

## CHAPTER V

## Kant

KANT's theory of ethics differs from Spinoza's and Hume's far more radically than these differ from each other. The most fundamental point of divergence is the following. For Spinoza and Hume the notions of good and evil are primary, those of right and wrong are derived from them, whilst that of duty or obligation is barely mentioned. A right action or intention is simply one that leads or is likely to lead to a good result. For Kant the notion of duty or obligation and the notions of right and wrong are fundamental. A good man is one who habitually acts rightly, and a right action is one that is done from a sense of duty. There is a second absolutely fundamental difference between Kant and Hume, at any rate, which may be mentioned at once. Ethics for Hume is concerned simply with mankind. It deals with the purely contingent fact that men have a disposition to feel emotions of approval and disapproval, and the equally contingent fact that in men this disposition is excited by contemplating the happiness or misery of human beings. Kant, on the other hand, holds that the fundamental laws of morality are the same for every rational being, whether man, angel, or God, since the ultimate criterion of rightness is deducible from the concept of a rational being as such. The relation of Kant to Spinoza on this point cannot be stated briefly; it will suffice to say here that both, in their very different ways, thought that

the double nature of man, as being partly instinctive and partly rational, was of vital importance in human ethics. After these preliminaries I will now give a critical account of Kant's theory.

The theory may be summed up in the following propositions. (1) Nothing is intrinsically good but a good will. Kant tries to prove this by taking other alleged intrinsic goods, such as happiness, intellectual eminence, etc., and showing that each may be worthless or positively evil when not combined with a good will. This argument is fallacious. If we accept the alleged facts they prove only that a good will is a *necessary constituent* of any whole which is intrinsically good. It does not follow, though it may of course be true, that a good will has itself any intrinsic value.

(2) A good will is one that habitually wills rightly.

(3) The rightness or wrongness of a volition depends wholly on the nature of its motive. It does not depend on its actual consequences. And it does not depend on its intended consequences except in so far as the expectation of these forms part of the motive. Of course a mere idle wish is of no moral value. But, provided we genuinely try to carry out our intention, and provided our motive is right, then the volition is right no matter what its consequences may be.

(4) The next question that arises is therefore: "What is the criterion of rightness of motive?" Before answering this question we must draw some distinctions among voluntary actions. In the first place we may divide them into *Actions on Impulse* and *Actions on Principle*. I will begin by illustrating this distinction. Suppose that I want to relieve a certain man who is in distress, simply because

I like him personally or because the sight of his distress makes me feel uncomfortable. Then I might not want to relieve a precisely similar man in a precisely similar situation if I did not happen to like him or if his distress were not thrust under my nose. This kind of voluntary action is impulsive. No doubt it has *causes*; there is something in the particular case which excites some conative disposition in me. But it is not, strictly speaking, done for a *reason* or on any *principle* which goes beyond this particular case. Now contrast this with the case of a member of the Charity Organisation Society giving relief to a complete stranger. He analyses the situation to see whether it does or does not come under a certain rule or principle of action which he has accepted. If it does, he gives relief; if it does not, he refuses it. And he would treat in exactly the same way any other man whose case had the same features. This is an example of action on principle. The agent had a *reason* for his action. And, if he stated his reason, his statement would always take the following form: "This case has such and such characteristics; and *any* case having these characteristics ought to be treated in such and such a way." Now Kant holds that an action cannot be right unless it is done on some general principle, which the agent accepts.

(5) This, however, is not a sufficient criterion of rightness. Kant divides principles or maxims of conduct into two classes, which he calls *Hypothetical* and *Categorical Imperatives*. A hypothetical imperative is a principle of conduct which is accepted, not on its own merits, but simply as a rule for gaining some desired end. Suppose that I refuse to make a certain statement on a certain occasion, for the reason that it would be a lie, and that lies ought not to be

told. Suppose that my ground for believing that lies ought not to be told is that they undermine confidence and thus reduce human happiness. Then the principle that lies ought not to be told would be, for me, a merely hypothetical imperative. It is accepted as a rule for maintaining human happiness, and not on its own merits. It is thus both contingent and derivative. It is contingent, because conditions are conceivable in which lying would not reduce human happiness, and in such conditions I should no longer accept the principle. And it is derivative, because my acceptance of it in existing circumstances depends on my desire for human happiness. The latter is my ultimate motive for not lying. A categorical imperative would be one that is accepted on its own merits, and not as a rule for gaining some desired end. If an action were done on a principle which is a categorical imperative we might say that it was done *for* a principle, and not merely *on* a principle. In fact we can distinguish three cases, viz., action *in accordance with*, action *on*, and action *for* a principle. An impulsive action might happen to be in accordance with a principle, though it could not be done on principle nor for principle. Now Kant holds that there are categorical as well as hypothetical imperatives; a view which many philosophers would reject. And he holds that an action is right if and only if it is done on a principle which is a categorical imperative, *i.e.*, if it is done *for* a principle.

