende seu "rigorismo" moral, ou seja, a posição de que não há nenhum espaço intermediário entre um agente bom e um agente mau. De acordo com esta passagem (KANT 1793, B10-11, Ak 6:23, nota), como "a Lei Moral é em nós um motivo impulsor (*Triebfeder*), (...) a falta de consonância do arbítrio com a mesma só é possível como consequência de uma determinação real (*realiter*) e oposta do arbítrio, isto é, de uma resistência (*Widerstrebung*) do mesmo". Ora, parece razoável identificar esta "determinação real" do arbítrio, pela qual ele se opõe ou resiste à Lei Moral, - parece razoável identificar essa "resistência ativa" a uma escolha contrária à Lei, efetuada pelo arbítrio do sujeito inteligível. Em outras palavras, parece haver alguma base textual para atribuir a Kant a noção de uma escolha indiferente – quer dizer, indiferente à exigência de justificação racional.

Ao afirmar, na passagem da *Introdução à Metafísica dos Costumes* acima citada, que não podemos conceber como é possível o sujeito racional escolher contra sua razão legisladora, Kant ecoa afirmações semelhantes feitas na *Religião*, quanto ao caráter incompreensível e insondável da presença das máximas más (em sentido amplo) no sujeito racional. Estas afirmações, entretanto, não devem nos levar a desistir do esforço de propor algum tipo de articulação conceitual para esclarecer a presença de tais máximas. Caso a noção de uma escolha indiferente se revele a única opção possível de uma mínima articulação teórica, parece recomendável preferir tal noção às alegações dogmáticas de um mistério insondável, teoricamente inarticulável. Assim,

é preciso avaliar se há outro modo de compreender a presença das máximas más (em sentido amplo) no sujeito racional.

Para proceder ao exame agora proposto, vale a pena inverter a ordem que nos conduziu à noção de escolha indiferente, e começar agora pela análise da compreensibilidade das máximas que, embora más, são “menos más” do que as máximas malignas. Trata-se das máximas impuras e das máximas frágeis, que exprimem a impureza e a fraqueza do coração humano (KANT 1793, B 21-22, Ak 6: 29-30). Como expressões de um grau menor de maldade, as máximas frágeis e impuras talvez abram espaço para outro tipo de compreensão da maldade (em sentido amplo) do coração humano, distinta da compreensão que envolve a noção de uma escolha do arbítrio positivamente contrária à Lei prescrita e querida pela vontade. A hipótese que gostaríamos de defender a partir de agora é a seguinte. Talvez as máximas frágeis e impuras possam ser compreendidas como resultado de uma derivação desatenta e negligente, tomada como um híbrido de falha intelectual e falha volitiva, ou seja, tomada como uma falha na qual a desatenção e a negligência acabam permitindo que princípios maus sejam de algum modo derivados da Lei Moral, sem necessidade da intervenção de uma capacidade propriamente dita de opor-se à Lei, que parece ser o aspecto da “escolha indiferente” rejeitado por Kant. E talvez as máximas malignas possam na sequência ser compreendidas como um acirramento deste processo de derivação desatenta e negligente, de tal modo que, embora a distância e divergência em relação à Lei Moral sejam muito mais nítidas, tampouco seria necessário recorrer à noção de uma capacidade de escolher opor-se à Lei, de modo consciente e ativo. Nossa hipótese, portanto, é a de que os graus sucessivos da maldade (em sentido amplo) do coração humano corresponderiam a um acirramento progressivo do processo de derivação desatenta e negligente de máximas (progressivamente) más a partir da Lei Moral prescrita e querida pela razão prática pura.

