Chapter 4
Fading Foundations and the Emergence of Justification

Abstract
A probabilistic regress, if benign, is characterized by the feature of fading foundations: the effect of the foundational term in a finite chain diminishes as the chain becomes longer, and completely dies away in the limit. This feature implies that in an infinite chain the justification of the target arises exclusively from the joint intermediate links; a foundation or ground is not needed. The phenomenon of fading foundations sheds light on the difference between propositional and doxastic justification, and it helps us settle the question whether justification is transmitted from one link in the chain to another, as foundationalists claim, or whether it emerges from a chain or network as a whole, as is maintained by coherentists and infinitists.

4.1 Fading Foundations

In the previous chapter we have introduced the idea of a probabilistic regress, and we have seen that such regresses are in general unproblematic: they mostly have a calculable limit, thus providing the target proposition, q, with a unique probability value. In all but a few exceptional cases there is no conceptual problem in saying that q is probabilistically supported by an epistemic chain of infinite length.

An important part of our argument concerned the rôle of the foundational or grounding proposition, p. In calculating the unconditional probability of the target, q, we managed to eliminate all the unconditional probabilities — except that of p. The factor $P(p)$ remained the only term in the chain of which the value was unknown. Consider the finite chain

D. Atkinson, J. Peijnenburg, *Fading Foundations*, Synthese Library 383,
DOI 10.1007/978-3-319-58295-5_4

$$q \longleftarrow A_1 \longleftarrow A_2 \longleftarrow \ldots \longleftarrow A_{m-1} \longleftarrow A_m \longleftarrow p,$$

where q is probabilistically supported by A_1, which is probabilistically supported by $A_2, \ldots,$ and so on, until A_m, which is probabilistically supported by the grounding proposition or belief p.

In any finite chain, we need to know the value of value of $P(p)$ in order to calculate $P(q)$. However, the *importance* of the unknown $P(p)$ for the probability of the target, $P(q)$, lessens as m gets bigger. If the chain is very short, consisting only of two propositions, q and p, then the importance of $P(p)$ for $P(q)$ is at its height: all the support for q comes from p (together with the pair of conditional probabilities that connect the one to the other). But now imagine that the chain is a little bit longer, consisting of three propositions:

$$q \longleftarrow A_1 \longleftarrow p.$$

In terms of nested rules of total probability this becomes:

$$P(q) = P(q|\neg A_1) + [P(q|A_1) - P(q|\neg A_1)]\{P(A_1|\neg p) \\ +[P(A_1|p) - P(A_1|\neg p)]P(p)\}. \tag{4.1}$$

In (4.1) the importance of $P(p)$ has somewhat decreased. It is still the case that it largely determines $P(q)$, but the influence of the conditional probabilities has become greater. In general it is so that, as the chain becomes longer, the support provided by the totality of the conditional probabilities increases, while that given by the foundation decreases. In other words, as m in A_m grows larger and larger, a law of diminishing returns come into force: the influence of $P(p)$ on $P(q)$ tapers off with each link, until it finally fades away completely. In the limit that m tends to infinity, all the probabilistic support for q comes from the conditional probabilities together, and none from the ground or foundation. This characteristic, that is essential to a probabilistic regress as we defined it, we call the feature of *fading foundations*. As we add more and more links to the chain the influence of $P(p)$ tails off, and $P(q)$ draws closer and closer to its final value.

The feature of fading foundations can be illustrated by our story about Barbara bacterium in the previous chapter. Recall that q is the proposition 'Barbara has trait T', A_n is 'Barbara's ancestor in the nth generation has T', and p is 'Barbara's primordial mother has T'. Now imagine that long and extensive empirical research in our laboratory has taught us that the probability that a bacterium has T is 0.99 when her mother has T, and that it is 0.04 when her mother lacks T:

$$P(q|A_1) = P(A_1|A_2) = \ldots = P(A_{m-1}|A_m) = P(A_m|p) = 0.99$$
$$P(q|\neg A_1) = P(A_1|\neg A_2)\ldots = P(A_{m-1}|\neg A_m) = P(A_m|\neg p) = 0.04$$

Let us further take for the unconditional probability of p the value 0.7. With the numbers we have chosen for the conditional probabilities, 0.99 and 0.04, the computed values for the unconditional probability of q are listed in the following table:

Table 4.1 Probability of q when the probability of p is 0.7

Number of A_n	1	2	5	10	25	50	75	100	∞
Probability of q	.710	.714	.726	.743	.774	.793	.798	.799	.8