Why did Kant hold this view? His reason appears to be this. It seems evident to him that any action which, in a given situation, is right or wrong at all must be right or wrong, in that situation, for *any* rational being whatever, no matter what his particular tastes and inclinations may be. Now any hypothetical imperative presupposes a desire

for some particular kind of object. But different rational beings, or different species of rational beings, might like different kinds of objects. All men, *e.g.*, dislike the kind of sensation which we call toothache. But this fact has no necessary connexion with their rationality. There is nothing impossible in the supposition that there might be rational beings who liked the sensation of toothache as much as most men like the scent of roses. And it is conceivable that there might be rational beings who have no sensations at all; indeed many people would hold that this possibility is realised in the case of angels. Therefore no hypothetical imperative would be accepted by all rational beings as such. Hence, if there be any principles of conduct which would be accepted by all rational beings as such, they must be accepted on their own merits and must therefore be categorical imperatives.

(6) We come now to the final question: "What characteristic must a principle of conduct have in order to be accepted on its own merits by every rational being as such?" Kant's answer is that the feature which is common and peculiar to such principles must be a certain characteristic *form*, and not anything characteristic in their *content*. And the formal criterion is this. It is necessary and sufficient that the principle shall be such that anyone who accepts it as *his* principle of conduct can consistently desire that every one else should also make it *their* principle of conduct and should act upon it. This supreme criterion Kant often calls "*the Categorical Imperative*" or "*the Moral Law*". It would be better to call it the "Supreme Principle of Categorical Imperatives". For it is a second-order principle which states the necessary and sufficient conditions that must be fulfilled by any first-order principle if the latter is

to be a categorical imperative and action determined by it is to be morally right.

We may now sum up the theory. An action is right if and only if the agent's sufficient motive in doing it is the fact that he recognises it to be required in the circumstances by a right principle of conduct. A principle of conduct is right if and only if it would be accepted on its own merits by any rational being, no matter what his special tastes and inclinations might be. It must therefore be a principle which is acceptable to rational beings simply because of its intrinsic form, and not because it is a rule for gaining some desired end. And a principle will be acceptable to all rational beings if and only if each could consistently will that all should adopt it and act on it. This is the essence of Kant's theory, as I understand it; and I will now make some explanations and criticisms before considering the further developments of the theory. I will begin with some explanations.

(1) What are we to say about actions which are determined by a mixture of causes? Suppose I refrain from telling a lie to a certain man on a certain occasion. All the following three causes may be moving me in the same direction. I may have a special feeling of love or respect for him. I may desire human happiness, and believe that lying under the given circumstances would tend to diminish it. And I may accept the principle that lies ought not to be told as a categorical imperative. Does my action cease to be right because the first two cause-factors are present and are moving me in the same direction as the third? Kant certainly talks as if this were so. But I do not think that he need have taken this extreme view if he had recognised a certain ambiguity in the notion of "mixed

motives". Suppose that three cause-factors, x, y, and z, are all moving me in the same direction. It may be that they are severally necessary and jointly sufficient to determine my action. If so, the situation would properly be described by saying that I have a *single* motive which is *internally complex*. On the other hand, it may be that one of these motive-factors, e.g., x, would have sufficed to determine my action even if the others had been absent. Now all that Kant needs to maintain is that, when there is a plurality of cause-factors all moving the agent in the same direction, the action would be right if and only if it *would still* have been done for a principle even though the other factors had been absent.

(2) Kant has sometimes been counted as an extreme advocate of the infallibility of the individual conscience. This is a peculiarly foolish accusation. He nowhere suggests that a single first-order moral principle is self-evident. On the contrary the essence of his theory is to offer a single necessary and sufficient criterion by which every suggested principle of conduct must be tested and judged before it can rightly be accepted and acted upon.