Em conformidade com esta hipótese, devemos agora começar pelo primeiro grau da propensão ao mal –uma propensão que, como Allison bem ensina, equivale a uma espécie de máxima suprema, ou disposição original; trata-se do fundamento subjetivo original de todas as máximas (más) de algum modo atuantes nas ações (ALLISON, 1990, p.153). Para marcar a diferença, podemos chamar à primeira de “máxima disposicional”, e às segundas de “máximas de ações”. Assim, o primeiro grau da maldade da máxima disposicional consiste na fraqueza ou fragilidade, que equivale à falha evidenciada em máximas de ação que, embora incorporem a lei moral, e incorporem inclusive a exigência de que ela seja móbil suficiente das ações,

incorporam-na de um modo frágil, de tal modo que as ações derivadas destas máximas de ação frágeis são muitas vezes más, equivalem a descumprimento da lei moral. Kant menciona e explica a “queixa de um Apóstolo” (São Paulo, na Carta aos Romanos): “Tenho, sem dúvida, o querer, mas falta o cumprir, isto é, admito o bem (a lei) na máxima do meu arbítrio; mas este bem, que objetivamente, na ideia (*in thesi*), é um móbil insuperável, é subjetivamente (*in hypothesi*), quando a máxima deve ser seguida, o mais fraco (em comparação com a inclinação)” (KANT 1793, B22, Ak 6:29).

Esta descrição, como já foi dito, corresponde bem ao fenômeno usualmente chamado de fraqueza da vontade. Cabe enfatizar, junto com Thomas Hill Jr., que a fraqueza do querer moral “em comparação com a inclinação” não deve ser entendida como uma incapacidade natural de cumprir a lei, determinada pela maior força física das inclinações. Nas suas palavras, “a comparação com a força e a fraqueza física leva a equívocos aqui” (HILL, 2008, p.221). Hill acentua ainda que, à diferença da malignidade ou perversão, a fraqueza moral em sentido amplo - que parece abranger os dois primeiros estágios da propensão ao mal, a fraqueza em sentido estrito e a impureza -, a fraqueza em sentido amplo é compatível com um comprometimento razoavelmente sincero de cumprir a Lei Moral; de acordo com ele, “aqueles que agem mal devido à fraqueza moral [em sentido amplo - A.S.B.] agem em determinado momento *a partir de uma máxima [de ação - A.S.B.] imoral*, mas contrária à sua razão prática (*Wille*) e ao seu comprometimento geral (*de Willkür*) de agir corretamente” (IBIDEM, p.221. Os grifos são meus).

Retomemos dois pontos da interpretação que estamos tentando defender neste momento. Em primeiro lugar, no quadro da filosofia prática kantiana, o fenômeno usualmente chamado de fraqueza da vontade (no caso uma ação que exprime falta de cumprimento da lei moral sinceramente querida pelo sujeito) pode muitas vezes ser compreendido como resultado do primeiro estágio da propensão ao mal, ou seja, resultado de uma máxima disposicional fraca ou frágil. Em segundo lugar, esta máxima é compatível não apenas com o fato de a lei moral ser querida pela vontade em sentido estrito, mas também com o fato de o arbítrio do sujeito ainda manter um compromisso sincero de cumprir a Lei moral em todas as situações; - a fraqueza da máxima disposicional equivaleria, justamente, à fraqueza deste comprometimento do arbítrio, no sentido de um compromisso frágil, pouco consistente. Seguindo os comentários de Thomas Hill sobre os dois modos de analisar “a metáfora da força de vontade” (IBIDEM, p.221-224), gostaria agora de elaborar esta interpretação, sugerindo que a fragilidade do comprometimento

equivale a uma falta de atenção a respeito do grau de firmeza e de decisão que o cumprimento da lei pode exigir em determinadas circunstâncias. Assim, em virtude desta desatenção, uma máxima disposicional frágil seria derivada da Lei moral, sem ser ativamente escolhida pelo sujeito racional. Trata-se de uma máxima caracterizada por um comprometimento decerto sincero com o cumprimento da lei, mas um comprometimento relativamente frágil, quer dizer, frágil em relação às exigências típicas de certas circunstâncias. E, nas referidas circunstâncias, desta máxima disposicional frágil seriam derivadas máximas de ação em certo sentido imorais - no sentido, a saber, de serem máximas das quais são derivadas ações que se desviam da lei moral.