The first entry in this table refers to the chain $q \longleftarrow A_1 \longleftarrow p$, where there is only one A. With the values that we have chosen in our example, the probability of the target proposition q yielded by this chain is 0.709. The second entry corresponds to the chain $q \longleftarrow A_1 \longleftarrow A_2 \longleftarrow p$. Here there are two A's, so the probabilistic support for q has grown, resulting in a probability for q that is somewhat higher, namely 0.714. The third entry refers to a chain of seven propositions: the target proposition q, five A's and the grounding proposition p. The support is still further augmented, and the probability of q equals 0.726. By including more and more A's we observe that the probabilistic support for q grows. The final entry corresponds to the situation where the chain is infinitely long. Here the probabilistic support for q has reached its maximum, culminating in the unconditional probability $P(q) = 0.8$. The latter can considered to be the 'real' value for the probability of q relative to the numbers chosen for the conditional probabilities.[1]

But now look at the second table, 4.2, where the conditional probabilities are the same as in Table 4.1, but where the unconditional probability of p is 0.95. There are two things that should be noted about these two tables. Firstly, the probability of q in Table 4.2 culminates in a limiting value that is the same as that in Table 4.1, namely 0.8. Secondly, while the numbers in

[1] In this table as well as in the following one, the values of the conditional probabilities are uniform, remaining the same throughout the chain. As has been explained in the previous chapter, and more in detail in the appendices, this is however not essential to the phenomenon of fading foundations. The argument goes through, in the usual class, when the values of the conditional probabilities differ from link to link.

Table 4.1 steadily increase as the number of links becomes larger, those in Table 4.2 go down. How can we understand these facts?

Table 4.2 Probability of q when the probability of p is 0.95

Number of A_n	1	2	5	10	25	50	75	100	∞
Probability of q	.935	.929	.910	.885	.840	.811	.803	.801	.8

The answer is provided by the feature of fading foundations. As the chain lengthens, the role of the foundation p becomes less and less important until it dies out completely. At the end of the day, the probability of q is fully determined by the conditional probabilities; everything comes from them and the influence of the foundation p has completely disappeared from the picture. The reason why the numbers in Table 4.1 go up, while those in Table 4.2 go down, is because in the first case the probability p is lower than the final real value of $P(q)$, relative to the chosen conditional probabulities, while in the second case it is higher. This is exactly what is to be expected as the foundational influence gradually peters out.

Lewis and Russell were right that, in a probabilistic regress, *something* goes to zero if m goes to infinity. However, this 'something' is not the value of $P(q)$, as they thought. Rather it is the influence that the foundation p has on the target q. This is not to say that p itself has become highly improbable, for p may have any probability value at all. It is rather that, in the limit, the effect of the would-be foundation p has faded away completely: the support it gives to q is nil.[2]

4.2 Propositions versus Beliefs

Up to this point we have not distinguished between propositional and doxastic justification: q, the A's, and p could be either propositions or beliefs.

[2] The fading influence of the foundation p should not be confused with the familiar washing out of the prior in Bayesian reasoning. In Bayesian updating, the prior probability becomes less and less important under the influence of new pieces of information coming in, until it washes out completely. Although this looks rather like the phenomenon of fading foundations, where the influence of p similarly diminishes, the two phenomena are actually quite different, as we explain in Appendix C.

However, it has often been pointed out that the distinction *is* relevant when we talk about justification, especially if we discuss the possibility of infinite justificatory chains. In this section we will look at a debate between Michael Bergmann and Peter Klein in order to explain how the phenomenon of fading foundations can shed light on the subject.[3]

Bergmann has critized Peter Klein's infinitism by arguing that, although propositional justification might go on and on, doxastic justification must always come to a stop; infinite epistemic chains and doxastic justification simply seem incompatible.[4] In a reply to Bergmann, Klein has acknowledged that, unlike propositional justification, doxastic justification is always finite. As he wryly notes, "We get tired. We have to eat. We have satisfied the enquirers. We die".[5] He does not regard this as a difficulty for infinitism, however, since the stop is merely contextual or pragmatic. According to Klein, "doxastic justification is parasitic on propositional justification": in principle it can go on, but in practice it ends.[6]

Bergmann, however, believes that Klein's position is untenable, arguing as follows.[7] In order to reject foundationalism, Klein must endorse the following view:

K_1: For a belief B_i to be doxastically justified, it must be based on some other belief B_j.

