

Christ as Symbol in Kant's Religion

Kant's philosophy of religion has often been considered marginal in the context of his broad philosophical enterprise. Even when Kant's philosophy of religion is taken seriously, some of the arguments adopted by Kant in his writings on religion are sometimes treated as little more than oddities, and as having little (if any) relevance to the way Kant's legacy has been received and elaborated by post-Kantian thinkers. From this angle, focusing on Christ as symbol in Kantian religion, as I do in this chapter, might seem relevant merely in exegetical terms. However, I contend that, if properly reconstructed, the Kantian conception of Christ as symbol of sacrifice is indicative of the innovations and limits of the Kantian philosophical project in the context of religion, and it paves the way for similar (albeit more sophisticated) philosophical moves in Hegel's philosophy of religion.

Religious notions, for Kant, have to be understood as symbolic presentations (*Darstellungen*) to be used to apply the moral law to the world. As such, they can be used properly or improperly. For instance, Kant rejects the sacrificial symbol of the *Akedah* and Abraham's willingness to sacrifice his son Isaac because of a direct command by God.¹ In the case of the sacrifice of Christ, however, the question is more complex. Kant gives no shortage of warnings regarding improper ways of interpreting (Christ's) sacrifice, but in one place, he seems to suggest that we must ourselves *participate* in Christ's sacrifice. This place is in *Religion within the Bounds of Bare Reason*, Second Piece, Section One, Subsection C, when Kant addresses the third of the three ethical difficulties that arise out of any belief in divine assistance, or grace. Kant's conception of grace,

traditionally underplayed in favor of a reading that supports Kant's emphasis on salvation by works, has recently been the subject of a set of studies that has debated and clarified this issue, showing the relevance of grace in the context of Kant's view on religion.² Although Kant's conception of grace inevitably enters into the analysis carried on here, the primary goal of this chapter is not to offer a contribution to the debate on the notion of grace but rather to focus on Kant's conception of Christ as symbol, to clarify its relevance in the context of Kant's philosophy, and to point toward significant philosophical issues that are implied in this conception and that were inherited and transformed in the idealist tradition, particularly in Hegel's philosophy of religion.

First, I address Kant's notion of conversion, which provides the framework within which Kant advances his idea of the "new man" being a symbolic Christ on behalf of the "old man," or "former self." In this context, Kant's conception of proper sacrifice emerges in connection with a representation of Christ conceived as the model for that sacrifice. Then, I analyze the two ways in which Christ should be considered as a model for sacrifice: acceptance of vicarious punishment and openness to forgiveness. I also underline how this openness can be related to the Kantian idea of truthfulness. Finally, in the conclusion, I outline both the limits and the relevance of Kant's conception of Christ as symbol and its implications. In the next chapter, I will argue that, among the post-Kantians, Hegel built on Kant's account in such a way that allows a possible solution to the unsolved aspects of Kant's conception.

Conversion and Christ as Exemplar

In *Religion within the Bounds of Bare Reason*, Kant argues that the human being is *originally* good but *radically* evil. In order to become a morally good subject—that is, in Kant's own words, a subject "who, when he cognizes something as a duty, requires no other incentive beyond this presentation of duty itself"³—the human being must transform her underlying disposition from evil to good. To refer to this transformation, Kant also uses the terms "revolution" (*Revolution*),⁴ "conversion" (*Übergang*),⁵ "change of mentality" (*Sinnesiinderung*),⁶ and "change of heart" (*Änderung des Herzens*).⁷ The latter two expressions are probably those that best capture Kant's conception of conversion as a "change of disposition" (*Gesinnung*)—a conception that is consistent with the original meaning of the Greek expression μετανοεῖτε

used in the Gospels (Mark 1:15), which properly means “change your mind” or “change your way of thinking.” The idea implied here is that the “revolution” in one’s way of thinking, and the consequent adoption of a different perspective and underlying disposition, represents a turn to a new principle of *action*.

Religion within the Bounds of Bare Reason is not the only work in which Kant addresses the notion of conversion. The notion is also discussed in *Anthropology from a Pragmatic Point of View*. In this work, sometimes underestimated, which originated from Kant’s lectures on anthropology, Kant is mainly concerned with knowledge from a pragmatic perspective—that is, knowledge of humans as “freely acting beings.” In the *Anthropology*, Kant refers to conversion as “revolution” and “rebirth” and emphasizes its “instantaneousness”: the “firmness and persistence in principles,” he claims, cannot be affected “gradually,” but only “by an explosion [*Explosion*] which happens one time.”⁸