Em conformidade com a interpretação que estamos tentando defender, podemos propor que o segundo estágio desta derivação desatenta equivale ao segundo estágio da maldade (em sentido amplo) da máxima disposicional, a “impureza” (*Unlauterkeit*). Kant caracteriza este estágio da seguinte maneira: “a máxima é decerto boa segundo o objeto (o seguimento intencionado da Lei), e talvez seja também suficientemente forte para a execução, mas não é puramente moral, isto é, não acolheu em si, como deveria, a Lei sozinha como móbil suficiente; ao contrário, na maioria dos casos (talvez sempre) precisa ainda de outros móveis para determinar o arbítrio àquilo que o dever exige” (KANT 1793, B 22, Ak 6:30).

Ainda seguindo as sugestões de Thomas Hill Jr. acima indicadas, gostaríamos de propor a seguinte interpretação da impureza. Como efeito de um processo de derivação marcado pelo acirramento da desatenção e negligência, da máxima disposicional frágil acaba sendo derivada uma máxima disposicional impura, a qual, mais uma vez, não seria ativamente escolhida pelo sujeito racional. Na máxima disposicional impura, a fragilidade é acrescida de um elemento de vagueza e indeterminação no comprometimento com a Lei Moral, um elemento que abre espaço para um grau maior de autoilusão, de tal modo que, embora este comprometimento ainda possa ser qualificado de sincero, ele é menos límpido do que o da fragilidade, e por isso o agente disfarçadamente permite e até precisa que móveis não morais venham complementar os morais em algumas circunstâncias especiais. Pode-se discutir até que ponto a impureza da máxima disposicional seria mais (ou menos) prejudicial do que a fragilidade para a conformidade externa das ações à Lei. Entretanto, ainda que, algumas vezes, da máxima disposicional impura possam indiretamente derivar-se ações externamente (aparentemente) conformes à Lei, as máximas de ação diretamente derivadas dela são sempre más, pois abrem espaço para a necessidade de motivações não morais, que exprimem uma deterioração do comprometimento com a Lei Moral.



Além disto, a necessidade de que móveis não morais complementem os morais faz com que muitas vezes o agente não cumpra sequer exteriormente a Lei - naquelas vezes, a saber, em que as circunstâncias não permitem a atuação adequada dos móveis não morais.

Podemos agora rematar nossa interpretação voltando ao problema de como compreender as ações e máximas de ação malignas, perversas. Vimos acima que, em princípio, a única maneira de compreender tais máximas e ações consistiria no recurso a uma escolha conscientemente contrária à Lei Moral, efetuada pelo sujeito racional suprasensível, como sujeito liberto não apenas da determinação por impulsos naturais, mas liberto também da influência irracional (perturbadora) que tais impulsos poderiam exercer sobre sua consciência e reconhecimento de que a Lei Moral é a melhor razão para agir, a única que se justifica por si mesma. Em outras palavras, seria preciso recorrer à ideia de um sujeito racional que, tendo plena consciência de que a lei moral é a melhor razão para agir e tendo pleno poder de escolher a favor desta lei, mesmo assim escolhe opor-se a ela, numa escolha indiferente a toda exigência de justificação. Vimos que Kant é relutante a respeito desta noção, e por isso enveredamos pela tentativa de encontrar outra explicação possível.

A noção de derivação desatenta e negligente que procuramos apresentar talvez seja uma boa alternativa. Poderíamos repetir o procedimento proposto para a compreensão dos dois primeiros estágios da maldade em sentido amplo. Assim, como efeito de um processo de derivação marcado pelo acirramento da desatenção e negligência, da máxima disposicional impura acabaria sendo derivada uma máxima disposicional perversa, a qual, mais uma vez, não teria sido ativamente escolhida pelo sujeito racional, mas seria, sim, uma falha na qual ele cai ou incorre ao deixar-se progressivamente iludir quanto ao tipo de motivação e disposição que a Lei Moral exige. Como dito acima, esta falha equivaleria a um híbrido de falha volitiva e falha intelectual, e é isto que a expressão “deixar-se progressivamente iludir” pretende designar. Embora, como diz Kant “ainda possam sempre existir ações boas conforme a lei (legais), o modo de pensar é entretanto corrompido em sua raiz (no tocante à disposição moral)” (KANT 1793, B 23, Ak 6: 30). E esta corrupção do coração humano ocorre sem o exercício de qualquer capacidade de ativamente escolher opor-se à Lei, mas por obra apenas do “Mentiroso que habita em nós”, que aproveita a desatenção e a negligência para insidiosamente rematar sua obra de autoilusão.

