



Groundwork of the Metaphysics of Morals

IMMANUEL KANT

Plain English Translation

MOUSEIA



Groundwork of the Metaphysics of Morals



IMMANUEL KANT

Plain English Translation

Translated by Michael Berardi

MOUSEIA

THIS EDITION

Mouseia's complete machine-assisted Plain English translation, made independently and directly from the same German, with every section, page reference and author's note.

TRANSLATOR

Michael Berardi

THE WORK

Written 1785 · Moral philosophy

SOURCE

Original text from archive.org.

GAPS AND UNCERTAINTIES

This is the text printed in the 1911 Akademie-Ausgabe, not a diplomatic transcription of the 1785 first printing. AA 386 is blank.

Mouseia · mouseia.org

Edition dated 26 October 2026

Contents



Preface	1
Section I	8
Section II	22
Section III	62

Preface



385 Groundwork for the Metaphysics of Morals, by Immanuel Kant.

387 Preface.

Ancient Greek philosophy was divided into three sciences: *physics*, *ethics* and *logic*. This division fits the nature of the subject perfectly. It needs no improvement, except perhaps to add the principle on which it rests. In this way we can both make sure that the division is complete and determine the necessary subdivisions correctly.

All cognition by reason is either *material* or *formal*. Material cognition considers some object. Formal cognition deals only with the form of the understanding and of reason itself, and with the general rules of thinking as such, without distinguishing between objects. Formal philosophy is called *logic*. Material philosophy deals with determinate objects and with the laws they are subject to, and it is in turn twofold. For these laws are either laws of *nature* or laws of *freedom*. The science of the first is called *physics*, and the science of the second is *ethics*. The first is also called the doctrine of nature, and the second the doctrine of morals.

Logic cannot have an empirical part. By that I mean a part in which the universal and necessary laws of thinking would rest on grounds taken from experience. If it had such a part, it would not be logic, which is a canon for the understanding or reason that holds for all thinking and has to be demonstrated.

Natural philosophy and moral philosophy, by contrast, can each have an empirical part. The reason is that natural philosophy must lay down its laws for nature as an object of experience, and moral philosophy must lay down its laws for the human will, insofar as nature affects it. The laws of the first are laws by which everything

388 happens. The laws of the second are laws by which everything ought to happen. Yet they also take into account the conditions under which it often does not happen.

All philosophy, insofar as it rests on grounds of experience, we can call *empirical* philosophy. Philosophy that presents its teachings solely from principles a priori we can call *pure* philosophy. When pure philosophy is only formal, it is called logic. When it is limited to determinate objects of the understanding, it is called metaphysics.

In this way the idea of a twofold metaphysics arises: a *metaphysics of nature* and a *metaphysics of morals*. Physics will therefore have its empirical part, but also a rational part. Ethics will too. Here, though, the empirical part could be called in particular *practical anthropology*, and the rational part properly *morals*.

All trades, crafts and arts have gained from the division of labor. One person does not make everything. Instead each person sticks to a certain kind of work that differs clearly from the others in how it is handled, so as to do it with the greatest perfection and with more ease. Where work is not distinguished and divided in this way, where everyone is a jack-of-all-trades, the trades are still in the deepest barbarism. It would be worth considering, in itself, whether pure philosophy in all its parts does not call for its own particular person, and whether the whole scholarly trade would not be better off if a warning were given to certain people. These people are used to selling the empirical mixed with the rational, to suit the public's taste, in all sorts of proportions unknown even to themselves. They call themselves independent thinkers, and they call others, who prepare only the rational part, brooders. The warning would be not to carry on two businesses at once. The two are very different in how they are handled. Each may perhaps need its own special talent. Joining them in one

person produces only bunglers. But here I ask only this: does the nature of the science not require that the empirical part always be carefully separated from the rational part? Before physics proper (the empirical kind), a metaphysics of nature should be put first. Before practical anthropology, a metaphysics of morals should be put first. Both would have to be carefully cleaned of everything empirical, so that we can know how much pure reason in both cases

389 can achieve, and from what sources it draws this teaching of its own a priori. It does not matter whether the latter task is taken up by all teachers of morals (whose name is Legion) or only by some who feel called to it.

My aim here is really moral philosophy, so I limit the question just put to this one: does one not think that it is extremely necessary to work out, at last, a pure moral philosophy, one completely cleaned of everything that may be merely empirical and belongs to anthropology? That there must be such a philosophy is obvious by itself from the ordinary idea of duty and of moral laws. Everyone must admit the following. A law, if it is to count as moral, that is, as the ground of an obligation, must carry absolute necessity with it. The command "You must not lie" does not hold merely for human beings, as if other rational beings need not heed it. The same goes for all the other laws that are properly moral. It follows that the ground of obligation here must not be sought in the nature of the human being, or in the circumstances of the world in which the human being is placed. Instead it must be sought a priori, solely in concepts of pure reason. And any other instruction that is founded on principles of mere experience can be called a practical rule, but never a moral law. This holds even for an instruction that is general in a certain respect, insofar as it rests on empirical grounds in even the smallest part, perhaps only in respect of one motive.

Moral laws, together with their principles, therefore differ not only in essence from everything else in practical cognition that has anything empirical in it. All moral philosophy also rests entirely on its pure part. When it is applied to human beings, it borrows nothing at all from knowledge of them (anthropology). Instead it gives them, as rational beings, laws a priori. These laws do, to be sure, still require a power of judgment sharpened by experience. It is needed

partly to tell in which cases the laws apply, and partly to win them entry into the human will and force for carrying them out. For human beings are themselves affected by so many inclinations. They are capable of the idea of a practical pure reason, but they are not so easily able to make it effective in their way of life in concreto.

A metaphysics of morals is therefore indispensable and necessary. It is not needed merely for a speculative motive, to investigate the source of the a

390 priori practical principles that lie in our reason. It is needed because morals themselves remain subject to all kinds of corruption as long as that guide and highest norm for judging them correctly is missing. For something that is to be morally good, it is not enough that it be *in accordance* with the moral law. It must also be done *for the law's sake*. Otherwise the accordance is only very accidental and uncertain, because the immoral ground will now and then produce actions that fit the law, but more than once actions that go against it. Now the moral law, in its purity and genuineness (which is just what matters most in practice), is to be sought nowhere else than in a pure philosophy. So this philosophy (metaphysics) must come first. Without it there can be no moral philosophy at all. Indeed, whatever mixes those pure principles in with the empirical ones does not deserve the name of philosophy. (For this is just how philosophy differs from ordinary cognition by reason: it presents in a separate science what ordinary cognition grasps only in a mixed-up way.) Much less does it deserve the name of moral philosophy, because by this very mixing it damages even the purity of morals themselves, and it works against its own purpose.

Let no one think that what is asked for here is already found in the preparatory study that the famous *Wolff* put before his moral philosophy, the one he called *general practical philosophy*, and that a wholly new field is therefore not really being entered here. Precisely because it was meant to be a general practical philosophy, it takes up no will of any particular kind. It does not take up, for example, a will that is determined entirely from principles a priori, without any empirical motives, which one could call a pure will. Instead it looks at willing in general, with all the actions and conditions that belong to willing in this gen-

eral sense. This is how it differs from a metaphysics of morals, just as general logic differs from transcendental philosophy. General logic presents the actions and rules of thinking *in general*. Transcendental philosophy presents only the particular actions and rules of *pure* thinking, that is, of the thinking by which objects are known entirely a priori. For the metaphysics of morals is meant to investigate the idea and the principles of a possible *pure* will. It is not meant to investigate the actions and conditions of human willing in general, most of which are drawn from psychology. The fact that in the general

391 practical philosophy moral laws and duty are also spoken of (though with no right to do so) is no objection to my claim. For the authors of that science stay true to their idea of it here too. They do not distinguish the motives that, as such, are represented entirely a priori by reason alone and that are properly moral, from the empirical motives that the understanding raises to general concepts merely by comparing experiences. Instead they consider the motives without paying attention to the difference in their sources. They look only at how large or small the total of them is, since all are regarded as being of the same kind. In this way they form their concept of *obligation*. This concept is certainly anything but moral. Still, it is as good as can be demanded in a philosophy that passes no judgment at all on the *origin* of all possible practical concepts, whether they occur a priori or merely a posteriori.

I intend to supply a metaphysics of morals some day, and I let this groundwork come first. Strictly speaking, there is no other foundation for such a metaphysics than the critique of a *pure practical reason*, just as the critique of pure speculative reason, which has already been supplied, is the foundation for metaphysics. But for one thing, the former is not as extremely necessary as the latter. In morals, human reason can easily be brought to great correctness and completeness, even with the most ordinary understanding. In theoretical but pure use, by contrast, it is dialectical through and through. For another thing, I require of a critique of pure practical reason that, if it is to be complete, its unity with speculative reason must be able to be shown at the same time in a common principle. For in the end there can be only one and the same reason,

which must differ merely in its application. I could not yet reach such completeness here without bringing in considerations of a quite different kind and confusing the reader. For that reason I have used the title *Groundwork for the Metaphysics of Morals* instead of the title *Critique of Pure Practical Reason*.

Third, even a metaphysics of morals is nevertheless capable, despite its forbidding title, of a high degree of popularity and of suitability for the ordinary understanding. So I find it useful to separate this preliminary work on the foundation from it, so that the subtle part

392 that is unavoidable in it need not be added later to teachings that are easier to grasp.

The present groundwork is nothing more than the search for and establishment of the *supreme principle of morality*. This alone makes up a task that is complete in its aim and that must be kept separate from all other moral inquiry. True, applying the same principle to the whole system would shed much light on my claims about this important main question, which so far has by no means been discussed satisfactorily. The adequacy that the principle shows everywhere would also strongly confirm those claims. But I had to give up this advantage. At bottom it would also serve self-love more than the common good. The ease of using a principle and its apparent adequacy do not give fully certain proof that the principle is correct. They rather arouse a certain partiality toward not examining and weighing it strictly for itself, without any regard to what follows from it.

In this work I have chosen my method as the one I believe most suitable, if one wants to take the road analytically from ordinary cognition to the determination of its supreme principle, and then back again synthetically from the examination of this principle and its sources to ordinary cognition, where the principle is found in use. The division has therefore come out like this:

¹ *First Section:* Transition from ordinary moral cognition by reason to philosophical cognition by reason.

- ² *Second Section:* Transition from popular moral philosophy to the metaphysics of morals.
- ³ *Third Section:* Final step from the metaphysics of morals to the critique of pure practical reason.

Section I



393 First Section

Transition from ordinary moral cognition by reason to philosophical cognition
Nothing anywhere in the world, or even beyond it, can be thought of as good without qualification except a *good will*. Understanding, wit, judgment, and whatever else the mind's *talents* may be called, or courage, determination, and persistence in a purpose as qualities of *temperament*, are certainly good and desirable in many ways. But they can also become extremely bad and harmful if the will that is to use these natural gifts is not good. The distinctive makeup of this will is therefore called *character*. The same holds for the *gifts of fortune*. Power, wealth, honor, even health and the whole sense of well-being and satisfaction with one's condition that goes by the name of *happiness*, give us courage, and often arrogance as well, where no good will is present to correct their influence on the mind, and with it the whole principle of acting, and to make that principle fit for purposes in general. There is also this to consider: a reasonable, impartial observer can never take pleasure even in the sight of the uninterrupted prosperity of someone who shows no sign of a pure and good will. So a good will seems to be the indispensable condition even of being worthy of happiness.
Some qualities even help this good will itself and can make its work much easier. But despite that,

394 they have no intrinsic, unconditional worth. They still presuppose a good will, which limits the high regard we otherwise rightly have for them and does not let us count them as good without qualification. Moderation in emotions and passions, self-control, and sober reflection are not only good in many ways. They even seem to form part of a person's *intrinsic* worth. Yet they are far from being good without qualification, however unconditionally the ancient thinkers praised them. Without the principles of a good will, they can become extremely bad. A villain's coolness makes him not only far more dangerous but also, in our eyes, directly more detestable than we would have thought him without it.

A good will is not good because of what it achieves or accomplishes, or because it is fit to achieve some proposed purpose. It is good only because of its willing—that is, good in itself. Considered on its own, it deserves to be valued incomparably more highly than anything it could ever bring about to satisfy some inclination, or even, if you like, the sum of all inclinations. Suppose that unusually bad luck, or the meager resources supplied by a nature acting like a neglectful stepmother, left this will completely unable to carry out its intention. Suppose that despite its greatest efforts it accomplished nothing, and only the good will remained—not, of course, as a mere wish, but as the use of every means within our power. It would still shine like a jewel, on its own, as something that has its full worth in itself. Usefulness or failure to produce results can neither add to nor subtract from that worth. Usefulness would be only the setting, so to speak, making the jewel easier to handle in everyday dealings or drawing the attention of those who are not yet knowledgeable enough. It would not recommend it to those who know its worth or determine that worth.

Still, there is something so strange in this idea of the absolute worth of a will alone, with no account taken of its usefulness, that a suspicion must arise, even though ordinary reason itself fully agrees with the idea. Perhaps nothing but high-flown fantasy secretly underlies it, and perhaps we have misunderstood nature's purpose in giving our will

395 reason as its governor. So let us examine this idea from that point of view.

When we consider the natural makeup of an organized being—that is, one arranged to serve the purposes of life—we take this as a principle: no instrument for any purpose is found in it unless that instrument is also the best suited and most appropriate to that purpose. Now suppose that, for a being with reason and a will, nature's real purpose were its *preservation*, its *well-being*, in a word, its *happiness*. Then nature would have made a very poor choice in selecting the creature's reason to carry out this intention. Instinct would have laid out all the actions it needed to perform for this purpose, and the whole rule of its conduct, far more precisely. Instinct could have achieved that purpose far more reliably than reason ever can. And if reason had been given to this favored creature as an added gift, it would have had to serve only to reflect on the fortunate arrangement of its nature, to admire and enjoy it, and to be grateful for it to the beneficent cause. It would not have served to subject its ability to desire to that weak and misleading guidance and to tamper with nature's purpose. In short, nature would have prevented reason from breaking out into *practical use* and having the presumption to devise, with its limited insight, its own plan for happiness and the means of reaching it. Nature would have taken on the choice not only of purposes but also of means, and with wise foresight would have entrusted both to instinct alone.

In fact, we also find that the more a cultivated reason concerns itself with enjoying life and achieving happiness, the further a person gets from true contentment. In many people, especially those most experienced in using reason, this produces a certain degree of *misology*, that is, hatred of reason—if only they are honest enough to admit it. They add up all the advantages they gain—I will not say merely from the invention of all the arts of ordinary luxury, but even from the sciences, which in the end also seem to them to be a luxury of the understanding. Yet they find that they have actually burdened themselves with more hardship

396 than they have gained in happiness. In the end they are more likely to envy than to look down on the more ordinary sort of people, who stay closer to the guidance of natural instinct alone and allow their reason little influence on what they do or leave undone. We must at least admit this much: the judgment of those who greatly reduce, or even reduce below zero, the boastful estimates of the advantages reason was supposed to bring us in happiness and contentment with life is by no means sour or ungrateful toward the goodness of the world's governance. Rather, these judgments rest on a hidden idea of another, much worthier purpose for their existence. Reason is properly meant for that purpose, not for happiness. As the highest condition, that purpose must therefore take precedence over people's private aims for the most part.