(3) Kant has sometimes been blamed because no particular rules of conduct can be deduced from his general principle. It is said to be "empty", "sterile", and "merely formal". Since Kant was perfectly well aware that his general principle is merely formal and since he plainly regarded this as its great merit, we may assume that this objection rests on a misunderstanding. The relation of the Moral Law to particular Categorical Imperatives, such as "Lies ought not to be told", is not supposed to be like the relation of the Law of Gravitation to Kepler's Laws of Planetary Motion. It is much more like the relation of the general principle: "All

arguments of the form 'all M is P and all S is M entail all S is P' are valid" to a particular bit of reasoning of that form, such as: "All men are mortal and all Greeks are men, therefore all Greeks are mortal." You cannot deduce any particular argument from the general principle of the syllogism; but, if any particular argument in syllogistic form claims to be valid, you can test its claims by seeing whether it does or does not have the formal structure required by the general principle. Kant would say, I think, that it is no more the business of ethics to provide rules of conduct than it is the business of logic to provide arguments. The business of ethics is to provide a test for rules of conduct, just as it is the business of logic to provide a test for arguments.

I have now, I hope, removed the more obvious misunderstandings that may arise about Kant's theory. Let me then begin to criticise it. (1) We must admit at once, as a plain matter of fact, that certain principles are accepted and acted upon by many people who do not accept or act upon them simply as hypothetical imperatives. It is perfectly certain that many people accept and act on the principle that lies ought not to be told, without thinking of whether the results of lying are desirable or undesirable. There are then imperatives which are here and now categorical for certain persons, and there is action for the sake of principles. To this extent Kant is right, and he has pointed out an important psychological fact which moralists like Spinoza and Hume tend to ignore. The utmost that a Utilitarian could honestly say of such facts is that the imperatives which are now categorical for Smith must once have been merely hypothetical, either for Smith himself or



for earlier members of his society or race. Accepted originally only as rules for gaining some desired end, they have now acquired such prestige that Smith accepts them for their own sake without thinking of their consequences. I am not at present concerned to criticise this theory of the origin of categorical imperatives. I wish simply to point out that there are imperatives which are here and now categorical for certain persons.

(2) We saw that the premise which is alleged to entail the most characteristic parts of Kant's theory is the following. Any action which, in a given situation, is right or wrong at all would be right or wrong for *any* rational being whatever in that situation, no matter what his special tastes and inclinations might be. Now this premise seems to me plainly false. I think it is true that *some* actions would be right, and that *some* would be wrong, in a given situation, quite independently of the tastes and inclinations of the agent. *E.g.*, if he were a member of a board of electors it would be his business to ignore his personal liking or disliking for any of the candidates. But it is equally certain that some actions would be right if done by an agent with one set of tastes and inclinations and wrong if done in precisely the same situation by an agent with certain other tastes and inclinations. If the agent, instead of having to decide whether to choose A or B for a professorship, had to decide whether to make a proposal of marriage to A or to B, it is perfectly obvious that his personal likings and dislikings would be relevant to the rightness of his action.

This conclusion may perhaps be reinforced by the following consideration. Every one admits that what is right or wrong for a given agent at a given moment depends

in part on the nature of the situation in which he is placed at the moment of acting. Now among the factors in the situation are the tastes and inclinations of the other rational beings with whom the agent is concerned. And, although these are not relevant to the rightness or wrongness of *some* actions, they quite certainly are relevant to the rightness or wrongness of others. Now it seems very far-fetched to suppose that, whilst the tastes and inclinations of all other rational beings are often ethically relevant, those of the agent are never so.

The most then that I could admit is that there may be some actions which would be right and some which would be wrong in a given situation no matter what might be the tastes and inclinations of the agent. Since there are certainly others of which this is not true, Kant's theory of ethics must at best be incomplete. For his criterion at best will apply only to this department of morality and not to morality as a whole. Perhaps this is the only part of morality for which any general criterion can be given; but that is another matter.

(3) Supposing that there are some actions which would be right and some which would be wrong in a given situation for *any* rational being, does it follow that the principles on which such actions are done must be categorical and not hypothetical imperatives? Kant's ground for asserting this is, as we have seen, that a hypothetical imperative is accepted only as a rule for gaining some desired end; and that there is no end which all rational beings as such must desire. This seems highly plausible. But it is necessary to draw a distinction between two different questions. (a) Is there any end which all rational beings who contemplated it would judge to be desirable? And (b) is there any end

such that one could infer from the concept of a rational being that any such being must judge it to be desirable?

The answer to the second question is, no doubt, in the negative. The concept of a rational being is the concept of a being who is capable of intuiting necessary connexions, of making inferences both deductive and problematic, and of forming *a priori* concepts. It is quite impossible to see directly or to infer deductively that such a being would find anything desirable, still less that it would find so-and-so—*e.g.*, general happiness—desirable. We could, however, infer the hypothetical proposition that, *if* anything—*e.g.*, general happiness—be intrinsically desirable, *then* such a being would be able to see this if he contemplated the notion of general happiness. For this hypothetical proposition does follow from the premise that the being is capable of intuiting necessary connexions, and this is part of the definition of a rational being.