Bergmann then introduces

[3] See Peijnenburg and Atkinson 2014b. We will say a bit more about the distinction between propositional and doxastic justification in the next chapter, when we discuss Klein's reply to the notorious finite mind objection. For the difference between propositional and doxastic justification, see also Turri 2010.
[4] Bergmann 2007. Jonathan Kvanvig has argued that Klein's infinitism has difficulties not only accounting for doxastic justification, but for propositional justification too (Kvanvig 2014). We will briefy come back to Kvanvig's criticism in the next chapter.
[5] Klein 2007a, 16. See Poston 2012, which contains a proposal for emerging justification on the basis of Jonathan Kvanvig's INUS conditions.
[6] Ibid., 8. Michael Williams (Williams 2014, 234-235) has noted that the distinction between doxastic and propositional justification was introduced by Roderick Firth (Firth 1978). He recalls that Firth, too, claims that doxastic justification is parasitic on propositional justification, but argues that Firth attaches a completely different meaning to this claim than does Klein. As Williams sees it, Klein tries to combine an infinitist conception of propositional justification with a contextual conception of doxastic justification — a venture that, according to Williams, is doomed to failure (Williams 2014, 236-238).
[7] Bergmann 2007, 22-23.

> K_2: A belief B_i can be doxastically justified by being based on some other belief B_j only if B_j is itself doxastically justified.

and subsequently tries to catch Klein on the horns of a dilemma. Klein must either accept or reject K_2. If he rejects it, then he must maintain that a belief B_i can be doxastically justified by another belief B_j even if the latter is itself unjustified. This would turn Klein into a defender of what Bergmann calls *the unjustified foundations view* — an outlook that is not particularly Kleinian, to say the least. On the other hand, if Klein accepts K_2 along with K_1, then he would run the risk of becoming a sceptic. For then "he is committed to requiring for doxastic justification an infinite number of actual beliefs. ... But it seems completely clear that none of us has an infinite number of actual beliefs".[8]

The phenomenon of fading foundations points to an escape route out of this dilemma, for it shows that there is another way to reject K_2. If doxastic justification indeed draws on propositional justification, as Klein claims, then the justification that one belief gives to another also diminishes as the distance between them increases. That is to say, a belief B_1 can be doxastically justified by a chain of other beliefs, B_2, B_3, to B_n, such that:

1. each B_m is conditionally justified by B_{m+1}, where $2 \leq m \leq n-1$;
2. B_n may be justified by another belief, or may justify itself, or may be unjustified;
3. the effect of B_n on B_1 becomes smaller as n becomes bigger and bigger.

In the limit that n goes to infinity, the justificatory support given by B_n to B_1 vanishes completely. In that case it does not matter for the doxastic justification of B_1 whether B_n is justified or not: B_1 can still be doxastically justified. Klein and Bergmann are of course right that we cannot forever go on justifying our beliefs. But the phenomenon of fading foundations manifests itself already in chains of finite length. Often we need only a few links to observe that the influence of the foundational belief on the target belief has diminished considerably. Of course, we can only be sure of what we seem to be observing in a finite chain if there exists a convergence proof for the corresponding infinite series, and a proof that the remainder term goes to zero: there needs to be knowledge of what happens in the infinite case in order for us to be certain that what we see in the finite case is a robust phenomenon rather than a mere fluctuation. But as we have seen such a proof can be provided. Klein, too, argues that "rejecting K_2 does not entail endorsing an

[8] Bergmann 2007, 23. See also Bergmann 2014.

unjustified foundationalist view" (Klein 2007b, 28). His argument is different from ours, in that it refers, among other things, to a reason's availability. We however believe that our reasoning about fading foundations can capture Klein's most important intuitions, and we will come back to availability in the next chapter.

Let us sum up. In doxastic justification the choice is not between indefinitely going on and the unjustified foundations view. There is a third possibility, provided by what we know about infinite chains. Once we have recognized that any justification that B_n gives to B_1 diminishes as the distance between the two is augmented, we might decide to stop at B_n because the justificatory contribution that any further belief would bestow on B_1 is deemed to be too small to be of interest. When exactly a justificatory contribution is considered to be negligible depends on pragmatic considerations, but our two tables show that we are able to make these considerations as precise as we wish.