For reason is not capable enough to guide the will reliably in dealing with its objects and in satisfying all our needs—some of which reason itself multiplies—while an inborn natural instinct would have led far more reliably to that goal. Yet reason has still been given to us as a practical ability, that is, as one meant to influence the *will*. So its true role must be to produce not a will that is good merely as a *means* to some other purpose, but a *will that is good in itself*. For that, reason was absolutely necessary, if nature has indeed acted purposefully throughout in distributing its capacities. So this will need not be the only and the complete good. But it must be the highest good, and the condition of everything else, even of every desire for happiness. In that case, the following observation fits quite well with nature's wisdom. The cultivation of reason needed for the first, unconditional purpose can limit the achievement of the second, always conditional purpose—happiness—in many ways, at least in this life. It can even reduce happiness below zero, without nature acting against its purpose in doing so. For reason recognizes that its highest practical role is to establish a good will. In reaching this goal it is capable only of a contentment of its own kind: contentment from fulfilling a purpose that, again, only reason determines. This holds even if doing so involves many setbacks to the purposes of inclination.

397 Let us now develop the concept of a will that deserves high regard in itself and is good without any further purpose. This concept is already present in natural, sound understanding and needs not so much to be taught as simply to be made clear. It always comes first in assessing the whole worth of our actions and is the condition of everything else. To develop it, we will consider the concept of *duty*. Duty contains the concept of a good will, though under certain subjective limits and obstacles. Far from hiding that good will and making it unrecognizable, these limits and obstacles bring it out by contrast and let it shine all the more brightly.

I pass over here all actions already recognized as contrary to duty, even if they may be useful for one purpose or another. With these, there is no question whether they might have been done *from duty*, since they actually conflict with it. I also set aside actions that really do accord with duty but toward which people have immediately *no inclination*, although they still perform them because another inclination drives them to do so. Here it is easy to tell whether the action that accords with duty was done *from duty* or for a selfish purpose. This distinction is much harder to notice when the action accords with duty and the person also has an *immediate* inclination toward it. For example, a shopkeeper certainly acts in accordance with duty by not overcharging an inexperienced buyer. Where there is a lot of trade, a sensible merchant does not overcharge either. Instead, he keeps a fixed general price for everyone, so that a child buys from him on the same terms as anyone else. People are therefore served honestly. But that is far from enough to make us believe that the merchant acted from duty and principles of honesty. His own advantage required it. We cannot assume here that he also had an immediate inclination toward his customers, so that, out of love as it were, he would give no one a better price than anyone else. So the action was done neither from duty nor from immediate inclination, but simply for personal gain.

By contrast, preserving one's life is a duty, and everyone also has an immediate inclination to do so. But the care, often full of anxiety, that most people take to preserve it still has no intrinsic worth, and the maxim behind it has no moral

398 content. They preserve their lives *in accordance with duty*, but not *from duty*. But suppose misfortunes and hopeless grief have completely taken away someone's taste for life. This unhappy person, strong in spirit and more outraged at his fate than fainthearted or dejected, wishes for death. Yet he preserves his life without loving it, not from inclination or fear but from duty. Then his maxim has moral content.

To help others when one can is a duty. Besides this, some people are so sympathetic by disposition that, even without any further motive of vanity or self-interest, they take an inner pleasure in spreading joy around them. They can delight in other people's contentment insofar as they have brought it about. But I maintain that in such a case an action of this kind, however much it accords with duty and however lovable it is, still has no true moral worth. It ranks alongside other inclinations, such as the inclination toward honor. When that inclination fortunately happens to aim at what really benefits everyone and accords with duty, and is therefore honorable, it deserves praise and encouragement, but not high regard. For the maxim lacks moral content, namely that of doing such actions not from inclination but *from duty*. Suppose, then, that the mind of this friend of humanity was clouded by his own grief, extinguishing all sympathy for the fate of others. He still had the means to help others in distress, but their distress did not move him because he was sufficiently occupied with his own. Now, when no inclination any longer prompted him to do it, suppose he nevertheless tore himself out of this deadly numbness and performed the action without any inclination, solely from duty. Only then does it have its genuine moral worth. More than that: suppose nature had put little sympathy at all into someone's heart. Suppose he was an honest man in other respects, but by temperament cold and indifferent to others' suffering—perhaps because, specially gifted with patience and endurance in facing his own, he assumes or even demands the same in everyone else. If nature had not really made such a man a friend of humanity—and he would certainly not be its worst product—would he not still find within himself a source from which to give himself a far higher worth than a good-natured temperament could have? Certainly! It is precisely here that the worth of character begins. This worth is moral

399 and, beyond all comparison, the highest. It lies in his helping others not from inclination but from duty.

Securing one's own happiness is a duty, at least indirectly. For discontent with one's condition, under the pressure of many worries and amid unsatisfied needs, could easily become a great *temptation to violate one's duties*. But even without considering duty here, all people already have the strongest and deepest inclination toward happiness, because it is precisely in this idea that all inclinations join together into a sum. Yet the prescription for happiness is usually such that it greatly interferes with some inclinations. At the same time, people cannot form a definite, reliable concept of the combined satisfaction of all inclinations under the name of happiness. So it is no surprise that a single inclination, definite in what it promises and in when it can be satisfied, can outweigh an uncertain idea. A person with gout, for example, can choose to enjoy what tastes good to him and suffer what he can bear. By his reckoning, he has at least not deprived himself of the present moment's enjoyment because of possibly groundless expectations of a happiness supposedly found in health. But even in this case, if the general inclination toward happiness did not determine his will, and if health, at least for him, did not figure so necessarily in his reckoning, a law still remains, here as in every other case: to promote his happiness not from inclination but from duty. Only then does his conduct have genuine moral worth.

This is certainly also how we should understand the passages of scripture that command us to love our neighbor, even our enemy. Love as an inclination cannot be commanded. But helping others from duty, even when no inclination at all prompts it, and even when a natural, unconquerable aversion opposes it, is *practical*, not *pathological*, love. It lies in the will, not in a tendency of feeling; in principles of action, not in tender sympathy. Only the former can be commanded.

The second proposition is this: an action from duty has its moral worth *not in the purpose* it is meant to achieve, but in the maxim by which it is decided on. So it does not depend on

400 the actual existence of the action's object, but only on the *principle of willing* by which the action was done, without regard to any objects of the ability to desire. What came before makes this clear: the purposes we may have in acting, and the effects of our actions, as goals and incentives of the will, cannot give actions unconditional, moral worth. Where, then, can this worth lie, if it is not to consist in the will's relation to the hoped-for effect? It can lie nowhere else *but in the principle of the will*, without regard to the purposes such an action can bring about. For the will stands between its a priori principle, which concerns form, and its a posteriori incentive, which concerns content, as if at a crossroads. Since something must determine it, it must be determined by the formal principle of willing in general when an action is done from duty, because every principle based on content has been taken away from it.

I would express the third proposition, as a consequence of the first two, like this: *Duty is the necessity of an action from respect for the law*. I can certainly have an *inclination* toward an object as the effect of my intended action, but *never respect*, precisely because it is only an effect and not an activity of a will. Likewise, I cannot have respect for inclination in general, whether my own or someone else's. At most, I can approve of it in the first case and sometimes even love it in the second—that is, see it as favorable to my own advantage. Only something connected with my will solely as a ground, never as an effect, can be an object of respect and thus a command. It must not serve my inclination but outweigh it, or at least entirely exclude inclination from being counted when a choice is made. Thus only the law by itself can be such an object. Now, an action from duty must completely set aside the influence of inclination, and with it every object of the will. So nothing remains that can determine the will except, objectively, the *law* and, subjectively, *pure respect* for this practical law—and therefore the maxim*) of

*) A *maxim* is the subjective principle of willing. The objective principle—that is, the one that would also serve subjectively as a practical principle for all rational beings if reason had complete power over the ability to desire—is the practical *law*.

401 following such a law even at the expense of all my inclinations.

So an action's moral worth does not lie in the effect expected from it, nor, therefore, in any principle of action that needs to borrow its motive from this expected effect. For all these effects—a pleasant condition for oneself, even the promotion of other people's happiness—could also have been brought about by other causes. They would not have needed the will of a rational being, even though only in that will can the highest, unconditional good be found. Therefore nothing but the *idea of the law* in itself, *which of course occurs only in a rational being*, can make up that outstanding good we call moral, insofar as this idea, not the hoped-for effect, determines the will. This good is already present in the person who acts according to that idea. It must not be expected only from the effect.*)

*) Someone might object that, behind the word *respect*, I am only taking refuge in an obscure feeling instead of giving a clear answer to the question through a concept of reason. But although respect is a feeling, it is not one *received* through an influence. It is a feeling *produced by oneself* through a concept of reason. So it differs in kind from all feelings of the first sort, which can be reduced to inclination or fear. What I directly recognize as a law for myself, I recognize with respect. Respect means only the awareness that my will is *subordinated* to a law, without any other influences on my feelings intervening. The direct determination of the will by the law, and the awareness of it, is called *respect*. So respect is regarded as the *effect* of the law on the subject, not as its *cause*. Strictly speaking, respect is the idea of a worth that diminishes my self-love. It is therefore something regarded as an object neither of inclination nor of fear, though it has something analogous to both at once. The *object* of respect is thus only the *law*, specifically the law that *we impose on ourselves*, yet impose as necessary in itself. As a law, we are subject to it without consulting self-love. As imposed on us by ourselves, it is nevertheless a consequence of our will. In the first respect it is analogous to fear; in the second, to inclination. All respect for a person is really only respect for the law—of integrity, and so on—of which that person gives us an example. Because we also see the development of our talents as a duty, in a talented person we also picture, so to speak, the

example of a law: to become like that person in this respect through practice. This is what makes up our respect. All so-called moral *interest* consists solely in *respect* for the law.

- 402 But what sort of law can this be, whose idea must determine the will, even without regard to the effect expected from it, if the will is to be called good absolutely and without qualification? I have taken away from the will every incentive that could come to it from following any law. So nothing remains but the universal conformity of actions to law as such, which alone is to serve as the will's principle. That is, I must never act except in such a way *that I can also will that my maxim should become a universal law*. Here, then, simply conforming to law as such—without taking as a basis any law directed at particular actions—is what serves, and must serve, as the will's principle, unless duty is to be nothing but an empty delusion and an imaginary concept. Ordinary human reason agrees completely with this in its practical judgments and always keeps this principle in view.

Suppose, for example, the question is this: when I am in a tight spot, may I make a promise with the intention of not keeping it? I can easily distinguish two possible meanings of the question: whether making a false promise is prudent, or whether it accords with duty. The first can certainly often be the case. I do see that it is not enough to get myself out of a present difficulty by this evasion. I must carefully consider whether this lie might later create much greater difficulties for me than those I am now escaping. Despite all my supposed *cleverness*, the consequences are not so easy to foresee that losing people's trust once could not harm me far more than all the trouble I now intend to avoid. So I must consider whether it would not be *more prudent* to follow a general maxim here and make a habit of promising nothing except with the intention of keeping it. But I soon see that such a maxim still rests only on the consequences I fear. Now, being truthful from duty is something quite different from being truthful out of fear of harmful consequences. In the first case, the concept of the action itself already contains a law for me. In the second, I must

first look elsewhere to find what effects might be connected with it that would be unwelcome to me. For if I depart from the principle of duty, that is quite certainly bad. But if I abandon my maxim of

403 prudence, that may sometimes be very advantageous to me, although it is certainly safer to stay with it. Still, there is a very short and yet reliable way to answer the question whether a lying promise accords with duty. I ask myself: would I really be satisfied if my maxim—to get myself out of a difficulty through an untrue promise—were to hold as a universal law for me as well as for others? Could I really say to myself: everyone may make an untrue promise when they are in a difficulty they cannot escape in any other way? I soon realize that I can indeed will the lie, but I cannot will a universal law to lie. For under such a law there would really be no promises at all. It would be pointless to tell others what I claim to intend to do in the future. They would not believe the claim, or, if they rashly did, they would pay me back in the same coin. So my maxim would have to destroy itself as soon as it was made a universal law.

I therefore need no great sharpness of mind to know what I must do for my willing to be morally good. Inexperienced in the ways of the world, unable to be prepared for every event that happens in it, I ask myself only: can you also will that your maxim should become a universal law? If not, it must be rejected. This is not because of any harm it threatens to you or even to others, but because it cannot fit as a principle into a possible system of universal lawmaking. Reason compels me to have immediate respect for such lawmaking. I do not yet *see* what that respect rests on—the philosopher may investigate that—but I at least understand this much: it is an assessment of worth that far outweighs all the worth of anything recommended by inclination. And the necessity of my actions from *pure* respect for the practical law is what makes up duty. Every other motive must give way to duty, because it is the condition of a will good *in itself*, whose worth surpasses everything.

So in the moral knowledge of ordinary human reason we have reached its principle. Admittedly, ordinary reason does not think of it separately in this general form. Yet it always really has it in view and uses it as the standard for its

judgments.

404 It would be easy to show how, with this compass in hand, ordinary reason knows very well how to distinguish in every case that arises what is good, what is bad, what accords with duty, and what is contrary to duty. We need only draw its attention to its own principle, as Socrates did without teaching it anything new at all, and how it follows that no science or philosophy is needed to know what one must do to be honest and good, or even wise and virtuous. We could well have expected this in advance: knowledge of what every person is obliged to do, and therefore also to know, will be within the reach of everyone, even the most ordinary person. Yet here we cannot help admiring how far the ability to judge in practical matters surpasses the ability to judge in theoretical matters in ordinary human understanding. In theoretical matters, when ordinary reason dares to leave the laws of experience and the perceptions of the senses behind, it falls into nothing but incomprehensible ideas and contradictions with itself, or at least into a chaos of uncertainty, obscurity, and instability. In practical matters, by contrast, the power of judgment only really begins to show itself to advantage when ordinary understanding excludes all incentives from the senses from practical laws. It then even becomes subtle. It may try to quibble with its conscience or with other claims about what should count as right; or it may sincerely try to determine the worth of actions for its own instruction. Most importantly, in the latter case it can hope to get things right just as much as any philosopher could. Indeed, it is almost more reliable in this than even the philosopher. For the philosopher can have no other principle than ordinary understanding does, but can easily confuse his judgment with a host of outside considerations that have nothing to do with the issue and lead it away from the straight path. Would it not, then, be wiser to leave moral matters to the judgment of ordinary reason? At most, we could bring in philosophy to present the system of morality more fully and clearly, and its rules in a form more convenient for use, and still more for debate. But we would not use phi-

losophy, even for practical purposes, to lead ordinary human understanding away from its fortunate simplicity and onto a new path of inquiry and instruction.

Innocence is a splendid thing. But it is also

405 very unfortunate that it is hard to preserve and easily led astray. That is why even wisdom—which otherwise consists more in what we do and leave undone than in what we know—still needs science, not to learn from it, but to give its prescriptions acceptance and lasting force. In our needs and inclinations we feel a powerful counterweight to all the commands of duty that reason presents to us as so worthy of respect. We bring together the full satisfaction of these needs and inclinations under the name of happiness. Now reason issues its prescriptions without relenting, yet without promising anything to the inclinations. It thus seems to set aside and disregard their demands, which are so forceful and at the same time seem so reasonable, and which refuse to be canceled by any command. This gives rise to a *natural dialectic*: a tendency to reason sophistically against those strict laws of duty, to cast doubt on their validity, or at least their purity and strictness, and, if possible, to make them better suit our wishes and inclinations. That means, at bottom, corrupting them and robbing them of all their dignity—something that even ordinary practical reason cannot ultimately approve.