It should now be evident that the negative answer which we have had to give to the second question has no bearing whatever on the first question. Let us take a parallel case from mathematics. We could not infer from the concept of a rational being that all rational beings are capable of seeing that the square-root of 2 is an irrational number. We could infer only that, *if* this be a necessary proposition, *then* all rational beings will be capable of seeing its truth and necessity. Yet, in point of fact, the proposition that the square-root of 2 is an irrational number *is* a necessary truth, and all rational beings who are properly trained and pay attention to the very simple proof of it can see this for themselves. In exactly the same way it might, *e.g.*, be the case that general happiness is intrinsically desirable. In that case every rational being who contemplated the notion of

general happiness with enough attention would be able to see that it is desirable, though it is certainly not deducible from the notion of a rational being that he should find general happiness desirable.

We see then that it is perfectly possible that there may be ends which every rational being who contemplated them *would in fact* recognise to be intrinsically desirable, although there are no ends with regard to which it could be *inferred* from the concept of a rational being that he would find them desirable. It is therefore possible that even those actions which would be right or wrong in a given situation regardless of the special tastes and inclinations of the agent may be done on principles which are accepted as hypothetical, and not as categorical, imperatives.

(4) Let us suppose, however, that there are some principles which are accepted by all rational beings as categorical, and not merely as hypothetical, imperatives. Kant, as we know, claims to infer from the concept of a rational being the necessary and sufficient conditions which such a principle must fulfil. Can this be done? It seems perfectly clear to me that it cannot. It appears possible only so long as the concept of a rational being is left unanalysed in an atmosphere misty with the incense of adoration. When it is brought into the common light of day and analysed, as we have done to it, we see that one can no more infer that a rational being would recognise any principle as right than that it would recognise any end as desirable. Still less could we infer from the concept of a rational being that it would accept all those principles and only those which answered to a certain formal condition.

Why did Kant imagine that he could infer such a criterion from the concept of a rational being? Presumably his mind

must have moved in the following way. If there be anything which a rational being as such might be expected to dislike, it will be logical inconsistency. So a rational being would *reject* any principle whose acceptance *would* involve him in logical inconsistency. Then Kant must have jumped, in some way which I cannot pretend to explain, from this proposition to the proposition that a rational being would *accept* any principle whose acceptance would *not* involve him in logical inconsistency. This is of course absolutely indefensible, and charity bids us turn our eyes from the painful spectacle.

The truth on this matter seems to me to be the following. There may be principles which would be accepted as categorical imperatives by all rational beings. But, if so, each is accepted because of its special content, and not because of any peculiarity in its form. I think that the principle that gratitude is due to our benefactors is a plausible example of such a principle. Now, if this would be accepted by any rational being who understood the meaning of the terms "gratitude" and "benefactor", it is because there is an intrinsic relation of *fittingness* between the former kind of emotion and the latter kind of object. It is accepted then, if at all, because a rational being can see that a certain relation necessarily relates those two special terms. It is not accepted because of anything in its general form.

Again, it is possible that there may be some characteristic which is common and peculiar to all the principles which would be accepted as categorical imperatives by all rational beings. If so, this characteristic might be abstracted and used in future as a test for any principle which claimed to be a categorical imperative acceptable to all rational beings.

But it is quite certain that such a criterion, even if it exists, could not be deduced from the concept of a rational being. If it exists and can be discovered at all, its discovery and establishment must take place in the following way. We should have to compare a number of admittedly categorical imperatives with each other, and contrast them with a number of principles which were admittedly not categorical imperatives. We might then discover that there is a certain characteristic common and peculiar to the former. Finally we might be able to see by an act of intuitive induction that *any* principle which had this characteristic would necessarily be a categorical imperative, and that the converse of this is also necessarily true.

(5) Let us now consider Kant's criterion in greater detail. The criterion is that a principle must be such that any rational being who proposes to accept it could consistently will that it should be accepted and acted upon by every one. As Kant points out, a principle might fail to pass this test in two different ways. In the first place, the very supposition of every one acting in a certain manner in certain circumstances might be self-contradictory. This case, he thinks, would be illustrated by the principle that every one should refuse to pay back money which was originally lent to him on promise of repayment. Secondly, the supposition of every one acting in a certain way in certain circumstances might not be self-contradictory, but it might be that the consequence of other people acting on this principle would be to hinder me from acting on it. In that case I could not consistently will that the principle should be generally accepted and acted upon. Kant thinks that this case would be illustrated by the principle that every one should seek to make himself as happy as possible

without regard to the happiness of any other man except in so far as this subserves his own happiness.