This third possibility goes unnoticed in the debate between Bergmann and Klein. Because the fact of fading foundations has not been taken into account, they fail to realize that the expression 'stopping at a belief B_n' can have more meanings than those that have been envisioned in the literature. It need not mean 'making an arbitrary move', as some coherentists have claimed. Nor need it imply that B_n is taken to be unjustified or self-justified. Rather, an agent can decide to stop at a belief B_n because she realizes that, for her purposes, B_{n+1} has become irrelevant for the justification of B_1. She finds that the degree of justification conferred upon B_1 by her beliefs B_2 to B_n is accurate enough, and she feels no call to make it more accurate by taking B_{n+1} into account. For her, the justificatory contribution that B_{n+1} gives to B_1 has become negligible, and with our tables she can precisely identify a point at which the role of B_n is small enough to be neglected, where we use the word 'justificatory' as before as meaning probabilistic support plus something else.

In this way we have given a more precise meaning to contextualist considerations that have been often expressed. For example Klein:

> The infinitist will take the belief that q to be doxastically justified for S just in case S has engaged in providing 'enough' reasons along the path of endless reasons. ...How far forward ...S need go seems to me a matter of the pragmatic features of the epistemic context.[9]

[9] Klein 2007a, 10.

> We don't have to traverse infinitely many steps on the endless path of reasons. There just must be such a path and we have to traverse as many as contextually required.[10]

And Nicholas Rescher:

> In any given context of deliberation the regress of reasons ultimately runs out into 'perfectly clear' considerations which are (contextually) so plain that there just is no point in going further. . . . Enough is enough.[11]

Our method differs however from what Klein and Rescher seem to have in mind. As we will explain in more detail in 5.3, where we argue for a view of justification as a kind of trade-off, the level of accuracy of the target can be decided upon in advance. Whether this level will be reached after we have arrived at proposition number three, four, sixteen, or more, depends on the structure of the series and on the chosen level. In no way does it depend on the question of how obvious proposition number three, four, sixteen, etc. is. Even if the proposition at issue is very obvious, and thus has a high probability, its contribution to the justification of the target might be small enough to be neglected. This is different from the contextualism of Klein and Rescher, according to which an agent stops when the next belief in the chain is sufficiently obvious and itself not in need of justification.

4.3 Emergence of Justification

It has been said that foundationalists and anti-foundationalists (that is coherentists and infinitists) conceive justification differently: the former gravitate towards an atomistic concept of justification, whereas the latter see it as a holistic notion.[12] Consequently, foundationalists regard justification as a property that can be *transmitted* or *transferred* from one proposition to another. The idea here is that justification somehow arises as a quality attached to a particular proposition, notably to the ground p, and then via inference is conveyed to the neighbouring proposition. The inferences themselves in no way affect the property that they transfer. They are just conduits, as McGrew and McGrew would have it, completely neutral in character, like wifi connecting two computers.[13]

[10] Ibid., 13.
[11] Rescher 2010, 47.
[12] Sosa 1980; Bonjour 1985; Dancy 1985.
[13] McGrew and McGrew 2008.

Anti-foundationalists, on the other hand, have a different outlook. For them justification is not a property that is transmitted from one link in the chain to another; rather it *emerges* gradually from the chain as a whole. In the words of Peter Klein:

> Foundationalists think of propositional justification as a property possessed autonomously by some propositions which, by inference, can then be transmitted to another proposition — just as a real property can be transmitted from one owner to another once its initial ownership is established. But of course, the infinitist, like the emergent coherentist, does not paint this picture of propositonal justification. . . . [T]he infinitist conceives of propositional justification of a proposition as emerging whenever there is an endless, non-repeating set of propositions available as reasons.[14]

> . . . the infinitist does not think of propositional justification as a property that is transferred from one proposition to another by such inference rules. Rather, the infinitist, like the coherentist, takes propositional justification to be what I called an emergent property that arises in sets of propositions.[15]

However, infinitists and coherentists experience great difficulty in explaining emergence. What exactly does it mean to say that justification emerges from a chain of propositions? How precisely does justification gradually arise from a chain or a web of beliefs? Champions of emergence illustrate their views by invoking arresting images, such as Neurath's boat or Sosa's raft. Although such metaphors are striking and helpful, they fail to inform us how exactly emergence can occur. It is one thing to claim that justification can emerge, but quite another to come up with a mechanism which explains how this can happen. Yet the latter is what we need. When emergence is called on to save the day for the anti-foundationalist, an account of the mechanism behind it ought to be specified in detail. Without such an account, emergence is in danger of being not much more than a name, and the appeal to it runs the risk of remaining gratuitous or *ad hoc*.

We believe that our concept of probabilistic support can help us here. For it carries with it the idea of fading foundations, which explains how justification can gradually emerge.[16] Look again at Table 4.1. It reveals the justification as it emerges from an infinite chain of reasons, and as a result we see the justification of q materializing in front of our eyes, as it were. The

[14] Klein 2007a, 16.