So *ordinary human reason* is driven to go beyond its own sphere and take a step into the field of *practical philosophy*. It is not driven by any need for speculation, which never comes over it as long as it is content to be simply sound reason. Rather, it is driven by practical grounds themselves. There it seeks information and clear guidance about the source of its principle and how to determine that principle correctly when set against maxims based on need and inclination. In this way it can escape its difficulty over the claims on both sides, without risking the loss of all genuine moral principles through the ambiguity into which it easily falls. Thus, as ordinary practical reason develops, a *dialectic*

arises unnoticed within it, just as happens in its theoretical use, forcing it to seek help in philosophy. So practical reason, like theoretical reason, will probably find rest nowhere except in a complete critique of our reason.

Section II



406 Second Section

Moving from popular moral philosophy to the metaphysics of morals

We have drawn our concept of duty so far from the ordinary use of our practical reason. But this certainly does not mean that we have treated it as a concept drawn from experience. In fact, when we look at what people do and fail to do, we often hear complaints that we ourselves admit are justified. No reliable examples can be given of a disposition to act purely from duty. Much may happen *in line with* what *duty* commands. But it always remains doubtful whether it really happens *from duty*, and so has moral worth. That is why there have always been philosophers who flatly denied that this disposition exists in human actions. They attributed everything to more or less refined self-love. Yet they did not question the validity of the concept of morality. Instead, they spoke with deep regret about the weakness and impurity of human nature. It is noble enough to make such a worthy idea its rule, but also too weak to follow it. It uses reason, which should give it laws, only to serve the interests of its inclinations, either one at a time or, at best, in the greatest possible harmony with one another.

407 In fact, experience cannot establish even one case with complete certainty in which the maxim of an action in line with duty rested solely on moral grounds and on the thought of one's duty. Sometimes, even after the closest self-exami-

nation, we find nothing besides the moral ground of duty that could have been powerful enough to move us to some good action and such great sacrifice. But this gives us no reliable basis for concluding that no hidden impulse of self-love was the real cause determining the will, with that idea serving as a mere disguise. We like to flatter ourselves by falsely claiming a nobler motive. In fact, even the most strenuous examination never lets us fully uncover our secret incentives. For moral worth does not depend on the actions we can see, but on their inner principles, which we cannot see.

Some people laugh at all morality as a mere fantasy of the human imagination, which overreaches itself through conceit. We could do them no greater favor than to concede that concepts of duty must be drawn only from experience. For convenience, we also like to persuade ourselves that the same holds for all other concepts. But this concession would guarantee those people a triumph. Out of love for humanity, I am willing to admit that most of our actions are in line with duty. But look more closely at people's thoughts and aims, and everywhere you meet the dear self, always pushing to the front. Their purpose rests on this, not on the strict command of duty, which would often require self-denial. You do not have to be an enemy of virtue to doubt at certain moments whether any true virtue is actually found in the world. You only need to be a calm observer who does not immediately mistake the strongest wish for goodness for its actual existence. This is especially so as you grow older and experience makes your judgment both wiser and sharper at observing. At this point, only one thing can keep us from abandoning our ideas of duty altogether and preserve a well-grounded respect for its law in our souls: the clear conviction that even if there had never been actions arising from such pure sources,

408 that is not the issue here at all. The issue is not whether this or that happens. Rather, reason, by itself and independently of all appearances, commands what ought to happen. So reason still unrelentingly commands actions of which the world may never yet have supplied an example, and whose very possibility someone who bases everything on experience might seriously doubt. For example, we can demand pure honesty in friendship from everyone just as firmly

even if there has never yet been an honest friend. For this duty, as duty in general, lies before all experience in the idea of a reason that determines the will on grounds independent of experience.

There is a further point. Unless we want to deny the concept of morality any truth or any relation to a possible object at all, we cannot deny that its law must have a very broad scope. It must hold not just for human beings, but for *all rational beings in general*, and not just under contingent conditions and with exceptions, but *with absolute necessity*. Clearly, no experience can give us grounds to infer even the possibility of such unconditionally necessary laws. What right have we to demand unlimited respect from every rational nature for something that may be valid only under the contingent conditions of being human? And how could laws determining *our* will count as laws determining the will of a rational being in general, and only as such also count for ours, if they were merely based on experience? They must originate entirely independently of experience in pure but practical reason.

Nor could we give morality worse advice than to try to borrow it from examples. Every example of morality presented to me must itself first be judged by moral principles. We must determine whether it deserves to serve as an original example, that is, a model. It can never be the primary source of the concept of morality. Even the Holy One of the Gospel must first be compared with our ideal of moral perfection before we recognize him as such. He also says of himself: "Why do you call me (whom you see) good? No one is good (the original model of goodness) except the one God (whom you do not see)." But where do we get the concept of God

409 as the highest good? Only from the *idea* of moral perfection that reason forms independently of experience and links inseparably with the concept of a free will. Imitation has no place at all in morality. Examples serve only to encourage us. That is, they remove doubt about whether we can do what the law commands. They show concretely what the practical rule expresses in more general terms. But they can never justify setting aside the true original, which lies in reason, and letting examples guide us instead.

There can be no genuine supreme principle of morality unless it rests on pure reason alone, independently of all experience. So I think we need not even ask whether it is good to present these concepts in general terms (in the abstract), as they and their associated principles stand established independently of experience. We must do this if our knowledge is to differ from ordinary knowledge and deserve the name philosophical. Yet in our time it may well be necessary to ask. Suppose we collected votes on whether to prefer pure rational knowledge, separated from everything based on experience—and so a metaphysics of morals—or popular practical philosophy. We can easily guess which side would win.

Bringing the subject down to popular concepts certainly deserves praise, provided we have first risen to the principles of pure reason and reached them with complete success. That would mean first grounding moral teaching in metaphysics, and then, once it is secure, making it popular so it can *gain acceptance*. But it is utterly unreasonable to try to satisfy this demand for popularity in the first investigation, on which the soundness of all the principles depends. This procedure can never claim the extremely rare achievement of true *philosophical popularity*. There is no skill in being understandable to everyone if you give up all thorough insight in the process. Worse, it produces a repulsive jumble of patched-together observations and half-reasoned principles. Shallow minds enjoy this, since it is very useful for everyday chatter. But people with insight feel confused and turn their eyes away, dissatisfied yet unable to help themselves. Meanwhile, philosophers who see perfectly well through the illusion find few listeners when they ask people to turn away for a while from this supposed

410 popularity. First we must gain definite insight; only then can we rightly make the subject popular.

Just look at attempts to discuss morality in this fashionable style. You soon find a strange mixture: here the special purpose of human nature (but sometimes also the idea of a rational nature in general), there perfection or happiness; here moral feeling, there fear of God; a little of this and a little of that. No one thinks to ask whether principles of morality should be sought in knowledge of

human nature at all, since that knowledge can come to us only from experience. Suppose they should not. Suppose these principles are found entirely independently of experience, free of everything empirical, in pure concepts of reason alone, and nowhere else—not even in the smallest part. Then we should decide to separate this investigation completely, as pure practical philosophy or, if we may use such a discredited name, as a metaphysics*) of morals. We should bring it to full completion on its own, and ask the public that demands popularity to wait until this undertaking is finished.

Such a completely isolated metaphysics of morals must be mixed with no anthropology, theology, physics, or hyperphysics, and still less with hidden qualities, which one might call hypophysical. It is not only an indispensable basis for all theoretical, clearly established knowledge of duties. It is also something of the highest importance if their rules are actually to be carried out. For the pure thought of duty, and of the moral law in general, without any foreign addition of incentives from experience, has a far more powerful influence on the human heart through reason alone (which here first becomes aware that it can also be practical on its own) than all other incentives*) that

*) If we wish, we can distinguish pure moral philosophy (metaphysics) from applied moral philosophy (applied, that is, to human nature), just as we distinguish pure from applied mathematics and pure from applied logic. This name also immediately reminds us that moral principles must not be based on the special features of human nature. They must stand on their own, independently of experience. But from them we must be able to derive practical rules for human nature, just as for every rational nature.

411 we might muster from the field of experience. This influence is so powerful that reason, conscious of its dignity, can scorn those other incentives and gradually master them. By contrast, a mixed moral teaching, made up of incentives from feelings and inclinations as well as concepts of reason, must leave the mind wavering between motives that cannot be brought under any single principle. They can lead to good only by chance, and often lead to evil as well.

What we have said makes the following clear. All moral concepts have their place and origin entirely in reason, independently of experience. This holds for the most ordinary human reason just as much as for the most highly theoretical reason. These concepts cannot be abstracted from any knowledge based on experience, which for that very reason is merely contingent. Their dignity, which makes them fit to serve us as supreme practical principles, lies precisely in this purity of their origin. Every addition of something empirical takes away just as much from their genuine influence and from the unconditional worth of actions. Drawing their concepts and laws from pure reason, and presenting them pure and unmixed, is not just an absolute necessity for theoretical purposes, when only abstract inquiry is at issue. It is also of the greatest practical importance. We must even determine the scope of this whole body of practical or pure rational knowledge—that is, the whole capacity of pure practical reason. In doing so, however, we must not—as theoretical philosophy permits and sometimes even finds necessary—make its principles depend on the particular

*) I have a letter from the late, excellent *Sulzer*, in which he asks me what the reason might be that teachings about virtue achieve so little, despite all their power to convince reason. My preparations to give him a complete answer delayed my reply. But the answer is simply that the teachers themselves have not made their concepts pure. They want to do too much. They hunt everywhere for motives to moral goodness to make the medicine really strong, and so spoil it. The most ordinary observation shows this. Present an honest action performed with a steadfast spirit, even under the greatest temptations of hardship or enticement, and separated from every thought of advantage in this world or another. It leaves far behind and outshines every similar action affected even slightly by a foreign incentive. It lifts the soul and awakens the wish to be able to act that way too. Even children in their middle years feel this impression, and duties should never be presented to them in any other way.

⁴¹² nature of human reason. Moral laws are supposed to hold for every rational being in general. We must therefore derive them from the general concept of a rational being as such. In this way we must first present morality completely,

independently of anthropology, as pure philosophy—that is, as metaphysics—even though morality needs anthropology for its *application* to human beings. A complete presentation is quite possible in this kind of wholly separate knowledge. We must be fully aware that without possessing this metaphysics, we cannot accurately determine, for theoretical judgment, what makes duty moral in everything that conforms to duty. More than that: even in ordinary practical use, especially in moral instruction, it is impossible to ground morals on their genuine principles without it. We cannot thereby produce pure moral dispositions and plant them in people’s minds for the greatest good of the world.

In this treatment, however, we want to advance through the natural stages beyond what we did before. We are not simply moving from ordinary moral judgment—which here deserves great respect—to philosophical judgment. We are moving from a popular philosophy, which can go no farther than it can grope its way by means of examples, all the way to metaphysics. Metaphysics lets nothing empirical hold it back. Since it must measure out the whole range of rational knowledge of this kind, it may go as far as ideas where even examples abandon us. To make this advance, we must trace and clearly present the practical power of reason, from its general rules for determining action to the point where the concept of duty arises from it.

Everything in nature works according to laws. Only a rational being can act *according to the thought* of laws—that is, according to principles—or has a *will*. Since deriving actions from laws requires *reason*, the will is nothing other than practical reason. If reason invariably determines the will, the actions of such a being that are recognized as objectively necessary are also subjectively necessary. That is, the will is a capacity to choose *only what* reason, independently of inclination, recognizes as practically necessary, or good. But suppose reason alone does not sufficiently determine the will. Suppose it is still subject to subjective conditions—certain incentives—that do not always agree with the objective ones. In

413 short, suppose the will is not *in itself* fully in line with reason, as is actually the case with human beings. Then actions recognized as objectively necessary are subjectively contingent. Determining such a will according to objective laws is *constraint*. That is, the relation of objective laws to a will that is not thoroughly good is thought of as determining the will of a rational being through reasons, although this will does not, by its nature, necessarily follow those reasons. The thought of an objective principle, insofar as it constrains a will, is called a command of reason. The formula of the command is called an *imperative*. All imperatives are expressed with an *ought*. In this way they indicate the relation of an objective law of reason to a will that is not necessarily determined by it because of its subjective makeup—a constraint. They say that doing or not doing something would be good. But they say this to a will that does not always do something simply because it is presented as good to do. What is practically *good* determines the will through the ideas of reason. It does so not through subjective causes, but objectively: through reasons valid for every rational being as such. It differs from what is *pleasant*. The pleasant influences the will only through sensation, from merely subjective causes that hold only for this or that person's senses, not as a principle of reason valid for everyone.*)

*) The dependence of the capacity for desire on sensations is called inclination, and so inclination always indicates a *need*. The dependence on principles of reason of a will that can be contingently determined is called an *interest*. Interest therefore exists only in a dependent will, which does not always conform to reason of its own accord. We cannot think of the divine will as having an interest. But even the human will can *take an interest* in something without *acting from interest*. The first means a *practical* interest in the action; the second, an interest *based on feeling* in the object of the action. The first indicates only the will's dependence on principles of reason in themselves. The second indicates its dependence on those principles in the service of inclination: reason merely supplies the practical rule for meeting the need of inclination. In the first case, the action interests me; in the second, the object of the action does, insofar as it is

pleasant to me. In the first section we saw that an action from duty must be considered in terms not of interest in its object, but only of interest in the action itself and its principle in reason, the law.

- 414 A perfectly good will would therefore also be subject to objective laws of the good. But we could not think of it as *constrained* by them to act in accordance with law. By its own subjective makeup, it can be determined only by the thought of the good. So no imperatives hold for the *divine* will, or for any *holy* will. *Ought* is out of place here, because *willing* already necessarily agrees with the law of its own accord. Imperatives are therefore only formulas expressing the relation between objective laws of willing in general and the subjective imperfection of the will of this or that rational being—for example, the human will.

All *imperatives* command either *hypothetically* or *categorically*. Hypothetical imperatives present a possible action as practically necessary as a means of reaching something else that we will, or at least might will. A categorical imperative would present an action as objectively necessary in itself, without reference to any other end.

Every practical law presents a possible action as good, and therefore as necessary for a subject whose actions can be determined by reason. So all imperatives are formulas for determining the action that is necessary according to the principle of a will that is good in some way. If the action would be good only as a *means* to something else, the imperative is *hypothetical*. If the action is presented as good *in itself*, and therefore as necessary in a will that in itself conforms to reason, as that will's principle, the imperative is *categorical*.

The imperative thus says which action within my power would be good. It presents the practical rule in relation to a will that does not immediately act simply because the action is good. This is partly because the subject does not always know that it is good. It is also partly because, even if the subject did know this, the subject's maxims might still conflict with the objective principles of practical reason.

The hypothetical imperative therefore says only that the action is good for some *possible* or *actual* aim. In the first

415 case it is a *practical principle for a possible goal*; in the second, a *practical principle for an actual goal*. The categorical imperative declares the action objectively necessary in itself, without reference to any aim or any other end. It counts as an *unconditionally necessary practical principle*.

Whatever is possible through the powers of any rational being can also be thought of as a possible aim for some will. So there are infinitely many principles of action that present an action as necessary for achieving a possible aim it can bring about. All sciences have some practical part. It consists of problems involving an end that is possible for us, and imperatives stating how to reach it. These can therefore be called imperatives of *skill* in general. Here the question is not at all whether the end is rational and good, but only what we must do to reach it. A doctor's instructions for making a person thoroughly healthy and a poisoner's instructions for killing that person reliably have equal worth in this respect: each serves to achieve its aim completely. In early youth we do not know what ends we may encounter in life. So parents mainly try to have their children learn *many different things*. They take care to develop their children's *skill* in using means for all sorts of *freely chosen* ends. They cannot tell whether any one of these will actually become their child's aim in the future, though the child might have it someday. This concern is so great that they generally neglect to develop and correct their children's judgment about the worth of the things they might choose as ends.