It is very difficult to think of any principle which would strictly be self-contradictory when generalised. I cannot see that Kant's example of promise-breaking is a case in point. If the principle were generally acted upon people in difficulties would, no doubt, soon cease to be able to get help from others by promises of repayment. So the real position is that the desire that every one who has got out of a difficulty by making a promise shall be allowed to break the promise afterwards is incompatible with the desire that every one who is in difficulties shall be able to get out of them by making promises. The incompatibility consists in the fact that, human memory and human motives being what they are, the fulfilment of the former desire would prevent that of the latter. In fact human nature is so constituted that, if the principle were generally acted upon, there would very soon be no more cases for it to apply to. This is plainly not, as it ought to be on Kant's theory, a case of *self-contradiction* or purely *formal* inconsistency.

The example of the second case is equally unfortunate, though in a different way. If it is to be relevant at all we must suppose that the principle of Egoism is accepted as a *categorical* imperative, and not as a mere rule for gaining maximum personal happiness. My acceptance of the principle therefore does not presuppose a desire for my own happiness or a belief that this is the most effective way to secure it. Now all that Kant shows is that the acceptance of this principle by others would be likely to lead to consequences detrimental to my happiness. Thus he shows only that my desire that every one should accept and act on

the principle of Egoism would be inconsistent with my desire for my own maximum happiness. And this is wholly irrelevant. For we ought to be testing the claims of Egoism to be a *categorical* imperative; and, as such, it does not presuppose the existence of a desire for my own happiness. And, so far as I can see, if anyone did propose to accept the principle of Egoism as a categorical, and not as a hypothetical, imperative, there would be no way of refuting him. I should claim to see by inspection that he was mistaken; but there I should have to leave the matter.

The only importance of Kant's criterion is as a means of avoiding personal bias. If I feel inclined to approve a certain action by myself in a certain situation it is always desirable to consider what I should think if the same kind of action were done in the same situation by another man. If I find that I should condemn it in another, and yet can see no relevant differences between him and me, the chances are that my approval of the action in my own case is due to some personal bias. But it is important to notice that this principle, like the Principle of Indifference in Probability, cannot be used mechanically. I have to judge for myself what differences between me and another are, and what are not, ethically relevant to this kind of action in this kind of situation. And, beyond a certain point, this cannot be reduced to general rules.

I have now criticised the most fundamental points in Kant's theory. I will therefore pass to the further developments of it. (1) Kant gives two other forms of the Supreme Principle of Morality. The second form is: "Treat every rational being, including yourself, always as an end and never as a mere means." The third form is: "A principle



of conduct is morally binding on me if and only if I can regard it as a law which I impose on myself." He regards the three forms of the Moral Law as logically equivalent, but thinks that each emphasises a different aspect of it. I cannot see that the three forms are logically equivalent; but the two additional principles are interesting, and deserve some slight comment.

(a) The second formula plainly contains an important truth, but it stands in need of some qualification. In the first place so far from being thought wrong, it is thought to be an act of specially heroic virtue in certain circumstances for a soldier to sacrifice his life for his country, or for a doctor to do so for his patients, or for a scientist to do so for the advancement of knowledge. It must be admitted, however, that, although we thus admire people in certain circumstances for treating *themselves* as mere means, we should not feel justified in treating them in that way without their consent. Again, there seem to be cases in which you must either treat A or treat B, not as an end, but as a means. If we isolate a man who is a carrier of typhoid, we are *pro tanto* treating him merely as a cause of infection to others. But, if we refuse to isolate him, we are treating other people *pro tanto* merely as means to his comfort and culture. The fact which this formula exaggerates seems to be the following. Every rational being (and, I should add, every sentient being) has as such certain claims to consideration which it is always wrong to ignore. But, although such claims must always be *considered*, they need not, and indeed cannot, all be *satisfied* in full. For they may conflict with each other, and then some compromise must be struck between them. And in certain cases we approve a man for voluntarily abating or renouncing his claims, though we

should not judge it right to impose this abatement or renunciation on him in the circumstances.