[15] Klein 2007b, 26.

[16] Frederik Herzberg also argues that our notion of probabilistic support can help explaining emergence (Herzberg 2013).

table enables us to give a precise interpretation of what Klein writes about justification as seen by infinitists (recall that for Klein doxastic justification is parasitic on propositional justification):

> ...the infinitist holds that propositional justification arises in sets of propositions with an infinite and non-repeating structure such that each new member serves as a reason for the preceding one. Consequently, an infinitist would seek to increase the doxastic justification of an initial belief – the belief requiring reasons – by calling forth more and more reasons. The more imbedded the initial belief, the greater its doxastic justification.[17]

Thus for Klein justification increases by lengthening the chain. A similar idea has been expressed by Jeremy Fantl:

> The infinitist [claims] that, for any particular series of reasons, the degree of justification can be increased by adding an adequate reason to the end of that series. Infinitism [claims]: ...the longer your series of adequate reasons for a proposition, the more justified it is for you.[18]

Our analysis can give a more precise meaning to these claims by Klein and Fantl. For it makes it clear that phrases like 'the emergence of justification' or 'the increase of justification' are in fact ambiguous. They can mean that, by adding more and more reasons, the value of the unconditional probability of q *becomes larger and larger*. But they can also mean that, by adding more reasons, the value of the unconditional probability of q *draws closer to its final value* (relative to the numbers chosen). It is the latter meaning that we are talking about here. In Table 4.1 it is the case that, every time we add an extra link to the chain, the probability of q rises until it reaches its maximum value. A rising value is however not essential for justification to emerge. This can be appreciated in Table 4.2, where the conditional probabilities are the same as those in Table 4.1, but where the unconditional probability of p is 0.95.

As in Table 4.1, in Table 4.2 the justification of q emerges as the number of A's gets bigger, for now q is, as Klein would say, more imbedded. However, it is not so that the probability of q rises with each step. As we

[17] Klein 2007b, 26.

[18] Fantl 2003, 554. Fantl defends infinitism on the grounds that, of all the theories of justification, it is best equipped to satisfy two requirements: the degree requirement ("a theory of the structure of justification should explain why or show how justification is a matter of degree") and the completeness requirement ("a theory of the structure of justification should explain why or how complete justification makes sense") — ibid., 538. That reasoning itself can generate justification has also been advocated by Mylan Engel (2014) and John Turri (2014).

add more and more reasons, the probability of q gets closer and closer to its final value, but numerically it goes down, namely from 0.935 to 0.8. Klein's phrase "[t]he more imbedded the initial belief, the greater its doxastic justification" or Fantl's phrase "the longer your series of adequate reasons for a proposition, the more justified it is for you" should therefore be properly interpreted. The phrases are correct under the interpretation: the longer the chain that justifies the target q, the more reliable the justification of q is, for the closer the unconditional probability of q is to its real value. What cannot be meant is: the longer the chain that justifies the target q, the greater the unconditional probability of q. The justification of q can ascend in reliability while the probability of q descends in numerical value. So we should be careful about what we mean when we say that justification emerges: we do not mean that the unconditional probability of the target proposition q necessarily increases numerically, rather we mean that this probability gradually moves towards its limit.

So far we have worked under the assumption that the values of $P(p)$ lay strictly between 0 and 1. Indeed, both Tables 4.1 and 4.2 respect this restriction. However, the assumption is neither necessary for fading foundations nor for the emergence of justification. The two tables below illustrate this point.

Table 4.3 Probability of q when the probability of p is 1

Number of A_n	1	2	5	10	25	50	75	100	∞
Probability of q	.981	.971	.947	.914	.853	.814	.804	.801	.8

Table 4.4 Probability of q when the probability of p is 0

Number of A_n	1	2	5	10	25	50	75	100	∞
Probability of q	.078	.114	.212	.345	.589	.742	.784	.796	.8

These tables are based on the same uniform conditional probabilities that we used before, that is 0.99 and 0.04. However, in Table 4.3 the unconditional probability of p is one and in Table 4.4 it is zero. They are extreme values, and admittedly they yield strange consequences. For example, if $P(p) = 0$, then p can scarcely be called a reason for q. And if $P(p) = 1$, then p cannot provide probabilistic support for any proposition (this is the root of the infamous problem of old evidence). Yet the tables reveal how ineffective the rôle of p is in the long run. For even with a $P(p)$ that is zero, the final probability of q is still 0.8; and justification can emerge when the foundation is nonexistent. Notwithstanding the extreme values of $P(p)$, the final probability of q is the same, and moreover the same as it was in Tables 4.1 and 4.2.[19]