The idea of a moral “revolution” that happens in a specific moment in time, and yet radically changes the subject, can seem surprising and perhaps not sufficiently grounded. Kant, however, has very good reasons to think this way. As explained by Frierson, “a change in moral status cannot be gradual but must be an absolute transformation from good to evil.”⁹ In fact evil is, in Kant’s view, the subordination of the moral law to the inclinations of the senses. A less frequent subordination of the moral law, or its subordination to very strong temptations, does not make a moral subject less evil. “Only a complete shift, such that the moral law assumes absolute priority,” Frierson continues, “constitutes genuine moral improvement on Kant’s account.” Frierson, however, believes that the moral revolution described in the *Anthropology* “is not identical to the moral revolution described in the *Religion*.”¹⁰ As evidence for this claim, he stresses the temporal and conscious¹¹ nature of the *Anthropology*’s revolution, as opposed to the atemporal and intelligible¹² nature of the *Religion*’s revolution.¹³ In emphasizing the difference between the two revolutions, however, Frierson fails, in my view, to take into consideration Kant’s *perspectival* attitude, which is here applied on both a logical and a theological level. On the logical level, the “change of heart”—a noumenal event that happens outside of the temporal chain of phenomenal events—determines a corresponding event in time, which initiates a gradual process of moral development; in other words, the change of heart is the condition of the possibility of the “gradual moral progress of man in time.”¹⁴ On the theological level, Kant is indeed talking about

the same conversion, but from two different points of view: the point of view of God in the *Religion* and the point of view of the moral subject in the *Anthropology*. As Hare puts it, “What God sees (by intellectual intuition) is revolution; what we experience is reform.”¹⁵ I will come back to the implications of Kant’s perspectival view of conversion at the end of the next section. For now, it is sufficient to underline here two aspects of Kant’s account of conversion. First, conversion is initiated by a single, “revolutionary” act, but this act then requires an ongoing lifelong commitment: “Virtue in this sense is acquired little by little, and means to some a long habituation (in observing the law), whereby the human being, through gradual reforms of his conduct and stabilization of his maxims, has passed over from the propensity to vice to an opposite propensity.”¹⁶ This is an important remark, which is meant to respond to the concern about whether conversion would not create “a disturbing discontinuity in someone’s empirical character”¹⁷—a concern that becomes more evident through the analysis of the former self / new self dynamics that Kant carries on in that context (more on this soon). Second, as an implication of the previous point, precisely because the moral conversion is a lifelong commitment, we can only *hope*¹⁸ to have changed our underlying disposition from evil to good, but we can never be completely sure that we have actually changed it.¹⁹

Thus, in Kant’s account, the innate propensity to evil can only be overcome through a “revolution” in one’s way of thinking (*Revolution für die Denkungsart*): this is a single event, which is, however, the necessary condition for a (gradual) reform of character (*für die Sinnesart*). The revolution is, therefore, the opening moment of a new life characterized by an ongoing progress toward goodness. As such, it is not merely an intellectual happening, as it involves a practical and unending process of reformation. There are substantial claims about anthropology and about the nature of morality operating here. Anthropology is conceived by Kant as the systematic doctrine containing our knowledge of man, and in contrast with the general optimism of his age, Kant puts at the center of his anthropology the idea of the innate propensity to evil. At the same time, the nature of morality is defined for Kant by the categorical imperative, which is the reason why—as I anticipated earlier—in order to become a morally good subject, the human being must transform her disposition (from evil to good). These issues also bear on discussions of Kant’s notion of virtue, given that virtue is acquired little by little. However, in the context of such discussions, the role of the “revolution” (which then leads

to conversion) is often downplayed, in the name of a more “secularized” conception of Kant’s moral philosophy.

Also, there is undoubtedly an issue of coherence. First, there clearly is some tension between the “revolutionary” aspect and the “reform of character.” Second, to some, Kant’s motif of a “rebirth,” that is, the description of moral conversion in terms of the birth of a “new man,” can appear as a mere “pictorial filler for a conceptual lacuna”²⁰—as if Kant has to resort to religious notions because of a lack of convincing arguments. However, as others have remarked,²¹ Kant’s move is far from a “retreat into the irrational.” On the contrary, Kant’s argument is very rigorous. If humans are radically evil, that is, if they have an underlying disposition to evil, a change toward the good would not be possible without a *supersensible* help—which, in theological language, is called “grace.”²² The charge of irrationality might be acceptable only if Kant had considered religious notions and narratives as idle and superfluous, but this is not the case: on the contrary, Kant maintains that the content of revealed religion is useful “in making up the theoretical deficiency which our pure rational belief admits it has (in the questions, for example, of the origin of evil, *the conversion from evil to good*, man’s assurance that he has become good, etc.) and helps—more or less, depending on the times and the person concerned—to satisfy a rational need.”²³ Clearly, Kant is very well aware that his move—that is, the attribution of the possibility of conversion to some sort of supersensible intervention—generates further questions: “how is rebirth (resulting from a conversion by which one becomes an other, new man) possible by God’s direct influence, and what must man do to bring it about?”²⁴

The answer to the first half of the question resides in the presence in us of a moral archetype, symbolized by Christ; the answer to the second half resides in the possibility of “activating” the archetype so that the “old man” (the former self) dies and the new self can come to birth. Therefore, the remaining part of this section will be devoted to an analysis of Christ, and then to the former self / new self transition process. In the context of this analysis, the notion of Christ as symbol and its relevance in the context of Kantian religion will start to emerge.