Still, there is *one* end that we can assume is actual for all rational beings insofar as imperatives apply to them—that is, as dependent beings. This is an aim that they do not merely *possibly* have. We can safely assume that they all have it by a natural necessity: the aim of *happiness*. The hypothetical imperative that presents an action as practically necessary as a means to promote happiness concerns an *actual goal*. It must be presented as necessary not just for an uncertain, merely possible aim, but for an aim we can safely assume in every human being independently of experience,

416 because it belongs to human nature. Now skill in choosing the means to one's own greatest well-being can be called *prudence* *) in the narrowest sense. So the imperative that concerns choosing means to one's own happiness—that is, the rule of prudence—is still *hypothetical*. It commands the action not without qualification, but only as a means to another aim.

Finally, there is an imperative that commands a kind of conduct directly, without making any other aim achievable through that conduct its underlying condition. This imperative is *categorical*. It concerns not the content of the action and what is to result from it, but the form and the principle from which the action itself follows. What is essentially good in the action lies in the disposition, whatever the outcome may be. We may call this the imperative of *morality*.

Willing according to these three kinds of principles is also clearly distinguished by the *differences* in how the will is constrained. To make this difference noticeable too, I think the most suitable names, in their order, would be *rules* of skill, *advice* of prudence, and *commands* (laws) of morality. Only a *law* carries the concept of an *unconditional necessity* that is objective and therefore universally valid. Commands are laws that must be obeyed—that is, followed even against inclination. *Advice* does involve necessity, but a necessity that can hold only under a subjective, contingent condition: whether this or that person counts this or that as part of their happiness. By contrast, no condition limits the categorical imperative. As absolutely necessary, although practically so, it can quite properly be called a command. We could also call the first kind of imperative *technical* (belonging to an art), the

*) The word prudence is used in two senses. In the first it may be called worldly prudence; in the second, personal prudence. The first is the skill of influencing other people so as to use them for one's own aims. The second is the insight needed to unite all these aims for one's own lasting advantage. The second is really what gives even the first its value. Someone who is prudent in the first way but not in the second might better be called clever and crafty, but still imprudent overall.

417 second *pragmatic* *) (concerning well-being), and the third *moral* (belonging to free conduct in general, that is, to morals).

The question now arises: how are all these imperatives possible? This does not ask how the action commanded by the imperative can be carried out. It asks only how we can think of the constraint on the will that the imperative expresses in setting the task. How an imperative of skill is possible hardly needs a separate discussion. Whoever wills the end also wills the indispensably necessary means to it that are within their power, insofar as reason has a decisive influence on their actions. As far as willing is concerned, this statement is analytic. For when I will an object as an effect of my action, I already think of my causality as an acting cause—that is, my use of means. The imperative draws the concept of the actions necessary for this end out of the very concept of willing this end. Determining the means themselves for an intended aim does require synthetic statements. But these do not concern the basis for making the act of will actual; they concern making its object actual. For example, to divide a line into two equal parts by a reliable principle, I must draw two intersecting arcs from its endpoints. Mathematics teaches this only through synthetic statements. But suppose I know that only this action can bring about the intended effect. Then, if I fully will the effect, I also will the action required for it. This is an analytic statement. For thinking of something as an effect I can bring about in a certain way is exactly the same as thinking of myself as acting in that way to bring it about.

If it were only easy to give a definite concept of happiness, imperatives of prudence would agree entirely with those of skill and would be just as analytic. Here as there, we would say: whoever wills the end also wills—necessarily, in accordance with reason—the only means to it within

*) It seems to me that this is the most precise way to define the proper meaning of the word *pragmatic*. For *sanctions* are called pragmatic when they arise not strictly from the rights of states as necessary laws, but from *concern* for the general welfare. A *history* is written *pragmatically* when it makes people *prudent*, that is, teaches the world how to look after its advantage better than earlier generations did, or at least just as well.

418 their power. Unfortunately, though, the concept of happiness is so indefinite that everyone wishes to reach it, yet no one can ever say definitely and consistently what they really wish and will. The reason is this: all the elements that belong to the concept of happiness come from experience. Yet the idea of happiness requires an absolute whole, a maximum of well-being in my present condition and every future condition. Now even the most insightful and powerful being, if still finite, cannot form a definite concept of what it really wills here. Does someone will wealth? How much worry, envy, and persecution this might bring down on them! Do they will extensive knowledge and insight? Perhaps this would only sharpen their sight, showing them all the more terribly the evils that are still hidden from them and yet cannot be avoided. Or it might burden their desires, which already give them enough trouble, with still more needs. Do they will a long life? Who can guarantee that it would not be a long misery? Do they will at least health? How often bodily discomfort has kept someone from excesses that unlimited health would have let them fall into, and so on. In short, no one can use any principle to determine with complete certainty what will make them truly happy. That would require knowing everything. So we cannot act on definite principles in order to be happy. We can act only on advice drawn from experience—for example, about diet, thrift, courtesy, restraint, and so on—which experience teaches generally does the most to promote well-being. It follows that, strictly speaking, imperatives of prudence cannot command at all. They cannot present actions objectively as practically *necessary*. They should be seen as recommendations (*consilia*) rather than commands (*praecepta*) of reason. The task of determining reliably and universally which action will promote a rational being's happiness is completely unsolvable. So no imperative about happiness is possible that commands, in the strict sense, doing what makes us happy. Happiness is not an ideal of reason but of the imagination. It rests only on grounds drawn from experience, from which it is hopeless

419 to expect a determination of an action that would achieve the whole of what is actually an infinite series of consequences. Still, if we assume that the means to happiness could be reliably specified, this imperative of prudence would be an analytic practical statement. It differs from the imperative of skill only in this: for skill the end is merely possible, while for prudence it is given. Both command only the means to something that we assume we will as an end. So in both cases, the imperative that commands someone who wills the end to will the means is analytic. Thus there is also no difficulty about the possibility of such an imperative.

By contrast, how the imperative of *morality* is possible is undoubtedly the only question that needs solving. It is not hypothetical at all. So the necessity it presents as objective cannot rest on any presupposition, as it can with hypothetical imperatives. Here, however, we must always keep one thing in mind. *No example*, and therefore no experience, can establish whether any such imperative exists at all. We must worry that all those that seem categorical may secretly be hypothetical. For example, take the command: "You ought not to make deceitful promises." Suppose we assume that the necessity of refraining from this is not just advice for avoiding some other evil. It does not mean, for instance: "You ought not to make lying promises, so that you do not lose your credit when the lie comes to light." Instead, an action of this kind must be regarded as bad in itself, and so the imperative forbidding it is categorical. Still, no example can show with certainty that the will is determined here solely by the law, without any other incentive, even if that is how it seems. Fear of disgrace, or perhaps a vague worry about other dangers, may always secretly influence the will. Who can prove through experience that a cause does not exist, when experience teaches us nothing more than that we do not perceive it? In such a case, the so-called moral imperative, which as such appears categorical and unconditional, would in fact be only a pragmatic rule. It would draw our attention to our advantage and teach us only to look after it. So we must investigate the possibility of a *categorical* imperative entirely independently of experience. Here we do not have the advantage

420 of having its actual existence given in experience. If we did, we would need to consider its possibility only to explain it, not to establish it. For now, though, we can see this much: only the categorical imperative has the force of a practical *law*. All the others can be called *principles* of the will, but not laws. What is necessary to do merely to reach an end of our choice can be regarded as contingent in itself. We can always free ourselves from the rule by giving up the aim. By contrast, the unconditional command leaves the will no choice to do the opposite. So it alone carries the kind of necessity we require in a law.

Second, there is also a very great source of difficulty in understanding the possibility of this categorical imperative, or law of morality. It is a synthetic practical statement*) independent of experience. Understanding the possibility of statements of this kind presents so much difficulty in theoretical knowledge that we can easily judge it will present no less in practical knowledge.

In tackling this problem, let us first try to see whether the mere concept of a categorical imperative might also supply its formula. This would contain the statement that alone can be a categorical imperative. For even once we know what such an absolute command says, understanding how it is possible will still require a separate, difficult effort. We will put that off until the last section.

When I think of a *hypothetical* imperative in general, I do not know in advance what it will contain until its condition is given to me. But when I think of a *categorical* imperative, I know at once what it contains. For besides the law, the imperative contains only the necessity that the maxim**) conform to this law,

*) Without presupposing a condition based on any inclination, I connect the deed with the will independently of experience, and therefore necessarily. This necessity is only objective, however: it falls under the idea of a reason that had complete power over all subjective motives. So this is a practical statement that does not analytically derive willing one action from another willing already presupposed—for we do not have such a perfect will. Instead, it connects willing an action directly with the concept of the will of a rational being, as something not contained in that concept.

**) A *maxim* is the subjective principle of acting. It must be distinguished from the *objective principle*, namely, the practical law. A maxim contains the practical rule that reason determines in accordance with the subject's conditions, often the subject's ignorance or inclinations. It is therefore the principle according to which the subject *acts*. The law, however, is the objective principle, valid for every rational being, and the principle according to which the subject *ought to act*—that is, an imperative.

421 and the law contains no condition that limits it. So nothing remains but the universality of a law in general, to which the maxim of the action must conform. This conformity alone is what the imperative actually presents as necessary.

There is therefore just one categorical imperative: *Act only on a maxim that you can at the same time will to become a universal law.*

Now suppose all imperatives of duty can be derived from this one imperative as their principle. Then at least we will be able to show what we mean by duty and what this concept is meant to express, even if we leave unsettled whether what we call duty might be an empty concept altogether.

The universality of the law according to which effects occur is what constitutes *nature* in the broadest sense, in terms of its form. That is, nature means the existence of things insofar as it is determined by universal laws. So the universal imperative of duty could also read: *Act as if the maxim of your action were to become, through your will, a universal law of nature.*

Let us now list some duties, following the usual division into duties to ourselves and to other people, and into perfect and imperfect duties.*)

1) Someone is weary of life because of a series of evils that have grown to the point of hopelessness. This person still

*) It is important to note here that I reserve the division of duties entirely for a future *Metaphysics of Morals*. The division here is therefore only a freely chosen one, to put my examples in order. Here, by a perfect duty, I mean one that permits no exception in the interest of inclination. In this sense I have not only

external but also internal *perfect duties*. This runs against the usage accepted in the schools. But I do not intend to defend it here, since it makes no difference to my aim whether it is granted to me or not.

422 has enough use of reason to ask whether taking their own life might be contrary to duty to themselves. They now test whether the maxim of their action could become a universal law of nature. Their maxim is: "From self-love, I make it my principle to shorten my life when its continuation threatens more evil than it promises pleasure." The only remaining question is whether this principle of self-love could become a universal law of nature. We soon see that a nature whose law was to destroy life through the very feeling whose purpose is to drive us to promote life would contradict itself. It could not exist as nature. So that maxim cannot possibly hold as a universal law of nature, and consequently conflicts completely with the supreme principle of all duty.

2) Someone else is driven by need to borrow money. They know very well that they will not be able to repay it. But they also see that no one will lend them anything unless they firmly promise to repay it at a specified time. They want to make such a promise. But they still have enough conscience to ask: "Isn't it forbidden and contrary to duty to get out of difficulty in this way?" Suppose they decided to do it anyway. Their maxim of action would be: "When I believe I need money, I will borrow money and promise to repay it, even though I know that will never happen." Now this principle of self-love or personal advantage might well be compatible with my whole future well-being. But the question now is whether it is right. So I turn the demand of self-love into a universal law and put the question this way: "What would happen if my maxim became a universal law?" I immediately see that it could never hold as a universal law of nature and agree with itself. It would necessarily contradict itself. For a universal law allowing everyone who believes they are in need to promise whatever they please, intending not to keep the promise, would make promising itself, and whatever end one might have in it, impossible. No one would believe that anything had been promised to them. They would laugh at all such statements as empty pretense.

3) A third person finds in themselves a talent that, with some

423 development, could make them useful in all sorts of ways. But they find themselves in comfortable circumstances and prefer to pursue pleasure rather than work at extending and improving their natural gifts. Still, they ask whether their maxim of neglecting these gifts agrees not only with their tendency to seek amusement, but also with what is called duty. They see that nature could still exist under such a universal law. Human beings could let their talents rust, like the inhabitants of the South Seas, and intend to spend their lives only in idleness, amusement, and reproduction—in a word, enjoyment. But they cannot possibly *will* this to become a universal law of nature, or to be implanted in us as such a law through natural instinct. For as a rational being this person necessarily wills all their abilities to be developed, since these serve them, and are given to them, for all sorts of possible aims.

A *fourth* person is doing well but sees others struggling with great hardships, whom they could also help. They think: “What has that to do with me? Let everyone be as happy as heaven wills, or as they can make themselves. I will take nothing from them, and will not even envy them. But I have no wish to contribute anything to their well-being or to help them in need!” Now if this way of thinking became a universal law of nature, the human race could certainly still exist. It would doubtless do better than if everyone chattered about sympathy and goodwill, and even eagerly practiced them on occasion, yet also cheated wherever possible, sold people’s rights, or otherwise damaged them. But although a universal law of nature could well exist according to that maxim, it is impossible to *will* such a principle to hold everywhere as a law of nature. For a will that decided this would conflict with itself. Many cases could arise in which the person needs the love and sympathy of others. Through such a law of nature, arising from their own will, they would rob themselves of all hope of the help they wish for.

These, then, are some of the many actual duties, or at least what we take to be duties. It is clear how their division follows from the one principle

424 stated above. We must *be able to will* that a maxim of our action become a universal law. This is the standard for judging the morality of the action in general. Some actions are such that their maxim cannot even be thought of as a universal law of nature without contradiction. Still less could we *will* that it *should* become one. Other actions do not involve this inner impossibility. But it is still impossible to will their maxim to be raised to the universality of a law of nature, because such a will would contradict itself. We can easily see that the first kind conflicts with strict or narrow duty, which allows no relaxation. The second conflicts only with broad duty, which carries merit. In this way, these examples fully present all duties in their dependence on the one principle, as regards the kind of obligation involved, not the object of their action.

Now whenever we break a duty, if we pay attention to ourselves, we find that we do not really will our maxim to become a universal law. That is impossible for us. Instead, we will that its opposite remain a universal law. We merely take the liberty of making an *exception* for ourselves, or even just for this occasion, in favor of our inclination. So if we considered everything from one and the same standpoint, namely reason, we would find a contradiction in our own will. A certain principle would be objectively necessary as a universal law, yet subjectively it should not hold universally but should allow exceptions. But we consider our action at one moment from the standpoint of a will entirely in line with reason, and at another from the standpoint of a will affected by inclination. So there is no actual contradiction here. There is, however, resistance by inclination to the rule of reason (*antagonismus*). This turns the principle's strict universality (*universalitas*) into mere general validity (*generalitas*), so that the practical principle of reason is supposed to meet the maxim halfway. Although our own impartial judgment cannot justify this, it still proves that we really do acknowledge the validity of the categorical imperative. With all our respect for it, we merely allow ourselves a few exceptions that seem to us insignificant and forced upon us.