(b) The third formula also contains an important truth expressed in an exaggerated form. It is not necessary that a principle of conduct should be "self-imposed"; indeed it is doubtful whether any clear meaning can be attached to this notion. But it is true that an action done for a principle has no moral value unless the agent freely and wittingly accepts the principle for which it is done. It is important to notice, however, that a principle may be freely and wittingly accepted in two quite different ways. (i) I may accept it directly, because, on inspection, I persuade myself that it is right to act in such and such a way in such and such circumstances. (ii) I may not be able to see this directly. But I may be told that it is so by someone whom I believe to have greater moral insight in general or in this special department of conduct than I have. Or, again, I might believe that it had been ordained by a good God for reasons which I cannot understand but which are certainly adequate. In such cases it would be my duty to accept the principle and act on it, even though I could not see its truth by direct inspection of the terms. The kind of case which Kant's third formula is meant to cut out is where the principle is accepted *merely* on tradition, or *merely* from the fear that God will punish me if I do not act in accordance with it. In such cases actions done from this principle would plainly have no moral value.

(2) Kant draws a very important distinction between the *intrinsically* good (*Summum Bonum*) and the *complete* good (*Bonum Consummatum*). As we have seen, he holds that a will which habitually wills rightly is intrinsically good and that nothing else is so. Pleasure and pain, *e.g.*, by

themselves are neither good nor bad. Nevertheless pleasure and pain are capable of adding to or detracting from the total value of a situation. A being who wills rightly *deserves* a certain degree of happiness, and one who wills wrongly deserves a certain degree of unhappiness. The *moral* value of each being who wills rightly is of course quite independent of whether he gets the amount of happiness which his right willing deserves. But the *total* value of a universe in which each being who willed rightly did get the amount and kind of happiness which his right willing deserved would plainly be greater than that of an otherwise similar universe in which happiness was absent or was not distributed on this principle. If we adopt a useful distinction of M'Taggart's we may say that the total value *in* the universe would be the same in both cases; but the total value *of* the universe would be greater in the first case than in the second. The *complete* good would be a system of perfectly virtuous beings, each enjoying that amount and kind of happiness which his virtue deserved; although the only *intrinsic* good is right willing, and although an action is never right if it be done *for the sake of* a reward.

This doctrine of Kant's is perfectly consistent, and I agree with a large part of it. I accept the notion of desert, and I agree that it is better that there should be virtue with the appropriate amount of happiness than the same degree of virtue with less happiness. I am not, however, convinced that pleasure and pain have no intrinsic value; though I am more inclined to think that pain is intrinsically evil than that pleasure is intrinsically good. Doubts on these points would introduce many complications into the elegant simplicity of Kant's doctrine; and it would be out of place to pursue the subject further here.

(3) The last point to be noticed is Kant's theory of Moral Obligation. Kant, like Spinoza, is greatly impressed with the double nature of man, as being partly a creature of passion, impulse, instinct, and sensation, and partly a rational being. Kant and Spinoza both held that the rational aspect of man's nature is the more fundamental. Neither of them gave a satisfactory account of the relations between the two; but for this no other philosopher is in a position to cast stones at them. Kant's theory is, roughly, that the non-rational aspect of a human mind is the way in which such a mind inevitably appears to itself. It is needless to waste much time over this theory. Either the human mind, as it really is, is purely active and rational or it is not. If it is not, the problem of the relation between the active rational side and the passive, emotional, and sensuous side remains where it was. But, if it is purely active and rational, the problem of how it comes to appear to itself as partly passive, emotional, and sensuous arises at once and is plainly insoluble. Most of Kant's theory of freedom consists of a rapid shuffle between one and the other horn of this dilemma, and resembles an unskilful performance of the three-card trick rather than a serious philosophical argument.

Still, the double nature of man remains a fact, whatever may be the right explanation of it. And both Kant and Spinoza held that the characteristic experiences of obligation and moral struggle are closely bound up with it. It will be remembered that Spinoza said that, if, *per impossibile*, a man were born with nothing but clear ideas and active emotions, he would not know the meaning of good and evil though he would in fact enjoy the highest good. Now Kant distinguishes between what he calls the *Good Will* and the

*Holy Will.* And he ascribes the experience of moral obligation in human beings to the fact that their wills are good, but not holy. A good will is one which always has the power to act on right principles, but is also susceptible to other kinds of solicitation, *e.g.*, special impulses and passions, desires for certain ends, and so on. The wills of human beings, in this life at any rate, are in this state. They *need* never act on these other solicitations, but they are always subject to them. A holy will would be one in which every tendency to action except for the sake of a right principle was absent. Such a will must be ascribed to God, and it may perhaps belong to angels and to just men made perfect. Now, in the case of a holy will, there is no question of duty or obligation. All obligation is the obligation of a being whose will is good, but not holy, to act *as if* its will were holy; *i.e.*, never to act on any motive but right principle, although other motives do in fact solicit it. Kant holds that the fact that we are under an obligation to act in this way implies that we always *could* have done so even in those cases where we in fact did not. If you ask him how this is possible his answer is that, as we really are, we are purely rational active beings and therefore can always behave as such. The particular inclinations, impulses, and passions are only aspects under which a being who is purely active and rational appears to itself.