An analysis of Christ as an archetype should appropriately start from the pure idea of moral perfection, an idea that—as Kant maintains in *The Metaphysics of Morals*—“reason frames a priori.”²⁵ However, the pure idea of moral perfection needs to be applied to the world and, as Kant remarks in the *Critique of Practical Reason*, “nothing corresponding to it

can be found in any sensible intuition.”²⁶ We need transitional forms to apply the pure principle of moral perfection to experience;²⁷ in other words, the pure idea of moral perfection is to be *symbolized* in an *archetype* (*Urbild*). The archetype is a discursive, image-dependent understanding (*intellectus ectypus*)²⁸ and “an object of intuition, insofar as it is the ground of imitation.”²⁹ The archetype should not be confused with a related idea, that of the “prototype” (*Vorbild*): the archetype is the original notion that makes something what it is, whereas the prototype is the first model that is adopted from that notion.³⁰

In the *Critique of Pure Reason*, Kant identifies the prototype of pure moral disposition in the wise man of the Stoics (the “sage”), claiming that “we have in us no other standard for our actions than the conduct of this divine human being, with which we can compare ourselves, judging ourselves and thereby improving ourselves, even though we can never reach the standard.”³¹ Later on, however, in the *Religion*, Kant rejects the Stoic sage as the prototype: in the new picture dominated by the idea of radical evil, the wise man representing the ancient ideals of *ataraxia* (absence of worries) and *aponia* (absence of passions) is clearly too abstract and does not fit with the corruption of human nature.³² Another prototype is to be adopted: Christ.

It should be stressed that in this identification of Christ as the prototype, Kant’s attitude remains strictly philosophical. Christ, being the prototype of pure moral disposition (that is, the first symbolic model that fully represents this notion), confirms the reality of the archetype (*Urbild*). The archetype is always already present in a human being, although its presence cannot be explained rationally: it is “a God-shaped hole in the heart of human reason.”³³ The presence of the archetype is, in other words, independent of faith in the historical Jesus (the prototype). Previous scholarship has already remarked that Kant was unconcerned about the historicity or “veracity” of the events of Christ’s life as they are narrated in the Gospels.³⁴ The idea of moral perfection is an idea of pure reason; being its symbolic representation, the archetype makes possible the application of that moral perfection in the world, but as such, it is not concerned with phenomenal reality. However, this remark, which is correct in itself, might easily lead one to conclude that Kant’s choice of Christ is arbitrary or *merely* culturally driven and that therefore Christ can in principle be substituted by some other figure. This is not the case:³⁵ there are at least two fundamental reasons why Kant identifies the prototype with Christ. First, Kant maintains that only God (whose idea is for Kant indissolubly

connected to another idea, that of moral perfection, which reason frames a priori) should properly be considered the archetype of the good:³⁶ being divine and human at the same time—God incarnate—Christ is the perfect symbol of pure moral perfection.³⁷ Second, as an implication of the previous point, Christ is thought of as “fraught with the very same needs and thus also the same sufferings, with the very same natural inclinations, and thus also the same kind of temptations to transgression as we are.”³⁸ The emphasis that Kant puts on this conception of Christ can perhaps be considered to be Kant's answer to a potential objection, that is, that Christ is an unreachable model because, being “superhuman,” he was immune from natural inclinations and temptations; but if Christ was *not* immune from human limitations, then the objection no longer stands, and his role as model should be taken seriously. The remark by Kant surely has the nature of a counter-objection; at the same time, however, its philosophical value is well beyond that of a specific counter-objection and touches the very strong connection that links, for Kant, Christianity and morality. In fact, the idea of a god becoming human and, as such, being exposed to the limitations and constraints connected with a finite nature, is clearly grasped by Kant as the novelty of Christianity, a novelty that allows Christ to be regarded as the *bodily exhibition*, or *incarnation*, of the good.³⁹

Kant's attitude toward the use of a moral example might be regarded as ambiguous. Although in the *Critique of Practical Reason* he briefly emphasizes the positive role of examples, especially in the context of moral education,⁴⁰ in the *Groundwork* he expresses strong doubts about the fruitfulness of empirical examples for the establishment or adoption of moral principles,⁴¹ claiming, “Nor could one give worse advice to morality than by wanting to derive it from examples,”⁴² and “[i]mitation has no place at all in matters of morality, and examples serve only for encouragement.”⁴³ Why then does Kant speak so highly of the role of Christ as a symbol in the *Religion*? Is Kant merely inconsistent? Not at all. Kant always maintains a negative attitude toward the use of empirical examples, precisely because they are *empirical*. As a symbol, Christ is different:⁴⁴ he should not be regarded as an example on which we are to model this or that particular moral action, but as an embodied *ideal* (*das Letztere*). The ideal component of this symbol resides in its divine nature (Christ is God), while at the same time Christ's human nature makes it accessible. Moreover, Christ evokes the adoption of a different “way of thinking,” whose possibility is already *in us* (as Kant claims, “in order to regard something as an archetype, we must first have an idea according