In sum, we have argued that, in a probabilistic model of epistemic justification, justification is not something that one proposition or belief receives lock, stock and barrel from another. Rather it gradually emerges from the chain as a whole. As the distance between the source p and the target q *increases*, the influence of the unconditional probability of p on the unconditional probability of q *decreases*; in the limit of an infinite chain, the probability of q reaches its final value, and the only contributions to this value come from the infinite set of conditional probabilities. So when we go probabilistic, a law of diminishing returns goes hand in hand with a law of emerging justification: the more the justification of the final proposition materializes, the less is the influence of the grounding proposition.

4.4 Where Does the Justification Come From?

In a finite probabilistic chain, part of the justification comes from the ground and part comes from the conditional probabilities that connect the ground to the target. If the series is infinite, then *all* of the justification is carried by the conditional probabilities, and none by the ground. One might however still be puzzled as to whence the justification comes. If justification does not have its origin in a foundation, then where does it come from? How can we make sense of there being justification without a ground?

Most people agree that having justification somehow involves making contact with the world; as we said in Chapter 2, to call our beliefs justified means acknowledging that they at least remotely indicate how things actually are. If one takes the view that contact with the world requires a ground, and that a ground is apprehended by a basic belief, and that a basic belief involves an unconditional probability, then it is puzzling indeed how infinite chains can do the job. Such a view would however be unduly restrictive. It assumes that notions like ‘applying to the real world’, ‘outside evidence’

[19] If $P(p)$ is zero or one, some of the conditional probabilities are not well-defined according to Kolmogorov’s prescription. Alternative approaches to probability theory exist however, in which conditional probabilities are the basic quantities, and we will come back to this in the next section. The important point here is that if $P(p)=1$ then $P(A_m)=P(A_m|p)$, and $P(A_m|\neg p)$, which does not have a Kolmogorovian definition, is not needed as an ingredient in the regress. Similarly, if $P(p)=0$ then $P(A_m)=P(A_m|\neg p)$, and $P(A_m|p)$ is not needed in the regress.

and 'empirical results' only makes sense within a framework of basic beliefs. This is questionable, since conditional probabilities are just as well equipped to carry the empirical burden.

One might object that conditional probabilities are built up from unconditional ones, and that one can only determine their values on the basis of unconditional probabilities. Such a complaint has in fact been made by Nicholas Rescher:

> There is ... a more direct argument against the thesis that one can never determine categorical probabilities but only conditional ones. This turns on the fact that conditional probabilities are by definition no other than ratios of unconditioned ones $P(q|p) = P(q\&p)/P(p)$. So unless conditional probabilities are somehow given by the Recording Angel they can be only be determined (or estimated) via our determination (or estimation) of categorical probabilities. And then if the latter cannot be assessed, neither can the former.[20]

It is true that, within standard probability theory, conditional and unconditional probabilities can be defined in terms of one another. It is also true that Kolmogorov himself saw the unconditional probabilities as the basic elements. However, three considerations should be taken into account here. First, one is free to make another choice, and many philosophers have done so. Rudolf Carnap, Karl Popper, Alan Hájek — they all plump for conditional probabilities as the more useful basic quantities. In fact taking conditional probabilities as primary has certain advantages: one can cover extreme cases that cannot be handled if unconditional probabilities are regarded as being fundamental.[21] Second, we have not claimed that unconditional probabilities can *only* be estimated via infinite regresses involving conditional probabilities: rather we have shown that they *can* be computed in that way. Third and most important, there is no objection whatever to questioning the conditional probabilities in turn. Up to this point we have considered them as being given, but that is only a pragmatic stance, motivated by expository considerations. It is perfectly possible to unpack the conditional probabilities and consider them as targets that are themselves justified by further probabilistic chains. This possibility will be briefly touched upon in Section 6.4

[20] Rescher 2010, 40, footnote 18 (we adapted Rescher's notation to ours).

[21] Carnap 1952; Popper 1959; Hájek 2011. Hájek mentions more philosophers who made this choice: De Finetti 1974/1990; Jeffreys 1939/1961; Johnson 1921; Keynes 1921; Rényi 1970/1998. One can define $P(q|p)$ as $P(q \wedge p)/P(p)$ only if $P(p) \neq 0$. If one adopts this Kolmogorovian definition, one is unable to make sense of $P(q|p)$ when $P(p) = 0$. The approach of the philosophers mentioned above is free from this difficulty.

and further explained in Section 8.5. But for the moment we ignore this refinement.