RESUMO

*O objetivo do presente artigo é analisar o problema da fraqueza da vontade no quadro conceitual da filosofia moral kantiana. A fim de realçar uma característica da filosofia kantiana relevante para a posição conceitual do problema, servimo-nos do contraste com as análises de Donald Davidson sobre o fenômeno da fraqueza da vontade. Depois de evidenciado o modo como o fenômeno pode ser admitido no sistema kantiano, procuramos defender uma hipótese quanto ao modo como ele poderia ser aí explicado. Trata-se da hipótese de que uma derivação desatenta e negligente a partir da Lei Moral poderia tornar compreensíveis os diversos estágios da maldade em sentido amplo do ser humano, e de que os dois primeiros estágios poderiam ser considerados correspondentes ao fenômeno usualmente chamado de fraqueza da vontade.*

**Palavras-Chave** *Kant; Fraqueza da Vontade; Propensão ao Mal; Vontade e Arbítrio.*

ABSTRACT

*The aim of the present paper is to examine the problem of weakness of will in the conceptual framework of Kantian's moral philosophy. In order to present a feature of Kantian's philosophy which is important for the theoretical placing of the problem, I draw on the contrast with Davidson's analyses of the phenomenon of weakness of will. After indicating the way in which the phenomenon can be admitted in the Kantian system, I try to defend a hypothesis about the way in which it could be explained in this system. It is the hypothesis that an inattentive and negligent derivation from the Moral Law could explain the different stages of the evilness in broad sense of the human being, and that the two first stages can be judged correspondent to the phenomenon usually called weakness of will.*

**Key-Words** *Kant; Weakness of Will; Propensity to Evil; Will and Choice.*

## Referências Bibliográficas

- ALLISON, H. 1990. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press, 1990.
- BROADIE, A. e PYBUS, E. 1982. Kant and Weakness of Will. *Kant-Studien*, Vol. 73, p.406-412.
- CHARLTON, W. 1988. *Weakness of Will*. Oxford: Basil Blackwell, 1988.
- DAVIDSON, D. 1970. How is Weakness of the Will Possible? In: *Essays on Actions and Events*. New York: Oxford University Press, 1980.
- HILL Jr, T. 2008. Kant on Weakness of Will. In: Hoffmann, T. (Ed) *Weakness of Will from Plato to the Present*. Washington: The Catholic University of America Press.
- HORN, C. 2006. Willensfreiheit und Willensschwäche: Über zwei Beiträge des Christentums zur Philosophiegeschichte. In: Elm, R. (Org.) *Vernunft und Freiheit in der Kultur Europas*. Freiburg/München: Karl Alber Verlag.
- \_\_\_\_\_. 2014. How Close is Augustine's 'liberum arbitrium' to the Concept of 'to eph'hêmin'? In: Destrée, P., Salles, R. e Zingano, M. (Eds.) *What is up to us? Studies on Agency and Responsibility in Ancient Philosophy*. Sankt Augustin: Academia Verlag.
- KANT, I. *Werke in zwölf Bänden*. Hrsg. von Wilhelm Weischedel. Frankfurt: Suhrkamp.
- \_\_\_\_\_. 1785. *Fundamentação da Metafísica dos Costumes*. Tradução, Introdução e Notas de Guido Antônio de Almeida. São Paulo: Discurso Editorial - Barcarolla, 2009.
- \_\_\_\_\_. 1793. *A Religião nos Limites da Simples Razão*. Tradução de Artur Morão. Lisboa: Edições 70, sem data.
- \_\_\_\_\_. 1797. *Metafísica dos Costumes*. Tradução de Clélia Aparecida Martins. Petrópolis: Vozes, 2013.

Recebido em julho de 2017  
Aprovado em outubro de 2017