HOW DID FREGE FALL INTO THE CONTRADICTION?

Peter M. Sullivan

Abstract

Quine made it conventional to portray the contradiction that destroyed Frege's logicism as some kind of act of God, a thunderbolt that descended from a clear blue sky. This portrayal suited the moral Quine was antecedently inclined to draw, that intuition is bankrupt, and that reliance on it must therefore be replaced by a pragmatic methodology. But the portrayal is grossly misleading, and Quine's moral simply false. In the person of others – Cantor, Dedekind, and Zermelo – intuition was working pretty well. It was in Frege that it suffered a local and temporary blindness. The question to ask, then, is not how Frege was overtaken by the contradiction, but how it is that he didn't see it coming. The paper offers one kind of answer to that question. Starting from the very close similarity between Frege's proof of infinity and the reasoning that leads to the contradiction, it asks: given his understanding of the first, why did Frege did not notice the second? The reason is traced, first, to a faulty generalization Frege made from the case of directions and parallel lines; and, through that, to Frege's having retained, and attempted incoherently to combine with his own, aspects of a pre-Fregean understanding of the generality of logical principles.

1

For all that has been written about it, it really is very puzzling that Frege should have fallen into the contradiction of his Basic Law V. It is puzzling, in particular, because what makes that Law contradictory[1] is so close to something that Frege very clearly understood.

One might present what Frege understood as having to do with a certain kind of process:

> (***Process***) Start with some range of objects_0; collect these objects_0 under concepts_1 according as they are equivalent in some respect; now collect together these concepts_1 according as ***they*** are equivalent in some respect; and finally, count each such collection_2 as an object_0.

One important instance of this process is:

> (***Numbers***) Start with a range of objects_0, collected in all the standard ways under concepts_1; collect together concepts_1 that are equivalent in respect of how many of the objects_0 they apply to; and count each such collection_2 as an object_0.

[1] On a very standard, and perhaps therefore a superficial, diagnosis. Others, most memorably George Boolos (1996) and Michael Dummett (1996), have discussed whether the contradiction has a deeper source, and if so what it might be. Theirs is a substantive concern: they aim to diagnose the contradiction itself. I am concerned here with a question that is more historical than substantive, and aim only to diagnose Frege's falling into the contradiction.

If you assume this process to begin with a range of $\boldsymbol{n}$ objects$_0$, it is very easily seen that there will be $\boldsymbol{n}+1$ collections$_2$. So the process requires that $\boldsymbol{n}+1 \leq \boldsymbol{n}$. So if your initial assumption was of an $\boldsymbol{n}$ for which this does not hold – if you assumed the initial range of objects to be finite – you were wrong. Frege of course knew this very well: it is just his proof of the infinity of the number series, casually written.

Some years later Russell, with a lot of help from Cantor, was reflecting on another instance of the process:

> (***Classes***) Start with a range of objects$_0$, collected in all the standard ways under concepts$_1$; collect together concepts$_1$ that apply to the same objects$_0$; and count each such collection$_2$ as an object$_0$.

It is just as easily seen that if this process begins with $\boldsymbol{n}$ objects$_0$ it yields $2^{\boldsymbol{n}}$ collections$_2$. So this process is feasible only if $2^{\boldsymbol{n}} \leq \boldsymbol{n}$, something that holds for no value of $\boldsymbol{n}$.

In devising his infinity-proof, then, Frege realized something that might be expressed as follows.

> (***Frege***) Some instances of the described process are disruptive of ***some*** (many) initial assumptions about the size the range of objects.

What Russell later realized and communicated to Frege is then just a word away:

> (***Russell***) Some instances of the described process are disruptive of ***all*** assumptions about the size the range of objects.

When the disastrous truth was so close, how is it that it never so much as entered Frege's considerations? For surely, if the thought here called (***Russell***) had so much as flittered through his reflective mind, he could never, understanding (***Frege***) as well as he did, have advanced Basic Law V.

Hume remarked that ideas are apt to induce others they resemble. (***Russell***) certainly resembles (***Frege***). But in Frege's mind sustained reflection on (***Frege***) did not induce the idea (***Russell***). Something must have blocked it out. Discovering what that something was would offer one kind of explanation of how Frege fell into contradiction.