Two final worries remain. First, how do we know that the conditional probabilities in our chain are ‘good’ ones, i.e. make contact with the world? What is the difference between our reasonings and those occurring in fiction, in the machinations of a liar, or in the hallucinations of a heroin addict? Or, applied to our example about bacteria, how can we distinguish the regress concerning Barbara and her ancestors from a fairy tale with the same structure in which, instead of the inheritable trait T, there is an inheritable magical power, M, to turn a prince into a frog?

The distinction is not far to seek. It lies in the mundane fact that in the former, but not in the latter, the conditional probabilities arise from observation and experiment. Research on many batches of bacteria have established the relevant conditional probabilities, α and β. These conditional probabilities are typically obtained by repeated experiments: they are measured by counting how many ‘successes’ there are in a given number of trials, and then by dividing one number by the other (e.g. the number of bacteria that carry a trait, divided by the total number of bacteria in a sample). In the fairy tale, on the other hand, the only ‘evidence’ that M is inheritable is contained in the story itself — outside the tale there is no evidence at all. When it comes to series of infinite length, conditional probability statements are the sole bearers of the empirical load. Together they work to confer upon the target proposition an unconditional probability that expresses the proposition’s degree of justification. It is by virtue of the conditional probabilities that an infinite chain is not just an arbitrary construct that displays mere coherence, but rather can provide real justification, albeit of a probabilistic character.

We realize perfectly well that this answer will not convince the confirmed sceptic, but our opponent after all is a particular kind of foundationalist, not the sceptic. We do not have the temerity to aim at refuting the claim that all our perceptions might be illusory, or at outlawing evil demon scenarios, old and new. We simply assume that there is a real world, and that empirical facts can justify certain propositions, or more generally can sanction the probabilities that certain propositions are true. Here we merely take issue with any foundationalist claim to the effect that only basic beliefs or unconditional probabilities can be candidates for connecting world and thought.

That brings us to the second worry. A foundationalist might not be persuaded by the above considerations, arguing that the erstwhile rôle of the basic belief is now being played by the set of conditional probabilities. Indeed, he might claim that we are worse off, for we seem to have traded one basic belief, viz. the remote starting point of the epistemic chain, for an infinite number of conditional probability statements.

We do not want to get involved in a verbal dispute here: we are not objecting to a type of foundationalism that acknowledges the empirical thrust of conditional probabilities as well as the importance of fading foundations. This should not blind us, however, to the difference between conditional probabilities and the traditional basic beliefs. The former are essentially relational in character: they say what is to be expected if something else is the case. The latter are by contrast categorical: they say that something is the case, or that something can be expected with a certain probability. There is a great difference between averring that 'A_n is true' (or that the probability of A_n is large) on the one hand, and holding that 'if A_{n+1} were true, the probability that A_n is true would be α', or 'if A_{n+1} were false, the probability that A_n is true would be β' on the other hand. Conditional probability talk is discourse about relationals and hypotheticals. Our use of an infinite number of conditional probabilities amounts to the introduction of an infinite number of relational statements. If all these statements satisfy the condition of probabilistic support as defined earlier, they can give rise to something that is no longer relational, but categorical. This categorical statement can in turn become the starting point of a new series of relational statements. And if this new series becomes sufficiently long, the influence of the categorical might die out, as we have seen.

The situation is somewhat comparable to what happens in science or in logic.[22] Scientists typically construct mathematical models on the basis of empirical input, and then employ these models to draw new conclusions about the world. Similarly, logicians make inferences on the basis of premises that contain empirical information, thus producing new conclusions as output. In both cases, the output can in turn become the input for other models and inferences. And in neither case can the machinery work without input: logicians need their premises and scientists need their data. Since every assumption that serves as input can itself be questioned in turn, there is in this sense a foundation behind every foundation. One may interpret that as support for foundationalism ('there is always a foundation!') or as support for anti-foundationalism ('every foundation is a pseudo-foundation!'). Rather than let ourselves be drawn into such a debate, it might be more fruit-

[22] Gijsbers 2015; Bewersdorf 2015.

ful to see what actually happens. And what happens is that a foundation becomes less important as it recedes from the target.[23]

4.5 Tour d'horizon

Let us take stock. The epistemological regress problem, as we have introduced it in Chapter 1, led to a discussion of epistemic justification in Chapter 2. The idea that epistemic justification has something to do with 'probabilification' is widespread among contemporary epistemologists: practically all agree that 'A_j justifies A_i' at least implies that A_i is made probable by A_j. Yet, as we have been arguing in Chapters 3 and 4, the far-reaching consequences of this unanimity about the regress problem in epistemology have been insufficiently understood.