425 We have at least shown this much: if duty is a concept that is to have meaning and give real laws for our actions, it can be expressed only in categorical imperatives, never in hypothetical ones. We have also clearly presented the content of the categorical imperative and made it definite enough for any use. This is already a great deal, since that imperative would have to contain the principle of all duty, if there is any such thing at all. But we have not yet reached the point of proving a priori that such an imperative actually exists: that there is a practical law that commands absolutely, on its own and without any incentives, and that following this law is a duty.

As we try to reach that point, one warning is extremely important. We must never think of trying to derive the reality of this principle from the *particular character of human nature*. Duty is supposed to be the practically unconditional necessity of an action. It must therefore hold for all rational beings to whom an imperative can apply at all. *Only for that reason* must it also be a law for every human will. By contrast, anything derived from humanity's particular natural constitution, from certain feelings and tendencies, or even, if possible, from a particular direction taken by human reason that need not hold for the will of every rational being, can give us a maxim, but not a law. It can give us a subjective principle, according to which we have a tendency and inclination to be allowed to act. It cannot give us an objective principle according to which we would be *required* to act even if all our tendencies, inclinations, and natural constitution opposed it. Indeed, the fewer subjective causes there are in favor of a duty, and the more there are against it, the more this shows the command's sublimity and inner dignity. Yet this does not in the least weaken the constraint imposed by the law or take anything away from its validity.

Here we see philosophy placed in a precarious position. It is supposed to stand firm, although it hangs from nothing in heaven and rests on nothing on earth. Here it must prove its purity as the keeper of its own laws, not as the herald of laws whispered to it by an implanted sense or by who knows what guardian nature. All such laws, however much better they may be

426 than nothing at all, can never supply the principles reason dictates. These must have their source entirely a priori, and with it their authority to command. They must expect nothing from human inclination, but everything from the supreme power of the law and the respect owed to it. Otherwise, they must condemn human beings to self-contempt and inward disgust.

Anything empirical is therefore not merely wholly unfit to be added to the principle of morality. It is also highly damaging to the purity of morals themselves. For the true worth of an absolutely good will, a worth beyond all price, consists precisely in this: the principle of its action is free from every influence of contingent grounds that only experience can supply. We cannot give too many or too frequent warnings against this carelessness, or even base way of thinking, that looks for the principle among empirical motives and laws. When human reason grows tired, it likes to rest on this cushion. Dreaming of sweet illusions, which make it embrace a cloud instead of Juno, it substitutes for morality a bastard patched together from parts of quite different origins. This looks like anything one wants to see in it—except virtue, to anyone who has once seen virtue in its true form.*)

So the question is this: is it a necessary law for *all rational beings* always to judge their actions by maxims that they can themselves will to serve as universal laws? If it is, then it must already be connected, entirely a priori, with the concept of the will of a rational being in general. To discover this connection, however much we resist, we must take a step beyond our present ground into metaphysics. We enter a part of it distinct from speculative philosophy: the metaphysics of morals.

*) To see virtue in its true form is simply to present morality stripped of every admixture of what belongs to the senses and of all the false decoration of reward or self-love. Anyone whose reason has not been completely spoiled for all abstraction can easily discover, with the slightest attempt, how much virtue then outshines everything else that appeals to the inclinations.

427 In a practical philosophy we are not concerned with accepting grounds for what *happens*. We are concerned with laws for what *ought to happen*, even if it never happens, that is, with objectively practical laws. So here we do not need to investigate the grounds for why something pleases or displeases. We do not need to ask how the pleasure of mere sensation differs from taste, or whether taste differs from a general satisfaction of reason. We do not need to ask what feeling of pleasure and displeasure rests on, or how desires and inclinations arise from it, and how maxims then arise from those through the cooperation of reason. All of that belongs to an empirical psychology, and an empirical psychology would form the second part of the doctrine of nature if it were treated as a *philosophy of nature*, insofar as that is founded on *empirical laws*. Here, though, we are speaking of the objectively practical law, and so of the relation of a will to itself, insofar as the will determines itself purely through reason. Anything that has a reference to the empirical then drops out by itself. For if *reason by itself alone* determines the conduct (and we now want to examine the possibility of exactly this), it must necessarily do so a priori.

The will is thought as a capacity to determine itself to act *in accordance with the representation of certain laws*. Such a capacity can be found only in rational beings. Now what serves the will as the objective ground of its self-determination is the *end*. When this end is given through reason alone, it must count equally for all rational beings. What instead contains only the ground of the possibility of the action whose effect is the end is called the *means*. The subjective ground of desire is the *incentive*, and the objective ground of willing is the *motive*. This gives the difference between subjective ends, which rest on incentives, and objective ends, which depend on motives that hold for every rational being. Practical principles are *formal* when they abstract from all subjective ends. They are *material* when they put these, and so certain incentives, at the ground. The ends that a rational being sets himself at will as *effects* of his action (material ends) are all only relative. Only their relation to a specially constituted faculty of desire in the subject gives them their worth. That worth is therefore never general, valid for all rational beings and for every act of willing, and necessary

428 principles, that is, practical laws, that it can offer. So all these relative ends are only the basis of hypothetical imperatives.

But suppose there were something whose *existence in itself* has an absolute value, something that could be the basis of definite laws as an *end in itself*. Then the basis of a possible categorical imperative, that is, of a practical law, would lie in that thing, and in it alone.

Now I say this: a human being, and in general every rational being, *exists* as an end in itself, *not merely as a means* to be used as this or that will pleases. In all of their actions, whether directed at themselves or at other rational beings, they must always be regarded *at the same time as an end*. All objects of inclinations have only a conditional value. If there were no inclinations, and no needs based on them, their object would have no value. The inclinations themselves, as the sources of need, have so little absolute value that nobody could wish to have them for their own sake. On the contrary, to be entirely free of them must be the universal wish of every rational being. So the value of all objects *to be acquired* through our action is always conditional. Beings whose existence rests not on our will but on nature still have only a relative value, as means, if they are beings without reason. That is why they are called *things*. Rational beings, on the other hand, are called *persons*, because their nature already marks them out as ends in themselves, that is, as something that must not be used merely as a means. And so, to that extent, it limits all arbitrary choice (and is an object of respect). These, then, are not merely subjective ends, whose existence, as an effect of our action, has a value for us. They are *objective ends*, that is, things whose existence is itself an end. And it is an end of such a kind that no other end can be put in its place, an end that these things would then serve *merely* as means. For without this, nothing of *absolute value* would be found anywhere at all. But if all value were conditional, and therefore contingent, no supreme practical principle could be found anywhere for reason.

So if there is to be a supreme practical principle, and, with regard to the human will, a categorical imperative, it must be one that comes from the idea of what is necessarily an end for everyone because it is an end in itself, and that forms an *objective*

429 principle of the will, and so can serve as a universal practical law. The basis of this principle is: *rational nature exists as an end in itself*. This is how a human being necessarily thinks of their own existence. To that extent it is a *subjective* principle of human actions. But every other rational being also thinks of their existence in this way, on the very same basis in reason that holds for me*). So it is at the same time an *objective* principle, from which, as a supreme practical basis, all laws of the will must be able to be derived. The practical imperative will therefore be this: *Act so that you use humanity, whether in your own person or in the person of anyone else, always at the same time as an end, never merely as a means*. Let us see whether this can be carried out.

To stay with the earlier examples:

First, by the concept of necessary duty to oneself, a person who is considering suicide will ask themselves whether their action can be compatible with the idea of humanity *as an end in itself*. If they destroy themselves to escape a burdensome condition, they are using a person merely *as a means* to keep up a bearable condition until the end of life. But a human being is not a thing. So a human being is not something that can be used *merely* as a means. In all their actions, they must always be regarded as an end in itself. So I cannot dispose of the human being in my own person by maiming, ruining, or killing them. (I must pass over here the closer definition of this principle that would prevent all misunderstanding, for example about having limbs cut off to save myself, or about the danger I expose my life to in order to preserve my life, and so on. That belongs to morality proper.)

Second, as for necessary or owed duty to others: a person who intends to make a lying promise to others will see at once that they want to use another human being *merely as a means*, without that person at the same time containing the end in themselves. For the person I want to use for my own purposes by such a promise cannot possibly agree to my way of treating them,

*) I set up this proposition here as a postulate. The reasons for it will be found in the last section.

430 and so cannot contain the end of this action in themselves. This conflict with the principle, where other people are concerned, is clearer when one brings in examples of attacks on the freedom and property of others. There it is plainly evident that a person who violates the rights of human beings means to use the persons of others merely as means. They do not take into account that others, as rational beings, should always be valued at the same time as ends, that is, only as beings who must also be able to contain in themselves the end of that very same action*).

Third, as for contingent (meritorious) duty to oneself, it is not enough that the action does not conflict with humanity in our person as an end in itself. It must also *harmonize with it*. Now humanity has natural capacities for greater perfection. They belong to nature's end for humanity in us as subjects. Neglecting them could perhaps be compatible with the *preservation* of humanity as an end in itself, but not with the *furthering* of this end.

Fourth, as for meritorious duty to others: the natural end that all human beings have is their own happiness. Now humanity could indeed continue to exist if no one contributed anything to the happiness of others, provided they did not deliberately take anything away from it. But this would be only a negative agreement, not a positive one, with *humanity as an end in itself*, if everyone also failed to strive, as far as they can, to further the ends of others. For the subject who is an end in itself must have ends that are, as far as possible, also *my* ends, if that idea is to have *its full* effect in me.

This principle of humanity, and of every rational nature in general, *as an end in itself* (which is the supreme limiting

*) Let no one think that the trivial saying "What you do not want done to you," and so on, can serve here as a guide or principle. For it is only derived from that principle, and with various limits. It cannot be a universal law, because it does not contain the basis of duties to oneself, or of duties of love toward others (for many a person would gladly agree that others should not do them good, if only they were excused from doing good to others), and finally not of owed duties toward one another. For on this basis the criminal would argue against the judges who punish them, and so on.

431 condition of the freedom of each human being's actions) is not borrowed from experience. There are two reasons. First, it is universal, since it applies to all rational beings in general, and no experience is enough to settle anything about that. Second, in it humanity is presented not as an end of human beings (subjectively), that is, not as an object that people of their own accord actually make their end. It is presented as an objective end, one that is meant, as a law, to be the supreme limiting condition of all subjective ends, whatever ends we may have. So it must arise from pure reason. For the basis of all practical lawgiving lies *objectively in the rule* and in the form of universality, which makes it capable of being a law (at least a law of nature). That is according to the first principle. *Subjectively*, however, it lies in the *end*. But the subject of all ends is every rational being, as an end in itself. That is according to the second principle. From this now follows the third practical principle of the will, as the supreme condition of its harmony with universal practical reason: the idea of *the will of every rational being as a universally lawgiving will*.

By this principle, all maxims are rejected that cannot be compatible with the will's own universal lawgiving. So the will is not simply subject to the law. It is subject to it in such a way that it must be seen *as giving the law to itself*, and for that very reason, in the first place, as subject to the law (of which it can regard itself as the author).

The imperatives in the earlier way of presenting them were these: the universal lawfulness of actions, resembling an *order of nature*; and the universal *priority as ends* of rational beings in themselves. These did exclude from their commanding authority every mixing in of some interest as an incentive, precisely because they were presented as categorical. But they were only assumed to be categorical, because one had to assume something of that kind if one wanted to explain the concept of duty. That there are practical propositions that command categorically could not be proved by itself, just as it cannot be done anywhere in this section, and so cannot yet be done here either. But one thing could have been done, namely this. The renunciation of all interest in willing

from duty is the specific mark that distinguishes the categorical from the hypothetical imperative. That renunciation could have been indicated as well in the imperative itself, through

432 some provision that it contained. And this is what happens in the present third formula of the principle, namely the idea of the will of every rational being as a *universally lawgiving* will.

For if we think of such a will, then a will that *stands under laws* may still be bound to that law by means of some interest. But a will that is itself supremely lawgiving cannot possibly, to that extent, depend on any interest. For such a dependent will would itself still need another law, which would limit the interest of its self-love to the condition of being valid as a universal law.

So the *principle* of every human will, as *a will that is universally lawgiving through all its maxims **), would fit the categorical imperative very well, if it were otherwise sound. It would fit it in this: precisely because of the idea of universal lawgiving, it is *based on no interest*, and so it alone among all possible imperatives can be *unconditional*. Or, better still, we can turn the proposition around: if there is a categorical imperative (that is, a law for the will of every rational being), it can only command us to do everything from the maxim of our will as the will of one who could at the same time have itself, as universally lawgiving, for its object. For only then are the practical principle and the imperative that one obeys unconditional, because it can have no interest at all as its basis.

Now it is no wonder, when we look back on all the efforts ever made to find the principle of morality, why they all had to fail. People saw human beings as bound to laws by their duty. It never occurred to them that human beings are subject *only to their own lawgiving, which is nevertheless universal*, and that they are bound only to act in line with their own will, which, according to its natural end, is universally lawgiving. For if people thought of human beings only as subject to a law (whatever law it might be), then this law had to bring with it some

*) I can be excused here from giving examples to illustrate this principle, for those that first illustrated the categorical imperative and its formula can all serve the same purpose here.

433 interest as an attraction or a compulsion, because the law did not arise as law from *their own* will. Instead, this will was constrained by *something else*, in accordance with a law, to act in a certain way. By this entirely necessary consequence, though, all the work of finding a supreme basis of duty was lost beyond recovery. For what people got was never duty, but only the necessity of an action arising out of some interest. This interest might be one's own or someone else's. But then the imperative always had to turn out conditional, and it could not be suited at all to be a moral command. So I will call this principle the principle of the *autonomy* of the will, in contrast with every other principle, which I therefore count as *heteronomy*.

Every rational being must regard themselves, through all the maxims of their will, as universally lawgiving, so as to judge themselves and their actions from this point of view. This concept leads to another concept that is attached to it and is very fruitful, namely the concept of a *realm of ends*.

By a realm I mean the systematic union of different rational beings through common laws. Now laws determine ends according to their universal validity. So if we set aside the personal differences between rational beings, and likewise all the content of their private ends, we can think of a whole of all ends in systematic connection (both of the rational beings as ends in themselves, and of the particular ends that each may set for themselves). That is, we can think of a realm of ends, which is possible according to the principles above.

For all rational beings stand under the *law* that each of them should treat themselves and all others *never merely as a means*, but always at the same time *as an end in itself*. From this arises a systematic union of rational beings through common objective laws, that is, a realm. Because these laws have in view precisely the relation of these beings to one another as ends and means, this realm can be called a realm of ends (admittedly only an ideal).

A rational being belongs to the realm of ends as a *member* if they are universally lawgiving in it, but are also themselves subject to these laws. They belong to it as *head* if, as lawgiving, they are subject to no one else's will.

434 A rational being must always regard themselves as lawgiving in a realm of ends made possible by freedom of the will, whether as a member or as head. But they can hold the place of head not merely through the maxim of their will. They can hold it only if they are a completely independent being, without need and without any limit on a power that is adequate to their will.

Morality, then, consists in the relation of every action to the lawgiving through which alone a realm of ends is possible. This lawgiving must be found in every rational being themselves, and it must be able to arise from their will. The principle of this will is therefore never to act on a maxim unless it is also possible for that maxim to be a universal law. That is, to act only in such a way *that the will can regard itself, through its maxim, as at the same time universally lawgiving*. Now if the maxims do not already agree necessarily, by their very nature, with this objective principle of rational beings as universally lawgiving, then the necessity of acting according to that principle is called practical constraint, that is, *duty*. Duty does not apply to the head in the realm of ends, but it does apply to every member, and indeed to all of them in equal measure.