to which we can cognize the archetype⁴⁵), but which requires a revolutionary act of will⁴⁶ (which initiates the conversion) in order to be, as it were, “activated.” Since this act cannot be properly conceptualized or schematized (we are not in the realm of theoretical reason), it can only be *symbolized*, and there is, according to Kant, only *one* symbol suitable for the purpose: Christ. As Kant explains: “If we have an idea of something, e.g., of the highest morality, and now an object of intuition is given, someone is represented to us as *being congruent with this idea*, then we can say: this is the archetype, follow it!”⁴⁷ It is important to understand that, in Kant’s view, the role of Christ is not meant to be that of a *passive example*; in the *Religion*, he claims that the archetype “can give us power [*Kraft*]” to “elevate ourselves to this ideal of moral perfection.”⁴⁸ Kant is still arguing within a philosophical framework here, and he is careful not to trespass into a theological realm; therefore, he acknowledges that we cannot say whether the archetype in us has supernatural origins; he does claim, however, that “we are not the idea’s originators” and that this idea “has taken its place in the human being without our comprehending how human nature could have been so much as receptive to it.”⁴⁹

Now, we need to clarify what this different “way of thinking” is, and why Christ is the only symbol able to help us in the process of conversion. The answer to these questions resides precisely in the specific sacrificial dynamics that Christ symbolizes—that is, Kant’s emphasis on Christ’s willingness “to take upon himself all sufferings.”⁵⁰

Before proceeding to the unpacking of the way of thinking of the converted man, we should first clarify the context in which Kant regards this sacrifice as happening: the death of the “old man” (the former self) and the birth of the moral “new man.” Kant provides such a context when he addresses the third of the three “Difficulties Opposing the Reality of This Idea,” that is, difficulties threatening the reality of the archetype.⁵¹ The third difficulty wonders how a just God can overlook pre-conversion evil. Given that conversion, as we have seen, marks a turning point in the moral status of a human being, Kant refers to the pre-conversion subject as “the old man” and to the post-conversion subject as “the new man.” The argument that follows from here is, in itself, fairly simple. The sins of the former self have not been punished, but justice *requires* such a punishment, because the good deeds of the “new man” are not enough to compensate for the evil done by the former self. Who is going to pay for those sins? The former self is, morally speaking, “dead”; at the same time, the “new self” is innocent and deserves no punishment. Therefore, the

punishment cannot occur *before* the conversion (in that case, there would be no need for conversion in the first place), but it cannot occur *after* the conversion either (precisely because the post-conversion subject is a “new man,” and holding him responsible for the sins of the former self would imply that no radical change has occurred). Therefore, Kant maintains that the punishment must happen *during* the conversion—or, better, *in* the act of conversion itself. Thus, the new man indeed takes upon himself the punishment, but only as a spontaneous (that is, not externally required, but freely accepted) and yet necessary (because without it no conversion is possible) sacrifice, *modeled* on Christ's sacrifice. Kant summarizes these dynamics in a passage that is worth quoting in full:

The exit from the corrupted into the good attitude (as “the dying of the old human being,” “crucifying of the flesh”) is in itself already a sacrifice and an entrance upon a long series of life's ills that the new human being takes upon himself in the attitude of the Son of God—in other words, merely for the sake of the good—but that yet were properly deserved by a different human being, namely the old one, as punishment (for, the old one is morally a different human being).⁵²

If the argument in itself is simple, understanding its various implications is difficult. First of all, let us clear the air about what Kant does *not* mean with this argument. First, this is not a classic theory of salvation for vicarious or substitutionary atonement—that is, a theory according to which the historical Christ bore our sins, was punished, and died as a substitute for all humankind.⁵³ Kant does not argue that the crucifixion of Christ *in itself* has a salvific power that redeems us. Second, Kant is not merely suggesting that the punishment is a sacrificial offering, in the form of suffering and troubles, that satisfies an angry God who can now look at our moral debt as discharged.⁵⁴ Such an account would suggest that there is a way to please God other than by pure moral conduct—not with prayers or praises, but with sufferings and struggles; but Kant always strongly rejects any way to influence God's benevolence. It is true that Kant argues that, because of this sacrifice, humans “can hope to appear before their judge as justified”;⁵⁵ but first, this is only a hope, not a knowledge; and second (and even more importantly), the justification evoked here is something more than a mere discharge. I will come back to the issue of justification in the next section.