In broad outline, of course, we already know what it was. Frege saw (***Frege***) as a consequence of the legitimacy of (***Numbers***), and he saw (***Numbers***) as legitimate because its dominant association in his mind was with a range of undoubtedly sound principles, principles that identify directions as what parallel lines have in common, heights as what things equally tall have in common, shapes as what similar figures have in common, and so on. The mental mechanics were so set up that the combined pull of all those sound principles effectively severed any associative connection on the other side. What we need to do, then, is to fill this outline in: shifting metaphors, we need to explain why it is that Frege saw (***Numbers***), and then by association (***Classes***), always in the reflected light of all those sound principles, and never under the cloud cast by (***Russell***).

2

It would be nice if we could turn to Frege's discussion of (***Numbers***) for the explanation. Notoriously, though, Frege's most extended discussion of (***Numbers***) is not explicitly a discussion of (***Numbers***) at all. Instead it is a discussion of one of those sound principles he took it to resemble, namely:

(***Directions***) The direction of line $\boldsymbol{a}$ = the direction of line $\boldsymbol{b} \leftrightarrow \boldsymbol{a}$ is parallel to $\boldsymbol{b}$.

The core idea Frege is defending in that discussion is that we can be brought to acknowledge a new kind of thing, here a direction, as the kind of thing that items we already acknowledge have in common when they stand in some equivalence relation, here parallelism. Everything it takes for there to be this new kind of thing is already required to obtain by the judgements we already make about the old kind of item. And to recognize that, all we have to do is to re-articulate those judgements. As Frege puts it:

> The judgement
>
> $$\boldsymbol{a} \,//\, \boldsymbol{b}$$
>
> can be taken as an identity. If we do this, we obtain the concept of direction, and say: 'the direction of line $\boldsymbol{a}$ is identical with the direction of line $\boldsymbol{b}$'. Thus we replace the symbol // by the more generic symbol =, through removing what is specific in the content of the former and dividing it between $\boldsymbol{a}$ and $\boldsymbol{b}$. (1884, §64)

Now evidently this core idea has nothing in particular to do with lines and directions: Frege's indirect strategy of argument would be hopelessly inept if it did. So we have to take it that he is defending a more general commitment.

The obvious way to begin to recover that general commitment is simply to quantify out the mention of lines and directions:

(***1^st^ Level Abstraction***) $\text{Equiv}(\boldsymbol{R}) \rightarrow \exists \boldsymbol{f}\, \forall \boldsymbol{x}\, \forall \boldsymbol{y}\, (\boldsymbol{R}(\boldsymbol{x},\boldsymbol{y}) \leftrightarrow \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{y}))$

Frege clearly believes this, and it is clearly true.[2] Just as clearly, though, this general principle can't have been what Frege was defending, because the case he is really interested in is not an instance of it. In that case, (***Numbers***), a judgement that a second-level equivalence relation holds between concepts is to be re-carved so as to yield a new kind of

---

[2] To see that just let '$\boldsymbol{f}$' mean, roughly, 'my favourite…', so that '$\boldsymbol{fx}$' picks out my favourite amongst the objects that are $\boldsymbol{R}$ to $\boldsymbol{x}$. (Or if, thinking of the likeness of socks, you doubt that I can always have a favourite, let it mean 'God's favourite'.)

thing, a number, that those concepts have in common. That would be an instance of a different order of principle:

$$(\textbf{\textit{2}}^{\textbf{\textit{nd}}}\ \textbf{\textit{Level Abstraction}})\ \ \text{Equiv}(\Psi) \rightarrow \exists\Phi\ \forall \boldsymbol{F}\ \forall \boldsymbol{G}\ (\Psi(\boldsymbol{F},\boldsymbol{G}) \leftrightarrow \Phi(\boldsymbol{F}) = \Phi(\boldsymbol{G})).$$

But now we can be sure (even ignoring its falsehood) that this wasn't the principle that Frege was aiming to defend either; because the case he actually discusses, that of lines and directions, is not an instance of this one, and that again would make his strategy of argument inept. Frege's discussion makes sense only if the general commitment he was defending was something more general than either of these: it cannot be either of the type-bound generalizations just canvassed, but has to be a 'type-crossing' principle, to the effect that that an equivalence relation of ***whatever*** level can ***always*** be taken as an identity.