A few exotic cases excluded, talk about probability is Kolmogorovian talk. One of the theorems of Kolmogorov's calculus is the rule of total probability, which enables us to determine the unconditional probability of q, namely $P(q)$. If $P(q)$ is made probable by an epistemic chain rather than a single proposition, then the value of $P(q)$ is obtained from an iterated rule of total probability. It has often been thought that such an iteration does not make sense if it continues indefinitely, but, as we have seen in Chapter 3, this is simply a mistake. In all but the exceptional cases $P(q)$ can be given a unique and well-defined value, even if the chain that supports it is infinitely long.

The iteration in question is a complex formula that consists of two parts. The first part is a series involving all the conditional probabilities, the second part is what we have called the remainder term, which contains information

[23] The phenomenon of fading foundations is not restricted to probabilistic chains in epistemology; it can be proved (although we will not do that here) that it also applies in modified form to infinite chains of propositions that are ranked in the sense of Spohn (Spohn 2012). Moreover, fading foundations occur in non-epistemic causal chains, as long as 'causality' is interpreted probabilistically. This fact may shed light on various philosophical debates, such as the one on rigid designators, i.e. expressions that denote the same object in every possible world. The objects themselves, at least for Saul Kripke, are identified by following causal chains backwards to the moment of baptism when they received their names. Gareth Evans noted a problem with this view: we can use proper names even if the causal chains are broken (his Madagascar-example in Evans 1973). In Addendum (e) to *Naming and Necessity* Kripke comments that he leaves this problem "for further work" (Kripke 1972/1980, 163); but with a probabilistic conception of causality Evans' problem disappears since the rôle and character of rigid designators change.

about the probability of the grounding proposition. What in this chapter we have called fading foundations arises if and only if the following two requirements are fulfilled:

1. the series involving the conditional probabilities converges
2. the remainder term goes to zero.

The first requirement is always fulfilled if the condition of probabilistic support has been satisfied for the entire chain; that is, if $P(A_i|A_j) > P(A_i|\neg A_j)$ for all the links. The second requirement is only fulfilled if we are dealing with what we have been calling the usual class, i.e. the class of probabilistic regresses that are benign. Informally, this means that the conditional probabilities must not tend too quickly to those appertaining to an entailment. Formally, it means that they comply with

$$\exists c > 0 \ \& \ \exists N > c : \ \forall n > N, \ 1 - \gamma_n > \frac{c}{n} \, .$$

Whereas conditional probabilities that obey this constraint belong to the usual class, those that violate it make up the exceptional class. The latter we also call the class of *quasi-bi-implication*. The conditional probabilities in this class resemble bi-implications, and they fail to meet the above asymptotic constraint. From this it follows that whenever we are dealing with a probabilistic regress in which the conditional probabilities are of the usual class, fading foundations will ensue. Indeed, the necessary and sufficient condition for fading foundations is membership of the usual class.

Despite the technicalities we needed to prove it, the result itself is actually very intuitive. If the conditional probabilities in a regress are very close to those corresponding to entailments, then we can only determine the truth value of the target if we know the truth value of the ground. Irrespective of the chain's length, and thus irrespective of whether the ground is very close to the target or is far removed from it, the ground continues to make a contribution, and then the age-old regress problem rears its ugly head. But if the regress contains genuine conditional probabilities, i.e. conditional probabilities that do not resemble implications, then the remainder term goes to zero, and the regress is benign.

Strictly speaking, as we noted in Chapter 3, footnote 29, in the usual class probabilistic support is not needed for convergence. But probabilistic support is important for three reasons. First, we are interested in epistemic justification, and this contains probabilistic support as a necessary element. Whatever it may mean to say that 'A_j justifies A_i', part of its meaning is that $P(A_i|A_j) > P(A_i|\neg A_j)$. Second, we like to see epistemic justification as something that amounts to striking a balance. In justifying our beliefs, we set up a trade-off between the number of reasons that we can handle with our finite minds and the level of accuracy that we want to reach. As we will explain in the next chapter, probabilistic support is needed for such a view of justification as a trade-off. Third and finally, the condition of probabilistic support is needed for the convergence of the networks that we discuss in Chapter 8.