The practical necessity of acting by this principle, that is, duty, does not rest at all on feelings, impulses, and inclinations. It rests solely on the relation of rational beings to one another. In this relation the will of a rational being must always be regarded at the same time as *lawgiving*, because otherwise they could not think of these beings as *ends in themselves*. So reason relates every maxim of the will, as universally lawgiving, to every other will, and also to every action toward oneself. It does this not for the sake of any other practical motive or future advantage, but from the idea of the *dignity* of a rational being, who obeys no law except the one they at the same time give themselves.

In the realm of ends, everything has either a *price* or a *dignity*. Whatever has a price can be replaced by something else as its *equivalent*. Whatever, on the other hand, is above all price, and so allows no equivalent, has a dignity.

What relates to the general inclinations and needs of human beings has a *market price*. What, even without presupposing a need, suits a certain taste, that is, a liking for the

435 mere purposeless play of the powers of our mind, has an *affection price*. But what makes up the condition under which alone something can be an end in itself has not merely a relative value, that is, a price, but an inner value, that is, *dignity*.

Now morality is the condition under which alone a rational being can be an end in itself, because only through it is it possible to be a lawgiving member in the realm of ends. So morality, and humanity so far as it is capable of morality, is the only thing that has dignity. Skill and diligence in work have a market price. Wit, lively imagination, and humor have an affection price. Faithfulness in promises and goodwill from principles (not from instinct), on the other hand, have an inner value. Neither nature nor art contains anything that could take their place if they were lacking. For their value does not consist in the effects that arise from them, in the advantage and benefit they produce. It consists in the attitudes of mind, that is, the maxims of the will, which are ready to show themselves in this way in actions, even if success did not favor them. These actions also need no recommendation from any subjective disposition or taste to look on them with immediate favor and liking, and no immediate tendency or feeling for them. They present the will that carries them out as an object of immediate respect. Nothing but reason is required for this, in order to *impose* them on the will, and not to *get them from it by flattery*, which would in any case be a contradiction where duties are concerned. This valuation therefore shows the value of such a way of thinking to be dignity. It places that way of thinking infinitely above all price. It can in no way be reckoned or compared with any price without, so to speak, doing violence to its holiness.

And what is it, then, that entitles the morally good attitude of mind, or virtue, to make such high claims? It is nothing less than the *share* that it gives the rational being *in universal lawgiving*. Through this share the being becomes fit to be a member in a possible realm of ends. Their own nature had already des-

tined them for this, as an end in itself and for that very reason as lawgiving in the realm of ends, as free with regard to all laws of nature, obeying only those that they themselves give, and according to which their maxims can belong to a universal lawgiving (to which they at the same time

436 subject themselves). For nothing has a value except the one that the law determines for it. But the lawgiving itself, which determines all value, must for that very reason have a dignity, that is, an unconditional, incomparable value. For this value the word *respect* alone supplies the fitting expression of the valuation that a rational being must make of it. *Autonomy* is therefore the basis of the dignity of human nature and of every rational nature.

The three ways given of presenting the principle of morality are, at bottom, just so many formulas of one and the same law. Any one of them unites the other two within itself, by itself. Still, there is a difference among them. It is more subjective than objectively practical: its point is to bring an idea of reason closer to intuition (by a certain analogy), and in this way closer to feeling. For all maxims have

1) a *form*, which consists in universality. Here the formula of the moral imperative is expressed this way: maxims must be chosen as if they were to hold as universal laws of nature;

2) a *matter*, namely an end. Here the formula says: the rational being, as an end by its nature, and so as an end in itself, must serve every maxim as the limiting condition of all merely relative and arbitrary ends;

3) a *complete determination* of all maxims by that formula, namely: all maxims, out of their own lawgiving, should harmonize into a possible realm of ends as a realm of nature*). The progress here goes as if through the categories of the *unity* of the form of the will (its universality), the *plurality* of the matter (of the objects, that is, of the ends), and the *allness* or totality of the system of them. But one does better, in moral *judging*, always to proceed by the strict method and to take as the basis the universal formula of the categorical imperative: *act on the maxim that can at the same time make itself a universal law*

*) Teleology considers nature as a realm of ends. Morality considers a possible realm of ends as a realm of nature. There, the realm of ends is a theoretical idea for explaining what exists. Here, it is a practical idea for bringing about what does not exist but can become real through what we do and refrain from doing, and indeed in keeping with this very idea.

⁴³⁷ *can make*. But if we also want the moral law to gain entry, it is very useful to run one and the same action through the three concepts mentioned. That brings the action, as far as possible, closer to intuition.

We can now end where we began: with the concept of an unconditionally good will. A *will* is *absolutely good* if it cannot be evil. So its maxim, when made a universal law, can never conflict with itself. This principle is therefore also its supreme law: always act on that maxim whose universality as a law you can at the same time will. This is the only condition under which a will can never be in conflict with itself. And such an imperative is categorical. The will's validity as a universal law for possible actions is analogous to the universal connection of the existence of things under universal laws, which is the formal side of nature in general. So the categorical imperative can also be expressed this way: *Act on maxims that can at the same time have themselves as universal laws of nature for their object*. This, then, is what the formula of an absolutely good will is like.

Rational nature stands out from the rest by setting an end for itself. This end would be the matter of every good will. But in the idea of a will that is absolutely good, without any limiting condition (the achievement of this or that end), we must set aside completely every end *to be brought about*. Such an end would make every will good only relatively. So here the end must be thought of not as one to be brought about but as a *self-standing* end, and therefore only negatively. That is, it is an end that must never be acted against. It must therefore never be valued merely as a means in any act of willing, but always at the same time as an end. This end can be nothing other than the subject of all possible ends itself. For that subject is also the subject of a possible absolutely good will, and such a will cannot, without contradiction, be ranked below any other

object. The principle: act toward every rational being (yourself and others) so that in your maxim it counts at the same time as an end in itself, is therefore basically the same as the principle: act on a maxim that, for every rational being,

438 at the same time contains its own universal validity within itself. At bottom the two are one and the same. For I must limit my maxim, in using means for any end, by the condition that it be valid universally as a law for every subject. That says just as much as this: the subject of ends, that is, the rational being itself, must never be made the basis of the maxims of action merely as a means. It must be made their basis as the supreme limiting condition on the use of all means—that is, always at the same time as an end.

It follows beyond dispute from this that every rational being, as an end in itself, must be able to regard itself as also giving universal law with respect to all the laws to which it may ever be subject. For this very suitability of its maxims for universal lawgiving marks it out as an end in itself. It also follows that its dignity (prerogative) above all mere natural beings brings with it this requirement: it must always take its maxims from its own standpoint and, at the same time, from the standpoint of every other rational being as a lawgiving being. (Such beings are for this reason also called persons.) In this way a world of rational beings (*mundus intelligibilis*) is possible as a realm of ends, and indeed through the lawgiving of all persons themselves, as members. So every rational being must act as if, through its maxims, it were always a lawgiving member in the universal realm of ends. The formal principle of these maxims is: act as if your maxim were also to serve as a universal law (for all rational beings). A realm of ends is therefore possible only by analogy with a realm of nature. Yet the first is possible only through maxims, that is, rules that one imposes on oneself, and the second only through laws of causes that produce effects under outside compulsion. Even so, we also call the whole of nature a realm of nature, although we regard it as a machine. We call it that on this ground: so far as it is related to rational beings as its ends. Such a realm of ends would really come about through maxims whose rule the categorical imperative prescribes to all rational beings, if they were *followed universally*. However, a rational being cannot

count on this: that if it follows this maxim punctually itself, every other being will therefore be faithful to the same maxim. Nor can it count on this: that the realm of nature and its purposive arrangement will harmonize with it, as a suitable member, toward a realm of ends possible through that being itself—that is, that its expectation of happiness

439 will be favored. Even so, that law remains in full force, because it commands categorically: act on the maxims of a universally lawgiving member of a merely possible realm of ends. And here lies the very paradox. It is that the mere dignity of humanity as rational nature, with no other end or advantage to be reached through it—and so mere respect for a bare idea—should nevertheless serve as a rule for the will that never lets up. It is also that the loftiness of the maxim consists precisely in its independence from all such incentives, and so does the worthiness of every rational subject to be a lawgiving member in the realm of ends. Otherwise that subject would have to be pictured as subject only to the law of nature that governs its needs. Suppose the realm of nature and the realm of ends were both thought of as united under one head. Then the realm of ends would no longer remain a bare idea but would gain true reality. This would indeed give it the addition of a strong incentive, but never an increase in its inner worth. For even so, this sole, unlimited lawgiver would still always have to be pictured as judging the worth of rational beings only by their unselfish conduct, which they prescribe to themselves purely from that idea. The essence of things does not change because of their external relations. What alone makes up a human being's absolute worth, leaving those relations out of account, is the standard by which that human being must be judged—by anyone at all, even by the highest being. *Morality* is therefore the relation of actions to the autonomy of the will, that is, to possible universal lawgiving through the will's maxims. An action that can coexist with the autonomy of the will is *permitted*. One that does not agree with it is *not permitted*. A will whose maxims necessarily agree with the laws of autonomy is a *holy* will, an absolutely good will. The

dependence of a will that is not absolutely good on the principle of autonomy (moral constraint) is *obligation*. Obligation therefore cannot be applied to a holy being. The objective necessity of an action from obligation is called *duty*.

From what was just said, it is now easy to explain how this happens. The concept of duty makes us think of being subject to the law. Yet by that very concept we also, at the same time,

440 picture a certain loftiness and *dignity* in the person who fulfills all their duties. For there is indeed no loftiness in that person in so far as they are *subject* to the moral law. But there is loftiness in so far as, with respect to that same law, they are at the same time *lawgiving*, and are subordinate to it only for that reason. We also showed above that neither fear nor inclination, but only respect for the law, is the incentive that can give an action moral worth. Our own will, in so far as it would act only under the condition of a universal lawgiving possible through its maxims—this will, which is possible for us in the idea—is the proper object of respect. The dignity of humanity consists in just this capacity to give universal law, though on the condition of also being itself subject to this same lawgiving.

The autonomy of the will

as the supreme principle of morality.

Autonomy of the will is the quality of the will by which it is a law to itself, independent of every quality of the objects of willing. The principle of autonomy is therefore: to choose only in such a way that the maxims of one's choice are also included, in that same willing, as a universal law. That this practical rule is an imperative—that is, that the will of every rational being is necessarily bound to it as a condition—cannot be proved by merely analyzing the concepts found in it, because it is a synthetic proposition. We would have to go beyond the cognition of objects and on to a critique of the subject, that is, of pure practical reason. For this synthetic proposition, which commands apodictically, must be capable of being cognized entirely a priori. But that task does not belong in the present section. That the principle of autonomy just stated is the only principle of morality, however, can very well be shown by merely analyz-

ing the concepts of morality. For such analysis finds that the principle of morality must be a categorical imperative, and that this imperative commands nothing more or less than precisely this autonomy.

441 The heteronomy of the will

as the source of all counterfeit principles of morality.

If the will looks for the law that is to determine it in *something else* than the fitness of its maxims for its own universal lawgiving—that is, if it goes beyond itself and looks for the law in the nature of one of its objects—the result is always *heteronomy*. Then the will does not give the law to itself. Instead the object gives the law to the will, through its relation to the will. Whether this relation rests on inclination or on ideas of reason, it allows only hypothetical imperatives to be possible: I ought to do something *because I want something else*. By contrast, the moral, and therefore categorical, imperative says: I ought to act in such and such a way, even if I wanted nothing else. For example, the first says: I ought not to lie, if I want to keep my honor. The second says: I ought not to lie, even if it would bring me not the slightest disgrace. The latter, then, must set aside every object so completely that the object has no *influence* at all on the will. The point is that practical reason (will) should not merely manage an interest that is not its own, but simply show its own commanding authority as supreme lawgiving. So, for example, I ought to try to promote the happiness of others, not as if its existence mattered to me (whether through direct inclination, or through some pleasure taken in it indirectly, by way of reason), but only because the maxim that excludes it cannot be included in one and the same willing as a universal law.

Division

of all possible principles of morality from the assumed basic concept of heteronomy.

As long as human reason lacks a critique, it does here what it does everywhere in its pure use: it first tries every possible wrong way before it succeeds in hitting on the one true way.

All principles that one may take from this standpoint are either *empirical* or *rational*. The *first*, which come from the principle

442 of *happiness*, rest on physical or moral feeling. The second, which come from the principle of *perfection*, rest either on the rational concept of perfection as a possible effect, or on the concept of a self-standing perfection (the will of God) as the cause that determines our will.

Empirical principles are altogether unsuited to be a foundation for moral laws. For the universality with which these laws are supposed to hold for all rational beings without distinction, and the unconditional practical necessity this places on them, drops away if their ground is taken from the *particular constitution of human nature* or from the accidental circumstances in which human nature is placed. Still, the principle of *one's own happiness* is the most objectionable of all. This is not merely because it is false, and because experience contradicts the claim that well-being always matches good conduct. It is also not merely because it contributes nothing to establishing morality, since it is one thing to make someone happy and another to make them good, and one thing to make them prudent and sharp about their own advantage and another to make them virtuous. It is most objectionable because it puts incentives under morality that are more likely to undermine it and destroy all its loftiness. They put the motives for virtue in the same class as those for vice. They teach only how to calculate better, and they wipe out entirely the specific difference between the two. On the other hand, take moral feeling, this supposed special sense*). To be sure, appealing to it is shallow. People who cannot *think* believe that *feeling* will see them through, even where only universal laws are at stake. Feelings also differ from one another infinitely in degree, by nature, so they give no equal standard of good and evil. And no one can judge validly for others by their own feeling at all. Even so, moral feeling stays closer to morality and its dignity. It does virtue the honor of ascribing the liking and high esteem for it to virtue *directly*, and it does not tell it, so to speak,

*) I count the principle of moral feeling under that of happiness, because every empirical interest promises a contribution to well-being through the mere pleasantness that something affords, whether this happens directly and without an eye to advantages, or with regard to them. Likewise, along with *Hutcheson*, one must count the principle of sympathy with the happiness of others under the same moral sense assumed by Hutcheson.

443 to its face that what ties us to it is not its beauty but only advantage.

Among the *rational*, or reason-based, grounds of morality is the ontological concept of *perfection*. This concept is empty and indeterminate, and so useless for finding, in the immense field of possible reality, the greatest sum that suits us. When it tries to tell the reality in question here apart specifically from every other, it has an unavoidable tendency to turn in a circle, and it cannot avoid secretly presupposing the very morality it is supposed to explain. Even so, it is better than the theological concept, which derives morality from a divine, most perfect will. It is better not merely because we do not, after all, intuit that will's perfection but can derive it only from our own concepts, of which the concept of morality is the foremost. It is better because of what follows if we do not do this. (If we did do it, it would be a gross circle in explaining.) The concept of that will that would then remain to us is built from the qualities of desire for honor and for domination, joined with fearful ideas of power and eagerness for revenge. It would have to be the foundation for a system of morals directly opposed to morality.

But if I had to choose between the concept of the moral sense and that of perfection in general (neither of which at least does damage to morality, although they are quite unfit to support it as foundations), I would decide for the latter. The reason is that it at least draws the decision of the question away from sensibility and before the court of pure reason. It decides nothing even there. Still, it keeps in reserve, unfalsified, the indeterminate idea (of a will that is good in itself) for closer determination.

For the rest, I believe I can be spared a lengthy refutation of all these doctrines. It is so easy. It is presumably also so well understood even by those whose office requires them to declare for one of these theories (because audiences do not much like a postponement of judgment) that the effort would be wasted. What interests us more here is to know this: that these principles everywhere set up nothing but heteronomy of the will as the first ground of morality, and for that very reason must necessarily miss their aim.