3

Someone might protest at this point, saying: 'It can't have been Frege's intention to defend such a principle, since it is utterly central to his logical thinking that there are no such principles to defend: what can be said of an object simply cannot be said of a concept, so there are no general principles that hold equally of both'. The premise of this protest is true. Unfortunately its conclusion is false. Frege ***was*** fundamentally committed to the impossibility of 'type-crossing' generalizations; all the same, he was also committed to various generalizations of that kind. Tactically it is best to demonstrate this point while steering clear of the more contentious regions of his thought; so at this stage I will take straightforward logical principles, rather than obscure philosophical ones, as illustrations of it.

(i) In ***Begriffsschrift*** Frege has a single axiom governing the quantifier, which reads '$\forall \boldsymbol{xFx} \rightarrow \boldsymbol{Fc}$', or 'what holds of everything holds of anything'. At several points in the proofs in Part III he cites this axiom in justification of second-order inferences, prompting critics to complain that he has used a first-order principle when what he needs is a second-order one. That is surely a mistake. Frege nowhere in ***Begriffsschrift*** explicitly states that his axiom is a first-order principle. So we have no clearer guide to the kind of principle he took it to be than the inferential use he makes of it. And that use dictates that it is a principle ***equally*** applicable in first- and second-order inference, hence a type-crossing principle.

(ii) Although the terminology of 'levels' or 'ranks' occurs already in ***Begriffsschrift*** (1879, §9), it is fair to observe that Frege's thinking on these matters is, in his first book, relatively unformed. In the mature system of ***Grundgesetze*** the single axiom gives way to a pair of axioms, one governing the first-order and one the second-order quantifier. That might lead one to think that Frege has rejected the idea of a 'type-crossing' principle as a mistake of his early work. Before we accept that conclusion, though, we should ask: Why just the two? This is, Frege explains, a convenience brought about by the introduction of value-ranges (or classes) in the later system: instead of having to quantify over higher-level functions or concepts, one can generalize instead about the associated value-ranges or classes.

> One might suppose that this would not nearly suffice; but we shall see that we can make do with this… It may be briefly observed that this economy is made possible by the fact that second-level functions can be represented in a certain manner by first-level functions, whereby the functions that appear as arguments of the former are represented by their courses of values. (1893, §25)

The currently relevant feature of this explanation is this: Frege does not actually formulate the third-order axiom he is talking about; but he clearly knows, and knows that his readers will know, which principle he is talking about, and saying that he does not need to implement it in an axiom. ***Which*** principle is it? And ***how*** does he know, and expect his readers to know, which principle it is? The principle is the third-order instance of the 'type-crossing' principle he had advanced in ***Begriffsschrift***. Frege expects his readers to know that, because he expects them to understand what his two axioms have in common, and what they each would have in common with axioms higher up the hierarchy, even though his official thinking by this time clearly insists that they can have nothing in common.

(iii) Between these two books comes ***Grundlagen***, where our discussion occurs. Earlier in this book Frege says:

> [it is] one of the requirements of reason [that it] must be able to embrace all first principles in a survey. (1884, §5)

If generalization at each of the infinitely many levels of Frege's hierarchy calls for a logically distinct principle, and if these infinitely many principles cannot be taken in at once, as instances of a more general principle, then the demand formulated here cannot be met. In endorsing the demand Frege very clearly indicates that this is not how things stand: evidently he thinks that he ***can*** take in all of those instances at once, however much his official thinking excludes that possibility.

4

With the protest deflected we can try to formulate what Frege's discussion of lines and directions is actually supposed to demonstrate. The obvious way to do that is by first explicitly indexing the expressions of (***1st Level Abstraction***), to yield

(***Indexed Abstraction 1***) $\text{Equiv}_2(\boldsymbol{R}_1) \rightarrow \exists \boldsymbol{f}_1\ \forall \boldsymbol{x}_0\ \forall \boldsymbol{y}_0\ (\boldsymbol{R}_1(\boldsymbol{x}_0,\boldsymbol{y}_0) \leftrightarrow \boldsymbol{f}_1(\boldsymbol{x}_0) = \boldsymbol{f}_1(\boldsymbol{y}_0))$,

and then generalizing through the indices:

(***Generalization 1***) $\text{Equiv}_{n+1}(\boldsymbol{R}_n) \rightarrow \exists \boldsymbol{f}_n\ \forall \boldsymbol{x}_{n-1}\ \forall \boldsymbol{y}_{n-1}\ (\boldsymbol{R}_n(\boldsymbol{x}_{n-1},\boldsymbol{y}_{n-1}) \leftrightarrow \boldsymbol{f}_n(\boldsymbol{x}_{n-1}) = \boldsymbol{f}_n(\boldsymbol{y}_{n-1}))$.