I have already shown that this solution is metaphysically impossible. It is equally unsatisfactory from an ethical point of view. Either what appear to me as my passive emotions and irrational impulses are *purely* delusive appearances, or they correspond to something in my real self. If it be a pure delusion that I have irrational impulses it must be a pure delusion that I ever act on them, and therefore

a pure delusion that I ever act wrongly. If, on the other hand, these appearances do correspond to something in my real self, then it is indeed possible that I should really act wrongly at times. But my self will *really* be mixed; and there is no explanation why a self which is in fact of mixed nature should always be able to act as if its nature were purely rational. The truth is that Kant takes the mixture to be real when he is dealing with purely ethical questions, and takes it to be delusive when he is trying to give a rational theory of the metaphysical consequences which he thinks are entailed by the ethical facts.

In Kant's theological works there are traces of a different, but no more satisfactory, theory. If it be admitted that the notion of duty has any application it must be admitted that some actions which I actually have done are actions that I ought not to have done, and that some which I failed to do ought to have been done by me. But to say that I did X but ought not to have done it implies that I could have avoided doing X. And to say that I omitted X but ought to have done it implies that I could have done it. It must therefore be admitted that the fact of duty and moral obligation implies that my present actions are not completely determined by my present character and situation. Yet my actions are events in time, and Kant claims to have proved in the *Critique of Pure Reason* that all events in time are completely determined. The solution is to distinguish between me as a noumenon and me as a phenomenon. Every act of mine could be predicted by a person who knew enough about my circumstances, my innate dispositions, my past actions, and the laws of empirical psychology. But, on the other hand, the whole series of my acts, emotions, etc., is the manifestation of a single

noumenal self which is not in time. Suppose now that my empirical self contains, among other events, certain wrong actions. These, as we have seen, must be so connected with the other events in my empirical self and the rest of the empirical world that they could have been predicted with complete certainty by anyone who had enough knowledge of empirical facts and laws. Nevertheless I am responsible for them. For my noumenal self *could have* manifested itself in time as an empirical self which did not contain these wrong actions but contained right actions instead. And I cannot disclaim responsibility for the fact that my noumenal self chose to appear as an empirical self which contains wrong actions rather than as an empirical self which contains nothing but right actions. For it is my real self; and the whole series of events which make up my empirical self is just one possible manifestation of my noumenal self, which the latter freely and timelessly chooses in preference to other possible manifestations.

This theory has at least the merit of admitting that the noumenal self can really choose wrongly. But the notion of a timeless and undetermined choice by a noumenon of the series of phenomena by which it shall be manifested in time is quite unintelligible. And there is a further difficulty. We have to suppose that each noumenon independently of all the rest makes an undetermined choice of the series of phenomena by which it shall be manifested in time. Yet these various independently chosen series of phenomena must all fit into each other in such a way that the whole phenomenal world forms a single system in which even the minutest items are subject to invariable rules of sequence and co-existence. This seems absolutely incredible. And so both Kant's attempts to reconcile complete determinism

in the empirical self with the possibility of real wrong and right doing seem to break down.

Nevertheless, from a purely ethical point of view, Kant's theory of moral obligation seems highly plausible, if we put it in the form that "duty" and "obligation" have meaning and application only to beings who are fundamentally, and yet not wholly, rational. And he is right in holding that an obligation to do or to abstain implies power to do or abstain. But it is a very difficult problem to say what precise meaning is to be given to the two highly ambiguous words "fundamentally" and "power"; and it is certain that Kant has failed to answer this question satisfactorily.

Kant holds that there is a peculiar kind of emotion which a being who has a good, but not a holy, will experiences when he contemplates moral purity. This emotion, which he calls *Achtung*, is a species of awe. On such occasions the being, who from his mixed nature belongs both to the world of sense and to the supersensible world, is getting a peep, and the only direct peep which he can get in this life, into the latter. This glimpse humbles and even frightens him, in so far as his nature is partly animal and sensuous; yet, at the same time, it exalts him, in so far as his nature is fundamentally rational, by reminding him that he is a citizen of the supersensible world. Here again it seems clear that Kant is describing a genuine fact in terms which most of us can understand and accept in outline, even though we might hesitate to follow him in points of detail.