Thus, in Kant's account, Christ is neither a vicarious victim nor a model for the discharge of our moral debt through suffering and struggles. In which sense then is Christ an embodied ideal? First, Christ symbolizes moral purity, and the new man is indeed pure. Second, Christ is the *redeemer* (*Erlöser*) and the *advocate* (*Sachverwalter*),⁵⁶ and the new man is called to play the same role in redeeming the former self and in advancing hope for salvation. Kant's "earlier denial of transmissible liability" is not "revised"⁵⁷ with the introduction of the idea of punishment as sacrifice: here Kant is proposing a brand new conception that, while it does not admit traditional transmissible liability, introduces important innovative elements.

To grasp the novelty of Kant's conception, consider two issues that, at first, might even appear to be problematic aspects of Kant's account: how the moral subject can really take upon herself the "long series of life's ills," if this sacrifice is supposed to happen *during* the conversion (and not *after* it); and how Kant can appeal to the former self / new self distinction without generating a lack of personal identity in the moral subject.

Regarding the former issue, we have already noted that, although the "change of heart" is supposed to be an instantaneous act of the will, Kant elsewhere stresses that a pure moral status is the objective of a lifelong, gradual improvement. This apparent contradiction can be explained only by appealing to Kant's perspectivism, which clearly appears in the context of the *Religion*, when Kant discusses the first difficulty. Two perspectives need to be taken into consideration: the temporal and finite point of view of the human subject, and the timeless and infinite point of view of God. With conversion, and the birth of the new man, our moral status appears as changed to God, who grasps the entirety of our moral life and can therefore see "the stable disposition of the heart";⁵⁸ but we live in time, and therefore we experience a gradual improvement, which never comes to an end. As Hare puts it: "What God sees (by intellectual intuition) is revolution; what we experience is reform."⁵⁹ Considered as such, sacrifice is "the culmination of the conversion itself":⁶⁰ the new man's willingness to sacrifice himself is really the object of the imitation of Christ, or, even better, the fulfillment of the "Christic disposition."⁶¹ However, we must remember that the archetype is always already (mysteriously) in us, and it is the idea that gives us the power to initiate the process of conversion; thus, what is required of us is, as Kant explicitly states in the *Conflict of the Faculties*, "to make room for it [*ihm Raum zu verschaffen*]."⁶²

Now, consider the problem of the lack of personal identity in the moral subject.⁶³ Although *physically* the same, the post-conversion subject

is *morally* “a new man.” This is just an implication of Kant’s conception of conversion as a revolutionary act of the will: if the change of heart is as radical as to imply a complete change of one’s way of thinking, a split in moral identity is unavoidable. However, Kant’s conception includes the solution to this danger. The death of the former self coincides with the birth of the new self, and in turn, the willingness of the new self to pay for the old self’s sins bridges the gap in moral identity. It would be misleading to regard this lack of moral identity as “merely figurative,” as suggested by Frierson.⁶⁴ The gap is, on the contrary, indubitably there; and yet, a sacrifice modeled on Christ is able to bridge it.⁶⁵ To regard the lack of moral identity as merely figurative means to downplay the status of Christ as a symbol of sacrifice; conversely, such status is essential precisely in order to avoid a risk of loss of moral identity that would otherwise present itself.⁶⁶

To summarize: a revolutionary act of the will causes the death of the old self, and the transformation of the underlying disposition from evil to good creates the condition for the birth of a new self. The process of conversion, however, is not completed without the willingness of the new self to take upon itself the evil of the earlier self, and this willingness can derive only from the archetype of pure moral perfection, which is always already in us, but which needs a symbol in order to be activated through imitation; this symbol is Christ. The process of conversion, therefore, culminates in the adoption of a new way of thinking and acting. Subjectively experienced, post-conversion life is therefore a continuous sacrifice, not only (and not primarily) in the sense of a quasi-legal retribution for pre-conversion evil. From this angle, the definition of the suffering of Christ (on which the sacrifice of the new man is meant to be based) as “no more than a symbol for the remorse and misery of the repentant sinner consequent upon his former misdeeds” is reductive at best.⁶⁷ The meaning of Christ’s sacrifice goes far beyond that for Kant. Therefore, the next section will be entirely devoted to a clarification of the meaning of Christ’s sacrifice, and to a discussion of what it means in Kant’s view for one to participate in (to become one with) that sacrifice.

Vicarious Punishment and Openness to Forgiveness

Christ, who is the symbol of the sacrifice that the new self performs on behalf of the former self, is not merely an appended metaphor but plays

an essential role in Kant's account of conversion. In fact, this sacrifice has to be performed if justification (forgiveness of pre-conversion evil) is to be rationally justifiable.⁶⁸ The sacrifice modeled on Christ is the pivotal point of the entire process of conversion, as Kant clearly states.⁶⁹ But recall that the relevance of Christ as symbol consists in the capacity of this symbol to "activate" the archetype and, consequently, to make the ideal of pure moral perfection applicable to the world. Therefore, the question is: what is the *meaning* of this sacrifice, and what consequences and implications does it have for the life and conduct of the post-conversion moral subject?