For discussion to proceed we have to concede to Frege's unofficial thinking that we understand this. But that is not much of a concession: a good reason for allowing that we understand it is that we know it to be false. The relevant falsifying instance is, of course,

$$\text{Equiv}_3(\text{Coextensiveness}_2) \rightarrow \exists\Phi_2\ \forall \boldsymbol{F}_1\ \forall \boldsymbol{G}_1\ (\text{Coextensive}_2(\boldsymbol{F}_1,\boldsymbol{G}_1) \leftrightarrow \Phi_2(\boldsymbol{F}_1) = \Phi_2(\boldsymbol{G}_1)),$$

whose detachable consequent is just the contradiction.

5

From this point someone could jump to an easy conclusion. Frege ran into ***the*** contradiction, they'd say, because of a broader contradiction in his thinking about logic. His official logical thinking introduced a hierarchical structuring of the principles of logic, one which altogether excludes there being any intelligible type-crossing principles. Even so, he imagined that, in his consideration of (***Directions***) and its ilk, he had latched onto a principle of precisely that kind. If he'd consistently adhered to the limitations his official thinking imposes he would have appreciated that the maximal generalization possible from (***Directions***) is (***1st Level Abstraction***), an innocent truth. Instead, he supposed himself able to transcend those limitations, and in (***Generalization 1***) paid the price for that inconsistency.

That is not, I think, wholly wrong. But it is partly wrong, and what remains is not very helpful. Dummett has already given an explanation of Frege's vulnerability to the contradiction that blames his 'amazing insouciance' (1991, p. 218) about the second-order quantifier, which led him to imagine that sound principles of first-order reasoning could be simply shifted up a level to yield further sound principles. What is right in the current diagnosis is not very helpful because it differs from that one mainly by being very much less detailed and developed. But I think there is more illumination to be had by considering where the current diagnosis is wrong. It suggests that Frege opened himself to the contradiction by taking (***1st Level Abstraction***) to instance a sound type-crossing principle. But that, as such, is not a mistake: it does. Frege's mistake was not simply that he generalized beyond (***1st Level Abstraction***), but that he generalized in the wrong way.

We represented Frege's generalizing move by indexing and then generalizing through these indices. Like Frege, though, we didn't make a thorough job of it. What we should have done gives first, as the indexed version of (***1st Level Abstaction***):

(***Indexed Abstraction 2***) $\text{Equiv}_2(\boldsymbol{R}_1) \rightarrow \exists \boldsymbol{f}_1\ \forall \boldsymbol{x}_0\ \forall \boldsymbol{y}_0\ (\boldsymbol{R}_1(\boldsymbol{x}_0,\boldsymbol{y}_0) \leftrightarrow [\boldsymbol{f}_1(\boldsymbol{x}_0)]_0 =_1 [\boldsymbol{f}_1(\boldsymbol{y}_0)]_0).$

Generalizing through the indices of this more thoroughly indexed claim then gives:

(***Generalization 2***) $\text{Equiv}_{n+1}(\boldsymbol{R}_n) \rightarrow \exists \boldsymbol{f}_n\ \forall \boldsymbol{x}_{n-1}\ \forall \boldsymbol{y}_{n-1}\ (\boldsymbol{R}_n(\boldsymbol{x}_{n-1},\boldsymbol{y}_{n-1}) \leftrightarrow [\boldsymbol{f}_n(\boldsymbol{x}_{n-1})]_{n-1} =_n [\boldsymbol{f}_n(\boldsymbol{y}_{n-1})]_{n-1}).$

And this, the true type-crossing generalization of (***1^st^ Level Abstraction***), is as entirely innocent as (***1st Level Abstraction***) itself.[3]

If the types of ***all*** the expressions in (***1st Level Abstraction***) are shifted together, that is, no harm results. So what we need to understand, to understand how Frege's type-shifting landed him in the contradiction, is why, in that process, he held some of the types fixed. In other words, why did he see (***1st Level Abstraction***) as an instance of (***Generalization 1***) rather than (***Generalization 2***)?