It remains to say something about Kant's ethical arguments for immortality and for the existence of God. The argument for immortality is as follows. We are under a



moral obligation, not merely to act rightly on all occasions, but also to make ourselves perfect. Now we should not be perfect until we had no inclination to act wrongly. And we shall never cease to have the inclination to act wrongly unless we become purely rational, and the sensuous, impulsive, passive side of our nature is completely eliminated. Now, in terms of time, this would take an infinite time to accomplish. But, since we are morally obliged to aim at this result, it must be possible to reach it. Therefore we must be immortal to allow us time to do so.

The following criticisms must be made on this argument :

(a) The command to make ourselves perfect is not to be taken literally. It is merely a rhetorical way of saying : "Never be contented with your present level of moral achievement." No doubt we always can improve our moral characters so long as we are alive. But this does not imply that we shall ever be able to make them perfect. (b) Kant's premises are really inconsistent with each other. One premise is that moral perfection must be attainable or it could not be our duty to seek it. The other premise is that it is attainable only after an unending time. And this is surely equivalent to saying that it is not attainable at all.

The ethical argument for the existence of God is as follows. Nothing is intrinsically good except virtue, which consists in doing right without any ulterior motive. Every one can always act virtuously whether God exists or not. But, although virtue is the only *intrinsic* good, it is not, as we have seen, the *complete* good. The complete good is composed of virtue with the appropriate amount of happiness. Now we can say of the complete good that it *ought* to exist. But what ought to be must be possible, and therefore the necessary conditions of its possibility must be actual. Now

there is no necessary connexion between virtue and happiness either logically or by way of ordinary natural causation. There is no logical connexion, because virtue cannot be defined in terms of happiness, and there are many other kinds of happiness beside the feeling of satisfaction with one's own virtue. And there is no causal connexion by the ordinary laws of Nature. Virtue depends wholly on oneself, and, on Kant's view, can always be realised no matter how unfavourable the conditions may be. But happiness depends largely on one's innate tastes and dispositions, on the state of one's bodily health, and on external circumstances. And a perfectly virtuous man has no more control over these than a vicious one. The position then is this. The complete good must be capable of existing, since it ought to exist. One of its factors, viz., perfect virtue, is possible under all circumstances. But the other factor, viz., the deserved amount of happiness, will be realised only if the course of Nature be deliberately overruled so as to secure it. And the only way in which we can conceive this happening is by supposing that Nature is dependent on a powerful, benevolent, and moral being, who arranges that in the long run virtue shall be rewarded by the appropriate amount of happiness.

This argument is open to the following criticisms :

(1) There are two different senses of "ought", and one of these involves factual possibility whilst the other involves only logical possibility. If I say : "You ought to do so and so," I do imply that you *could* do so and so in some sense which is not merely that there is no logical contradiction in the notion of your doing it. But, if I say : "So and so ought to exist," I imply only that it would involve no logical contradiction, and that any being who could bring it about ought to try to do so. But it does not imply

that there actually is any such being. Thus Kant is entitled only to the hypothetical proposition: "If a perfect God existed he would order the course of Nature so that virtue would receive its appropriate reward in happiness." He is not entitled to the categorical conclusion that such a being exists. (2) It seems to me that there is a certain inconsistency between Kant's position in this argument and his position in the argument for immortality. In the latter it is assumed that we shall not be morally perfect until we have completely got rid of the passive, sensuous, and emotional side of our nature. In the argument about God it is assumed that the happiness, which is an essential feature in the supreme good, is not the mere consciousness of virtue, but is something further added as a reward of virtue. But how could we feel any such happiness if we had no sensations or emotions left?

## CHAPTER VI

## Sidgwick

SIDGWICK'S *Methods of Ethics* seems to me to be on the whole the best treatise on moral theory that has ever been written, and to be one of the English philosophical classics. This does not of course imply that Sidgwick was a better man or an acuter thinker than the other writers with whose theories we have been dealing; for he inherited the results of their labours, and he thus had over them an advantage of the kind which any contemporary student of mathematics or physics has over Newton and Faraday. But, even when this advantage has been discounted, Sidgwick must continue to rank extremely high. He combined deep moral earnestness with complete coolness and absence of moral fanaticism. His capacity for seeing all sides of a question and estimating their relative importance was unrivalled; his power of analysis was very great; and he never allowed the natural desire to make up one's mind on important questions to hurry him into a decision where the evidence seemed inadequate or conflicting. Those who, like the present writer, never had the privilege of meeting Sidgwick can infer from his writings, and still more from the characteristic philosophic merits of such pupils of his as M'Taggart and Moore, how acute and painstaking a thinker and how inspiring a teacher he must have been. Yet he has grave defects as a writer which have certainly detracted from his fame. His style is heavy and involved, and he seldom allowed