There are two meanings of this sacrifice, coexisting in Kant's account; although it is possible to regard them as mutually compatible and even complementary, there is a tension between them, with Kant emphasizing now one, now the other, thus retaining some degree of unresolved ambiguity.

The first meaning of this sacrifice is *vicarious punishment*. The "third difficulty" properly wonders how a just God can overlook pre-conversion evil; from this angle, the answer is simply that God *cannot* overlook that evil, and therefore, someone has to be punished for that evil: the sacrifice of the "new man" makes atonement possible. This is the way in which Michalson interprets the new man's Christic sacrifice: it is "the locus of the punishment,"⁷⁰ with the "long series of life's ills" referring to the "ongoing process of moral struggle,"⁷¹ an "atoning sacrifice [. . .] mandated by the moral economy of Kant's universe, a universe requiring symmetry and correct proportion."⁷² According to this interpretative angle, Christ is the symbol of sacrifice because the "punishment for the old self is thus analogous to Christ's death on the cross."⁷³ There is no doubt that this is *one* aspect of Kant's account of the sacrifice of Christ as a model; to Michalson's account, one might add that the expiatory punishment extends across the whole life of the converted man,⁷⁴ but the punitive and retributive aspect of this sacrifice is clearly there. However, and this is the point, this is only *one* aspect of Christ's sacrifice. As already noted by Wood, Kant "does not in fact restrict himself to the (essentially irrelevant) question of punishment for man's guilt. [. . .] Alongside Kant's statements that the 'new man' (the disposition to progress) takes upon himself the sufferings and punishments due to the former self (the propensity to evil) and bears his guilt vicariously, we find also the statement that this disposition 'takes the place of the deed in its perfection (*vertritt . . . die Stelle der That in ihrer Vollendung*),' and that the 'new man' as advocate makes it possible for men to hope to appear before their judge as justified (*gerechtfertigt*)."⁷⁵ Wood does not, however, unpack the implications of the Kantian conception of

sacrifice as a “disposition”—which effectively represents the second (and most important) meaning of sacrifice.

In order to understand sacrifice as a disposition, we first have to underline an aspect of the Christic sacrifice that was not properly addressed during our previous analysis—*forgiveness*.⁷⁶ It has already been said that the sacrifice performed by the new man is a spontaneous and free act of the will (because the new man, insofar as he is innocent, is under no obligation to suffer for the evil made by the former self);⁷⁷ but the new self, as already stressed, is not only the “redeemer”; he is also the “advocate” for the former self. As such, the new (post-conversion) self *forgives* the earlier (pre-conversion) self; without this act of forgiveness, no justification can be possible, and the process of conversion cannot be completed. Christ is the symbol for this sacrifice, and this should not come as a surprise: when Kant refers to the redeemer and the advocate, he might well have in mind Christ dying on the cross saying, “Father, forgive them, for they do not know what they are doing” (Luke 23:34). As stressed by Palmquist, conversion for Kant is not a mere intellectual acknowledgment of radical evil; it is a “radical conversion of one’s disposition”⁷⁸ modeled on Christ’s incarnation. Consistent with the kenotic tradition, Christ’s first sacrifice is the incarnation itself: God renouncing his absoluteness to become human, accepting all the limitations and constraints that come attached to our finite condition.⁷⁹ Analogously, the self (the new, post-conversion person) changes her underlying disposition from a dedication to her own self-interest to “a recognition that she is not in herself any more valuable than anyone else.”⁸⁰ Only if we follow the archetype in adopting this attitude can we indeed fully participate in Christ’s sacrifice and become one with it.

In this sense, the sacrifice of the former self is not only punishment; it is primarily a lifelong attitude, and because of this (as I have already stressed), we can never be completely *sure* to appear as justified before God (the condition that, in theological language, is called “the state of grace”), but we can only *hope* to be so.⁸¹ This is what Kant calls “practical faith in this Son of God”:⁸² an attitude that is, at the same time, a lack of absolute certainty and *truthfulness*. Only the person who adopts this attitude is, Kant claims, “entitled to regard himself as the human being who is an object not unworthy of divine pleasure.”⁸³ This is important, because it shows how Kant’s account of religion in general, and of Christ in particular, is not severed from Kant’s epistemological preoccupations.⁸⁴

To clarify the relevance of the notion of truthfulness in relation to the connection between Kant’s religious *and* epistemological concerns,

it is useful to refer to another of Kant's works, the 1791 essay *On the Miscarriage of All Philosophical Trials in Theodicy*. Here Kant addresses the book of Job and, commenting on the "debate" that opposes Job and his not-so-friendly "friends," he shows a clear disregard for the "subtle or hypersubtle reasonings of the two sides," while—he claims—"the spirit in which they carry them out merits all the more attention." Job, in fact, "speaks as he thinks," whereas—Kant continues—his friends "speak as if they were being secretly listened to by the mighty one."⁸⁵ A few paragraphs later, Kant draws the important distinction between truth (*Wahrheit*) and truthfulness (*Wahrhaftigkeit*):