444 Everywhere that an object of the will must be laid down as the basis in order to prescribe to the will the rule that is to determine it, the rule is nothing but heteronomy. The imperative is conditional, namely: *if* or *because* one wants this object, one ought to act in such and such a way. So it can never command morally, that is, categorically. The object may determine the will by means of inclination, as in the principle of one's own happiness, or by means of reason directed at the objects of our possible willing in general, as in the principle of perfection. Either way, the will never determines itself *directly*, through the idea of the action alone. It is determined only through the incentive that the foreseen effect of the action has on the will. *I ought to do something because I want something else*. And here yet another law must be laid down as the basis in me as subject, according to which I necessarily want this something else. That law in turn needs an imperative to limit this maxim. For the impulse that the idea of an object possible through our powers is supposed to exert on the subject's will, according to the subject's natural makeup, belongs to the nature of the subject. It may belong to sensibility (inclination and taste), or to understanding and reason, which by the particular constitution of their nature exercise themselves with pleasure on an object. So it would really be nature that gives the law. Not only must that law be known and proved through experience, which makes it contingent in itself and so unfit to be an apodictic practical rule, as the moral rule must be. It is also *always only heteronomy* of the will: the will does not give itself the law, but a foreign impulse gives it the law, by means of a nature of the subject that is attuned to the will's receptivity.

The absolutely good will, whose principle must be a categorical imperative, will therefore be undetermined with respect to all objects. It will contain only the *form of willing* in general, and indeed as autonomy. That is, the fitness of the maxim of every good will to make itself a universal law is itself the only law that the will of every rational being imposes on itself, without putting any incentive or interest beneath it as a ground.

How such a synthetic practical a priori proposition is possible, and why it is necessary, is a task whose solution no longer lies within the limits of the metaphysics of morals. Nor have we

445 asserted its truth here, much less claimed to have a proof of it in our power. We showed only, by developing the concept of morality that is now generally current, that an autonomy of the will unavoidably attaches to it, or rather underlies it. Whoever therefore holds morality to be something, and not a fanciful idea without truth, must at the same time grant the principle of morality that was stated. This section was therefore, just like the first, merely analytic. Now, that morality is no figment of the mind follows if the categorical imperative, and with it the autonomy of the will, is true and absolutely necessary as an a priori principle. This requires a *possible synthetic use of pure practical reason*. But we may not venture on that use without first sending ahead a *critique* of this faculty of reason itself. In the last section we have to present the main outlines of that critique, enough for our purpose.

Section III



446 Third Section

Transition from the Metaphysics of Morals to the Critique of Pure Practical Reason

The concept of freedom is the key to explaining the autonomy of the will.

The *will* is a kind of causality that living beings have insofar as they are rational. *Freedom* would be the property of this causality that allows it to act independently of outside causes that *determine* it. In contrast, *natural necessity* is the property of the causality of all nonrational beings by which outside causes determine them to act.

This explanation of freedom is *negative*. It therefore does not help us understand what freedom is in itself. But it leads to a *positive* concept of freedom that is all the richer and more fruitful. The concept of causality includes that of *laws*: through something we call a cause, something else, its effect, must be brought about according to these laws. So although freedom is not a property of the will under laws of nature, it is not lawless. Instead, it must be a causality governed by unchanging laws, though laws of a special kind. Otherwise a free will would be an impossibility. Natural necessity was a heteronomy of acting causes. For every effect was possible only under the law that something else determined the acting cause to exercise its causality. What else, then, can freedom

447 of the will be but autonomy—that is, the will’s property of being a law to itself? But the statement that the will is a law to itself in all its actions expresses only this principle: act on no maxim except one that can also have itself, as a universal law, as its object. This is precisely the formula of the categorical imperative and the principle of morality. A free will and a will under moral laws are therefore one and the same.

If freedom of the will is assumed, then morality and its principle follow simply by analyzing the concept of freedom. Still, the latter principle remains a synthetic proposition: an absolutely good will is one whose maxim can always contain itself when considered as a universal law. For this property of the maxim cannot be found by analyzing the concept of an absolutely good will. Such synthetic propositions are possible only when the two items of knowledge are connected through their connection with a third, in which both can be found. The *positive* concept of freedom provides this third thing. It cannot be the nature of the world of sense, as it can be with physical causes. (The concept of that nature brings together the concepts of something as a cause and of *something else* as its effect.) We cannot yet say here what this third thing is, to which freedom points us and of which we have an a priori idea. Nor can we yet make the derivation of the concept of freedom from pure practical reason intelligible, and with it the possibility of a categorical imperative. This still requires some preparation.

Freedom must be assumed as a property of the will of all rational beings.

It is not enough to attribute freedom to our will, whatever our reason for doing so, unless we have sufficient grounds for attributing the very same freedom to all rational beings. Morality serves as a law for us only because we are *rational* beings. It must therefore hold for all rational beings as well. And since it must be derived solely from the property of freedom, freedom must also be proved to be a property of the will of all rational beings. It is not enough to establish it from certain supposed experiences

448 of human nature. (In fact, that is completely impossible: it can be established only a priori.) Instead, we must prove that freedom belongs to the activity of rational beings endowed with a will as such. I now say this: any being that cannot act except *under the idea of freedom* is, for that very reason, actually free in a practical respect. That is, all the laws inseparably connected with freedom hold for it, just as if its will had also been validly declared free in itself in theoretical philosophy.*) I further claim that we must necessarily give every rational being that has a will the idea of freedom, under which alone it acts. For in such a being we think of a reason that is practical—that is, that has causality in relation to its objects. But we cannot possibly think of a reason that is conscious of receiving direction for its judgments from somewhere else. In that case, the subject would attribute the determination of its judgment to an impulse, not to its reason. Reason must see itself as the source of its principles, independent of outside influences. So, as practical reason, or as the will of a rational being, it must regard itself as free. In other words, this being's will can be a will of its own only under the idea of freedom. For practical purposes, we must therefore attribute freedom to all rational beings.

On the interest attached to the ideas of morality

We have finally traced the definite concept of morality back to the idea of freedom. Yet we could not prove that freedom is something real, even in ourselves and in human nature. We

*) I take this approach—accepting freedom as sufficient for our purpose when rational beings merely base their actions on it as an idea—so that I need not undertake to prove freedom from a theoretical standpoint as well. Even if that theoretical question remains unsettled, the same laws hold for a being that cannot act except under the idea of its own freedom as would bind a being that really was free. Here, then, we can free ourselves from the burden that weighs on theory.

449 saw only that we must assume freedom if we are to think of a being as rational and as conscious of its causality in relation to actions—that is, as endowed with a will. We thus find that, on precisely the same grounds, we must attribute this

property to every being endowed with reason and will: it determines itself to act under the idea of its freedom.

But assuming these ideas also led to awareness of a law for acting: the subjective principles of actions, or maxims, must always be taken in such a way that they can also hold objectively—that is, universally—as principles. They can therefore serve for our own universal lawgiving. But why should I submit to this principle, and do so as a rational being as such, so that all other beings endowed with reason must do so too? I will grant that no interest *drives* me to do this, since that would not yield a categorical imperative. But I must still necessarily *take* an interest in it and understand how this happens. For this “ought” is really a willing that holds for every rational being on the condition that its reason is practical without obstacles. Beings like us are also affected by sensibility, which supplies motives of another kind. In us, what reason alone would do does not always happen. So for us, this necessity of action is called only an “ought,” and subjective necessity is distinguished from objective necessity.

It therefore seems that, in the idea of freedom, we have really only assumed the moral law itself, namely the principle of autonomy of the will. We seem unable to prove its reality and objective necessity on their own. We would still have gained something quite substantial: at least we would have defined the genuine principle more precisely than is usually done. But we would have made no progress on its validity or on the practical necessity of submitting to it. Suppose someone asked us why the universal validity of our maxim as a law must be the condition that limits our actions. Suppose they asked what grounds the value we give to this way of acting—a value said to be so great that there can be no higher interest at all—and how it happens that a person believes this alone lets them feel their personal worth,

450 beside which the value of a pleasant or unpleasant condition counts for nothing. We could give no satisfactory answer.

We do find that we can take an interest in a personal quality that brings no interest in our condition with it, provided this quality makes us fit to share in that condition if reason were to distribute it. In other words, simply being wor-

thy of happiness can interest us in itself, even without the motive of obtaining that happiness. But this judgment is really only an effect of the importance already assumed for moral laws, when we use the idea of freedom to separate ourselves from every empirical interest. We still cannot understand in this way why we should separate ourselves from that interest—that is, regard ourselves as free in acting—and yet hold ourselves subject to certain laws, in order to find a value merely in our person that could compensate us for the loss of everything that gives our condition a value. We thus still cannot understand how this is possible, and hence *what makes the moral law binding*.

We must freely admit that a kind of circle appears here, from which there seems to be no escape. We assume we are free in the order of acting causes so that we can think of ourselves as under moral laws in the order of ends. We then think of ourselves as subject to these laws because we have attributed freedom of the will to ourselves. For freedom and the will's own lawgiving are both autonomy, and so are interchangeable concepts. For that very reason, neither can be used to explain the other or give its ground. At most, they can be used for a logical purpose: to bring apparently different representations of the same object under a single concept, as different fractions with the same value are reduced to their simplest expressions.

But one way out remains. We can ask whether we take a different standpoint when we think of ourselves, through freedom, as causes that act a priori, from the standpoint we take when we picture ourselves through our actions as effects that we see before our eyes.

There is an observation that requires no subtle thinking. We can assume that even the most ordinary understanding can make it, though in its own way through an obscure distinction

⁴⁵¹ made by its judgment, which it calls feeling. The observation is this: all representations that come to us without our choosing them, such as those of the senses, let us know objects only as they affect us. What the objects may be in themselves remains unknown to us. So with this kind of representation, even the greatest attention and clarity that the understanding can add let us know

only *appearances*, never *things in themselves*. Once this distinction is made, something else follows on its own. Behind appearances we must admit and assume something else that is not appearance, namely things in themselves. We can draw the distinction, if only by noticing the difference between representations given to us from elsewhere, in which we are passive, and those we produce entirely from ourselves, in which we show our activity. At the same time, we freely acknowledge a limit: since we can never know these things, but only how they affect us, we cannot get any closer to them or ever know what they are in themselves. This must yield a distinction, though a rough one, between a *world of sense* and a *world of understanding*. The first can differ greatly with differences in the sensibility of the various observers of the world. The second, which underlies it, always remains the same. Even in knowing themselves through inner sensation, human beings must not claim to know themselves as they are in themselves. For people do not, so to speak, create themselves. They acquire their concept of themselves empirically, not a priori. So it is natural that they can learn about themselves only through inner sense, and thus only through the appearance of their nature and the way their consciousness is affected. Yet they must necessarily assume something else that underlies this character of their own subject, composed entirely of appearances: their self, as it may be in itself. So in respect of merely perceiving sensations and being receptive to them, they must count themselves as belonging to the *world of sense*. But in respect of whatever may be pure activity in them—whatever reaches consciousness directly, not through the senses being affected—they must count themselves as belonging to the *intellectual* world, which they nevertheless know no further. A thinking person must reach this kind of conclusion about all the things

⁴⁵² that may appear to them. The conclusion is probably also found in the most ordinary understanding. As we know, it is very inclined to expect something invisible, active on its own, behind the objects of the senses. But it spoils this by soon trying to give this invisible thing a sensory form again—that is, to make it an object of intuition. In this way it becomes no wiser at all.

Now people do actually find in themselves a capacity that distinguishes them from all other things, and even from themselves insofar as objects affect them. This capacity is *reason*. As pure activity of its own, reason is raised even above the *understanding* in the following way. Understanding, too, is activity of its own. Unlike sense, it does not merely contain representations that arise only when things affect us and we are therefore passive. Still, its activity can produce no concepts except those that serve to *bring sensory representations under rules* and so unite them in one consciousness. Without this use of sensibility it would think nothing at all. Reason, in contrast, displays such pure spontaneity in what are called ideas that it goes far beyond anything sensibility can supply. It shows its main work in distinguishing the world of sense from the world of understanding, and in this way setting limits for the understanding itself.

For this reason, a rational being must regard itself, as an *intelligence* and thus not from the side of its lower powers, as belonging to the world of understanding, not the world of sense. It therefore has two standpoints from which it can consider itself and know laws for using its powers, and thus for all its actions. *First*, insofar as it belongs to the world of sense, it is under laws of nature: heteronomy. *Second*, as belonging to the intelligible world, it is under laws that are independent of nature and based not on experience but on reason alone.

As a rational being, and thus one belonging to the intelligible world, a human being can think of the causality of their own will only under the idea of freedom. For independence from the determining causes of the world of sense—an independence reason must always attribute to itself—is freedom. Now the concept of *autonomy* is inseparably connected with the idea of freedom. With autonomy, in turn, is connected the universal principle of morality. In the idea, this principle underlies all the actions

⁴⁵³ of *rational* beings, just as the law of nature underlies all appearances.

We can now remove the suspicion we raised above: that there might be a hidden circle in our inference from freedom to autonomy and from autonomy to the moral law. The suspicion was that we might have made the idea of freedom our starting point only for the sake of the moral law, in order then to infer this

law from freedom in return. We would thus be unable to give any ground for the law. We would only be asking to have a principle granted to us. Well-disposed people might gladly grant it, but we could never put it forward as a provable proposition. For we now see that when we think of ourselves as free, we place ourselves in the world of understanding as its members and recognize the autonomy of the will, along with its consequence, morality. But when we think of ourselves as under obligation, we consider ourselves as belonging to the world of sense and yet also to the world of understanding.

How is a categorical imperative possible?

A rational being counts itself, as an intelligence, as belonging to the world of understanding. Only as an acting cause belonging to that world does it call its causality a *will*. On the other hand, it is also conscious of itself as a part of the world of sense. There its actions are found as mere appearances of that causality. But their possibility cannot be understood from this causality, which we do not know. Instead, those actions must be understood as determined by other appearances, namely desires and inclinations, and thus as belonging to the world of sense. If I were only a member of the world of understanding, all my actions would therefore agree perfectly with the principle of autonomy of the pure will. If I were only a part of the world of sense, my actions would have to be taken as agreeing entirely with the natural law of desires and inclinations, and thus with the heteronomy of nature. (The first would rest on the highest principle of morality, the second on that of happiness.) But *the world of understanding contains the ground of the world of sense, and thus of its laws as well*. It therefore gives laws directly to my will, which belongs entirely to the world of understanding, and must be thought of as giving laws in this way. So as an intelligence, although from the other side I am a being belonging to the

454 world of sense, I must recognize that I am subject to the law of the first world. That is, I am subject to reason, which contains that law in the idea of freedom, and so to autonomy of the will. I must therefore regard the laws of the world of understanding as imperatives for me, and actions that agree with this principle as duties.

This is how categorical imperatives are possible. The idea of freedom makes me a member of an intelligible world. If I were only that, all my actions always *would* agree with autonomy of the will. But since I also view myself as a member of the world of sense, they *ought* to agree with it. This *categorical* “ought” expresses a synthetic a priori proposition. For to my will as affected by sensory desires is added the idea of this very same will as pure, belonging to the world of understanding, and practical on its own. According to reason, this pure will contains the highest condition of the first will. This is roughly like the way concepts of the understanding are added to intuitions of the world of sense. These concepts by themselves mean nothing but the form of law in general. Their addition makes possible the synthetic a priori propositions on which all knowledge of a nature rests.