One cannot always stand by the truth of what one says to oneself or to another (for one can be mistaken); however, one can and must stand by the truthfulness of one's declaration or confession, because one has immediate consciousness of this. [. . .] I can indeed err in the judgment *in which I believe* to be right, for this belongs to the understanding which alone judges objectively (rightly or wrongly); but in the judgment *whether I in fact believe* to be right (or merely pretend it) I absolutely cannot be mistaken, for this judgment—or rather this proposition—merely says that I judge the object in such-and-such a way.⁸⁶

There is no doubt that the notion of truthfulness plays an important role in the context of Kant's practical philosophy. Kant defines truthfulness as "the greatest virtue in the world"⁸⁷ and as "the supreme *principium formale* of morality."⁸⁸ Kant's treatment of truthfulness has a very large set of ethical implications, which cannot be taken into consideration here. What really matters for our purposes is the distinction between truth and truthfulness. As Zupančič puts it, truth is about "the relation between our statements and the 'facts' to which they refer," whereas truthfulness is about "the agreement between our statements and our beliefs"—in other words, it is about the *will* to tell the truth.⁸⁹ In the 1797 essay *On a Supposed Right to Lie Because of Philanthropic Concerns*, Kant points out that the expression "to have a right to the truth" is "meaningless," and that "one must instead say one has a right to his own truthfulness (*veracitas*)."⁹⁰ Now, there are two ways of reading such a position with respect to the story of Job. On one hand, one can read it as meaning that Job's view is honest but not necessarily correct: as one commentator puts it, "Truthfulness, while better

than mendacity, is no guarantee of truth.”⁹¹ On the other hand, one can emphasize the inappropriateness (on an epistemological level) and arrogance (on a practical level) of a position that claims to have *direct* access to truth. Kant's conclusive statement about the story of Job mentioned earlier seems to go in the latter rather than in the former direction. The pretension of having a direct access to truth implies a naïve standpoint of self-certainty that, even when considered from a strictly epistemological point of view, does not serve the cause of truth and knowledge well; from this angle, it can be argued, as Hanna does, that “practical truthfulness is the enabling presupposition of all propositional truth”⁹²: “It is rather that we must be truthful, that is, non-cognitively aimed at truth (so this is truth-conation, or ‘the love of truth’), because this is a necessary condition of all the other intellectual and moral virtues that together constitute an ideally successful rational human life.”⁹³ On a practical level, the danger of confusing one's account of truth with *truth itself* (thus demanding a right to truth rather than to truthfulness) is also very dangerous, because hypocrisy toward others can be grounded on an inner self-deception.⁹⁴ Untruthfulness, the sin of which the so-called “friends” of Job are so evidently guilty in their arrogance to “speak the truth” and to see the world “from the point of view of God,” is nothing other than the “radical evil” of *Religion within the Bounds of Bare Reason*.⁹⁵

It is precisely at this point, however, that things take a surprising turn. For, at the end of the book of Job, God shows up, as God's voice is heard coming from a “cloud” showing God's power, proclaiming God's absolute freedom over God's creation, and eventually condemning Job's friends for their arrogance while, at the same time, commending Job for his truthful words. As one commentator has noted, this is one of the few places in Kant's works where God speaks;⁹⁶ and it is striking that Kant introduces God here, precisely in the context of one of his most powerful attacks against a pretension of embracing a God's-eye view. This performative contradiction is really meaningful in order to understand Kant's (philosophical) predicament. On one hand, Job is presented as the rightful man, one who has adopted the good as the underlying principle of his disposition and truthfulness as his ongoing conduct of action, to the point of defending his right to truthfulness even in the face of his pious “friends” who pretend to speak “from the point of view of God.” On the other hand, Kant *needs God* to intervene as a guarantor and to clarify that no one is entitled to offer an authentic interpretation of the world, *except God Godself*.⁹⁷ This state of affairs mirrors the role that Kant

attributes to God in the context of his analysis of divine grace. Recall that Kant says that the act of will can indeed activate the archetype and thus initiate the process of conversion, but that the sacrifice of which Christ is the symbol, which is expressed in a lifelong commitment to virtue and truthfulness, would not grant us redemption, unless God, who is able (through his intellectual intuition) to grasp the entirety of our moral efforts, makes that conversion real. In other words, “a God’s-eye view of our state is needed.”⁹⁸ Here one can recall Kant’s repeated insistence that the moral subject must listen to moral commands as if they were spoken by the voice of God.⁹⁹ The new man’s sacrifice, which signals the “entrance upon a long series of life’s ills that the new human being takes upon himself in the attitude of the Son of God,” is the path to moral perfection; but the very possibility of this moral perfection in the Kantian religion depends, both epistemologically and practically, on a conception of a God’s-eye view, unattainable by humans, that is yet required.

The Limits and Relevance of Kant’s Conception of Christ

We are now in a position to appreciate the philosophical meaning of the Kantian conception of Christ as symbol in the context of Kant’s theory of conversion. Therefore, we can now briefly examine the strengths and limits of this move.

Although occasionally referred to as “opaque” by some interpreters,¹⁰⁰ Kant’s solution to the question of conversion presents several aspects of philosophical innovation, organized in a coherent way. First, Kant relies on the idea of an internal archetype of moral perfection that needs to be “activated” by adopting a unique embodied ideal on which we model a sacrificial attitude: Christ thus becomes the symbol of a sacrifice that is conceived in an unorthodox manner, not only and not primarily as a punitive retribution but also, and more importantly, as the will to forgive. With this account, Kant succeeds in avoiding any irrationalist drift in conceiving the process of conversion (via the notion of the internal archetype), while at the same time he takes into consideration the religious element represented by Christ, treated not as an accidental metaphor but as a symbol that has a conceptual connection with the internal archetype. As Kant puts it in the *Conflict of the Faculties*: “Even the Bible seems to have nothing else in view: it seems to refer, not to supernatural experiences and fantastic feelings which should take reason’s place in bringing

about this revolution, but to the spirit of Christ, which he manifested in teachings and examples so that we might make it our own—or rather, since it is already present in us by our moral predisposition, so that we might simply make room for it [*ihm Raum zu verschaffen*].”¹⁰¹ Consider that here Kant adopts, once again, a kenotic language: the archetype is already in us, and what we should do when we encounter the triggering symbol of Christ is to “withdraw” and make room for it.

This is for Kant the path to conversion and to moral perfection. However, moral perfection is still an ethical notion; is this also the path to (religious) *redemption*? In Christian theology, in fact, redemption, as the deliverance from sin, is an essential element of salvation. Is the activation of the archetype through the encounter with Christ enough to deliver us from sin and *redeem* us?

Kantian scholarship is split on this point. Firestone and Jacobs argue that “[r]edemption is not brought about by a historical Christ-figure”; therefore, in their view, there is room for redemption in Kantian religion, but only for a redemption considered as “self-redemption worked out in our mirroring of the symbol that is Christ.”¹⁰² Conversely, Reardon argues that Kant’s moral theology “cannot admit the idea of redemption,” pointing out that “the word itself occurs but once in his book [the *Religion*], and then only in a context which depreciates its meaning.”¹⁰³ In fact, Kant uses the word “redemption” (*Erlösung*) only once in the *Religion*, in the context of a criticism toward those who pretend to know the precise mechanism whereby God brings salvation about.¹⁰⁴

There is a sense in which both positions are accurate, insofar as each captures different aspects of Kant’s theory. On one hand, Firestone and Jacobs are correct in their claim that the imitation of Christ’s attitude is *as close as it can be* to the idea of (self-)redemption: the converted person somehow redeems herself through the adoption of a sacrificial attitude. On the other hand, Reardon is right insofar as there is no room for *proper* (religious) redemption in Kantian religion. In fact, redemption implies, properly speaking, forgiveness for past sins: a forgiveness that is not meant to be merely a legal absolution. Kant’s God is a judge who provides legal absolution to the moral subject, so that the subject can appear justified in front of God (thanks to the subject’s sacrifice), but who does not grant proper forgiveness. This can be considered the major weakness of Kant’s theory of conversion. It has already been noted that Kant’s theory requires a God’s-eye view: sacrifice makes sense because there is a God who recognizes this sacrifice, and conversion is possible because God, grasping

our struggle toward moral perfection through God's intellectual intuition, makes us justified.¹⁰⁵ This God is, at the same time, the "supreme legislator" and the judge: as such, God is commanding "not mercifully," Kant says, and "not forbearingly (indulgently) toward the weakness of human beings."¹⁰⁶ As Wood correctly points out, Kant remarks in several places that "a 'pardoning' or 'beneficent' judge, one who judges leniently and not according to the law, is a contradiction in terms."¹⁰⁷ In other words, in order to work properly, Kant's theory needs a God's-eye view, that is, an external, judging standpoint from which the moral subject can be either condemned or absolved (and thus justified)—but not forgiven.

There is therefore, in Kant's theory of conversion, an emphasis on the need of the subject's participation in Christ's sacrifice, so that—to use Hare's words—"[w]hen God looks at us, he sees his Son, because he is imputing to us his Son's righteousness,"¹⁰⁸ but Christ is still treated as an external symbol, and this symbol's conceptual relation to the internal archetype is not fully pursued by Kant, thus resulting in the underdevelopment of an aspect that remains only suggested: a full participation in Christ's sacrifice, to the point of becoming one with it. The post-Kantians, and especially Hegel, subsequently provided alternative theories of (moral and/or religious) conversion, thus addressing various issues that appeared in Kant's philosophy and that have been briefly sketched here, such as the question of radical evil, the possibility of a kenotic sacrifice, the meaning and use of religious symbol, and the question of moral identity; but it is important to recall that it is upon Kant's theory of conversion that these theories have been built.