The practical use of ordinary human reason confirms that this derivation is correct. Take anyone, even the worst wrongdoer, as long as they are otherwise used to using reason. Show them examples of honesty in intentions, firmness in following good maxims, sympathy, and goodwill toward everyone, combined with great sacrifices of advantage and comfort. There is no one who would not wish to have that disposition too. But because of their inclinations and impulses, they cannot easily bring this about in themselves. Yet at the same time, they wish to be free of these inclinations, which are a burden even to them. This shows that, in thought, they place themselves with a will free from the impulses of sensibility in an entirely different order of things from that of their desires in the field of sensibility. For from that wish they can expect no gratification of their desires, and thus no condition that satisfies any of their actual or even conceivable inclinations. (That would rob the very idea that draws this wish from them of its superiority.) They can expect only a greater inner worth of their person. This better person

455 is what they believe themselves to be when they take the standpoint of a member of the world of understanding. The idea of freedom—that is, independence from the *determining* causes of the world of sense—forces them to this standpoint without their choosing it. Here they are conscious of a good will. By their

own admission, this good will provides the law for their bad will as a member of the world of sense. They recognize the law's authority even while breaking it. So the moral "ought" is their own necessary willing as a member of an intelligible world. They think of it as an "ought" only insofar as they also regard themselves as a member of the world of sense.

On the outermost limit of all practical philosophy

All human beings think of themselves as free in respect of their will. This is the source of all judgments about actions that *ought to have happened*, even though they *did not happen*. Yet this freedom is not a concept drawn from experience, nor can it be. For it always remains, even when experience shows the opposite of the requirements presented as necessary on the assumption of freedom. On the other hand, it is just as necessary that everything that happens be inevitably determined by laws of nature. This natural necessity is not a concept drawn from experience either, precisely because it includes the concept of necessity, and thus of a priori knowledge. But experience confirms this concept of a nature. It must itself inevitably be assumed if experience is to be possible—that is, knowledge of objects of the senses connected according to universal laws. Freedom is therefore only an idea of reason, whose objective reality is in itself doubtful. Nature, however, is a *concept of the understanding* that proves its reality in examples from experience, and must necessarily do so.

This gives rise to a dialectic of reason. The freedom attributed to the will seems to contradict natural necessity. At this crossroads, reason finds the path of natural necessity much better cleared and more usable for *speculative* purposes than the path of freedom. But for *practical* purposes, the footpath of freedom is the only one on which we can use our reason in what we do and leave

456 undone. So it is just as impossible for the most subtle philosophy as for the most ordinary human reason to argue freedom away. Reason must therefore assume that there is no real contradiction between freedom and the natural necessity of the very same human actions. For it can give up the concept of nature no more than it can give up the concept of freedom.

Still, this apparent contradiction must at least be convincingly removed, even if we can never comprehend how freedom is possible. For if even the thought of freedom contradicts itself, or contradicts nature, which is equally necessary, freedom would have to be given up entirely in favor of natural necessity.

But we cannot escape this contradiction if the subject that thinks itself free thinks of itself *in the same sense*, or *in exactly the same relation*, when it calls itself free as when it assumes it is subject to the law of nature with respect to the very same action. Speculative philosophy therefore has an unavoidable task. It must at least show that its mistake about the contradiction rests on this: we think of a human being in a different sense and relation when we call that person free than when we regard them as part of nature and subject to its laws. It must also show that the two can not only coexist perfectly well, but must be thought of as *necessarily united* in the same subject. Otherwise we could give no reason for burdening reason with an idea that involves us in a task that severely restricts reason in its theoretical use—even though this idea can be combined *without contradiction* with another that is well established. But this duty belongs only to speculative philosophy, so that it can clear the way for practical philosophy. The philosopher therefore cannot choose whether to remove the apparent conflict or leave it untouched. In the latter case, the theory on this point is ownerless property. The fatalist can rightfully take possession of it and drive all morality out of the supposed property that it occupied without title.

But we cannot yet say that the limit of practical philosophy begins here. For settling that dispute is not its business at all. It only requires speculative reason to end the disagreement in which it entangles itself in theoretical questions,

457 so that practical reason can have peace and security against outside attacks that might challenge its claim to the ground on which it wants to build.

But even ordinary human reason's rightful claim to freedom of the will rests on the awareness and the accepted assumption that reason is independent of merely subjectively determining causes. Taken together, these causes make up what belongs merely to sensation, and thus falls under the general name of sensibility. A human being who regards themselves in this way as an intelligence

thereby places themselves in a different order of things, and in relation to determining grounds of a quite different kind. This happens when they think of themselves as an intelligence endowed with a will, and thus with causality, rather than when they perceive themselves as an appearance in the world of sense—which they really are as well—and subject their causality to outside determination by laws of nature. Now they soon realize that both can, and indeed must, take place at once. There is not the slightest contradiction in a *thing as an appearance*, belonging to the world of sense, being subject to certain laws from which that very same *thing*, or being, *in itself* is independent. But they must represent and think of themselves in these two ways. The first is based on their awareness of themselves as an object affected through the senses. The second is based on their awareness of themselves as an intelligence: as independent of sensory impressions in their use of reason, and thus as belonging to the world of understanding.

This is why human beings claim a will that does not charge to its own account anything belonging merely to their desires and inclinations. On the other hand, they think of actions as possible through themselves, and even as necessary, that can take place only by setting aside all desires and sensory attractions. The causality of these actions lies in them as an intelligence, and in the laws of effects and actions according to principles of an intelligible world. They know nothing further about that world except that reason alone gives laws in it—specifically, pure reason, independent of sensibility. They also know that, since in that world they are their proper self only as an intelligence, while as a human being they are only an appearance of themselves, those laws concern them directly and categorically. Thus, what inclinations and impulses—and so the whole nature of the world of sense—urge them to do cannot impair the laws of their willing as an intelligence

458 —so much so that they do not hold themselves responsible for those inclinations and impulses, or attribute them to their proper self, their will. They do, however, hold themselves responsible for any indulgence they might show

them by allowing them to influence their maxims at the expense of the will's laws of reason.

Practical reason does not overstep its limits at all by *thinking* itself into a world of understanding. But it would do so if it tried to *see* or *feel* itself into that world. The former is only a negative thought in relation to the world of sense, which gives reason no laws for determining the will. It is positive only in this one point: that freedom, as a negative determination, is also connected with a positive capacity, and even with a causality of reason. We call this a will: the capacity to act so that the principle of actions agrees with the essential character of a rational cause, that is, with the condition that the maxim be universally valid as a law. But if reason also brought an *object of the will*, that is, a motive, from the world of understanding, it would overstep its limits and claim to know something about which it knows nothing. The concept of a world of understanding is therefore only a *standpoint* that reason finds itself forced to take outside appearances in order to *think of itself as practical*. This would be impossible if the influences of sensibility determined human beings. Yet it is necessary unless we deny them awareness of themselves as an intelligence, and thus as a rational cause active through reason—that is, a freely acting cause. This thought does of course bring with it the idea of an order and a lawgiving different from the mechanism of nature that applies to the world of sense. It makes necessary the concept of an intelligible world: the whole of rational beings as things in themselves. But it makes no claim to think here any further than in accordance with that world's merely *formal* condition. This is the universality of the will's maxim as a law, and thus the will's autonomy, which alone can coexist with its freedom. In contrast, all laws directed to an object yield heteronomy, which can be found only in laws of nature and can apply only to the world of sense.

But reason would overstep all its limits if it undertook to *explain how* pure reason can be practical.

⁴⁵⁹ That would be exactly the same task as explaining *how freedom is possible*.

For we can explain only what we can trace back to laws whose object can be given in some possible experience. But freedom is a mere idea. Its objective reality cannot be established in any way through laws of nature, and so cannot be established in any possible experience. No example can ever be supplied for it by any analogy. For this reason, it can never be comprehended or even understood. It holds only as a necessary assumption of reason in a being that believes itself conscious of a will. A will is a capacity distinct from the mere capacity to desire: it is the capacity to determine oneself to act as an intelligence, and thus according to laws of reason, independently of natural instincts. But where determination by laws of nature ends, all *explanation* ends too. Nothing remains but *defense*: rejecting the objections of those who claim to have looked deeper into the nature of things and therefore confidently declare freedom impossible. We can only show them where the contradiction they think they have discovered really lies. To apply the law of nature to human actions, they had to consider a human being as an appearance. Now they are asked to think of that person, as an intelligence, also as a thing in itself, but they still consider the person an appearance. On that view, of course, separating the person's causality, or will, from all laws of nature in the world of sense would involve a contradiction in one and the same subject. But the contradiction disappears if they reflect and admit, as they should, that behind appearances there must be things in themselves as their ground, though hidden. We cannot demand that the laws of their effects be the same as the laws under which their appearances stand.

The subjective impossibility of *explaining* freedom of the will is the same as the impossibility of finding and making intelligible an *interest* *)

*) Interest is that through which reason becomes practical—that is, a cause that determines the will. This is why we say only of a rational being that it takes an interest in something. Nonrational creatures feel only sensory impulses.

Reason takes a direct interest in an action only when the universal validity of the action's maxim is a sufficient ground for determining the will. Only an interest of this kind is pure. But if reason can determine the will only through another object of desire, or on the assumption of a particular feeling in the sub-

ject, it takes only an indirect interest in the action. Reason alone cannot, without experience, find either objects of the will or a particular feeling underlying it. So this latter interest would be merely empirical, not a pure interest of reason. Reason's logical interest—in advancing its insights—is never direct. It presupposes purposes for which reason is to be used.

460 that human beings could take in moral laws. Yet they actually do take an interest in them. We call the basis for this in us moral feeling. Some have wrongly presented it as the standard for our moral judgment. Instead, it must be regarded as the *subjective* effect that the law has on the will. Reason alone supplies the objective grounds for this effect.

For a rational being affected by the senses to will what reason alone prescribes as what it ought to do, reason must indeed have a capacity to *instill a feeling of pleasure* or satisfaction in fulfilling duty. Reason must therefore have a causality that determines sensibility in accordance with its principles. But it is completely impossible to understand—that is, to make intelligible a priori—how a mere thought, which itself contains nothing sensory, produces a sensation of pleasure or displeasure. For this is a special kind of causality. As with all causality, we can determine nothing about it a priori, but must consult experience alone. Yet experience can provide a relation of cause to effect only between two objects of experience. Here, however, pure reason, through mere ideas that provide no object for experience at all, is supposed to cause an effect that does lie in experience. So it is completely impossible for us human beings to explain how and why *the universality of a maxim as a law*, and thus morality, interests us. Only this much is certain: it is not valid for us *because it interests us* (for that is heteronomy and dependence of practical reason on sensibility, namely

461 on an underlying feeling, in which case reason could never give moral laws). Instead, it interests us because it is valid for us as human beings, since it has arisen from our will as an intelligence, and thus from our proper self. *But reason necessarily subordinates what belongs to mere appearance to the character of the thing in itself.*

We can therefore answer the question of how a categorical imperative is possible to this extent: we can state the one assumption under which alone it is possible, namely the idea of freedom. We can also understand the necessity of this assumption. This is enough for the *practical use* of reason—that is, to convince us of the *validity of this imperative*, and thus also of the moral law. But no human reason can ever understand how this assumption itself is possible. If we assume freedom of the will of an intelligence, however, its *autonomy* follows necessarily, as the formal condition under which alone that will can be determined. Assuming this freedom of the will is entirely *possible*, as speculative philosophy can show, without contradicting the principle of natural necessity in the connection of appearances in the world of sense. But it is also unconditionally *necessary* for a rational being to make freedom the practical condition, as an idea, underlying all its chosen actions, if that being is conscious of its causality through reason and thus of a will distinct from desires. Yet how can pure reason be practical on its own, without other motives taken from anywhere else? How can the mere *principle of the universal validity of all its maxims as laws*—which would indeed be the form of a pure practical reason—provide a motive on its own and produce an interest that would be called purely *moral*, without any matter or object of the will in which we might take an interest in advance? In other words, *how can pure reason be practical*? All human reason is completely unable to explain this, and every effort and labor spent seeking an explanation is wasted.

It is just as if I tried to discover how freedom itself is possible as the causality of a will. For then I leave behind the philosophical

462 ground of explanation and have no other. I could, of course, now wander around in the intelligible world that remains to me, the world of intelligences. But although I have an *idea* of it that is well grounded, I do not have the slightest *knowledge* of it. Nor can I ever reach such knowledge through all the efforts of my natural power of reason. The idea means only a something that remains when I have excluded everything belonging to the world of sense from the grounds that determine my will. Its only purpose is to restrict the principle of

motives drawn from the field of sensibility. I set a limit to that principle and show that it does not contain everything within itself: there is something more outside it. But I have no further knowledge of this something more. Once I remove all matter—that is, knowledge of objects—from the pure reason that thinks this ideal, nothing remains for me but its form. This is the practical law of the universal validity of maxims, and, in accordance with it, the thought of reason in relation to a pure world of understanding as a possible acting cause, that is, one that determines the will. Here a motive must be entirely absent, unless this idea of an intelligible world were itself the motive, or that in which reason originally took an interest. But making this intelligible is precisely the task we cannot solve.

Here, then, is the outermost limit of all moral inquiry. Defining it is already very important for two reasons. On the one hand, it keeps reason from searching around in the world of sense, in a way harmful to morality, for the highest motive and for a comprehensible but empirical interest. On the other, it keeps reason from beating its wings helplessly without moving forward in the space of transcendent concepts, empty for reason, under the name of the intelligible world, and losing itself among fantasies. Still, the idea of a pure world of understanding, as a whole of all intelligences to which we ourselves belong as rational beings—though we are also members of the world of sense from the other side—always remains a useful and permissible idea for the purpose of a rational faith, even though all knowledge ends at its boundary. Through the magnificent ideal of a universal kingdom of *ends in themselves*, or rational beings—a kingdom to which we can belong as members only if we carefully

463 conduct ourselves according to maxims of freedom as if they were laws of nature—this idea serves to produce in us a lively interest in the moral law.

Concluding Note

The speculative use of reason *with regard to nature* leads to the absolute necessity of some highest cause of the *world*. The practical use of reason *with regard to freedom* also leads to absolute necessity, but only that of the *laws of action* of a rational being as such. Now it is an essential *principle* of every use of our rea-

son to push its knowledge to the point of being conscious of that knowledge's *necessity*. Without this, it would not be knowledge of reason. But it is an equally essential *limitation* of this same reason that it can understand neither the *necessity* of what exists or happens, nor that of what ought to happen, unless it takes as its basis a *condition* under which that thing exists, happens, or ought to happen. In this way, however, reason's constant demand for the condition only postpones its satisfaction further and further. It therefore searches tirelessly for what is unconditionally necessary and finds itself forced to assume it, without any means of making it intelligible to itself. Reason is fortunate enough if it can merely find the concept that is compatible with this assumption. So the inability to make an unconditional practical law—such as the categorical imperative must be—intelligible in its absolute necessity is not a criticism of our derivation of the highest principle of morality. It is a charge that would have to be made against human reason in general. We cannot blame reason for being unwilling to do this through a condition, namely through some interest assumed as a basis. For then the law would not be a moral law, that is, the highest law of freedom. And so, although we do not comprehend the practical, unconditional necessity of the moral imperative, we do comprehend its *incomprehensibility*. This is all that can reasonably be required of a philosophy that strives, in its principles, to reach the limit of human reason.