6

A first, disappointing suggestion would simply observe that (***Generalization 1***) gives Frege what he wanted while (***Generalization 2***) does not. The point of (***Process***) is that, while it rises through levels as a ***means of finding*** new entities, the entities it finds are reckoned to belong to the initial level, where they can feed again into it. Frege's infinity proof and the contradiction are generated in exactly the same way by this feature. So, aiming to admit the one, he endorsed a generalization that admits the other. Now it is no doubt true that the prospect of success can hide from someone the obstacles in its way. Even so, this first answer does not help us to understand how Frege might have been thinking. In his explicit reasoning he was about as far as anyone could be from endorsing a 'regressive' method, one that adopts first principles on the basis of the desirable consequences they yield. That is not to say that his unformulated thinking could never have taken that shape. But it does imply, I think, that the hypothesis that it did so amounts to little more than the hypothesis that he made unmotivated mistake. It does not begin to explain how the mistake might have seemed to him not to be one.

A rather better suggestion notes that in the route to (***Generalization 1***) it is an ***identity*** formula that goes un-indexed, and that the identity relation is special to objects: 'only in the case of objects can there be any question of equality (identity)' (1979, p.182); 'the relation of equality between objects cannot be conceived as holding between concepts too' (1979, pp. 121-2). This is a point Frege makes frequently. But he often accompanies the point by explaining that there will be 'a corresponding relation' for entities of other types (1979, p. 122), the correspondence lying in its satisfaction of the (level-adjusted) laws of identity. Frege's unofficial thinking generalizes easily across these different relations, for instance in his endorsement of extensionalism and in the associated principle that 'sameness' of reference will guarantee inter-substitutability. (***Generalization 2***) would have been just another case of that kind.

---

[3] In it '***f***' can again mean something like 'my favourite…', so that, whenever items including ***x*** are grouped by an equivalence relation, '***fx***' will pick out my favourite among the items grouped with ***x***. Whatever the level of ***x***, there will always be such an item. Instead of the contradiction we have, from the relevant instance of (***Generalization 2***), a tautology,

$$\exists\Phi_2\ \forall \boldsymbol{F}_1\ \forall \boldsymbol{G}_1\ (\text{Coexstensive}_2(\boldsymbol{F}_1,\boldsymbol{G}_1) \leftrightarrow [\Phi_2(\boldsymbol{F}_1)]_1 =_2 [\Phi_2(\boldsymbol{G}_1)]_1),$$

I think we get a better answer by looking to the expressions that flank the identity sign. When the identity sign is indexed, these flanking terms also have to be indexed, and the result looks distinctly un-Fregean. In the notation I used in (***Generalization 2***), '$[\boldsymbol{f}_{\mathrm{m}}(\boldsymbol{x})_{\mathrm{n}}]_{\mathrm{p}}$', the type of $\boldsymbol{f}$ (m) is marked as that of a function from items of the type of its argument $\boldsymbol{x}$ (n) to items of the type of its value (p). But in Frege's scheme this is more typing than a function needs: for him the type of a function is settled by that of its argument(s) alone; we do not need to mention as well the type of the value, since for him ***any*** function has as its value an object. A correct generalization capturing what Frege took to be illustrated by (***Directions***) would be one that assures us, for any equivalence of whatever level, of the existence of functions mapping items thus equivalent to the 'same' entity. Given what kinds of functions Frege recognized, he would take that to be an assurance that there exist such functions with ***objects*** as values; that is, he would take it to be (***Generalization 1***) and not (***Generalization 2***).

7

So the question becomes: Why did Frege think only of functions of this restricted kind?

The way mathematicians typically talk of functions portrays them as a kind of ontological add-on. When for instance they say, 'Let $\boldsymbol{f}$ be any function from such-and-suches to so-and-soes', it's presumed we have a scheme of things (including such-and-suches and so-and-soes) already set out and categorized; the functions are then just laid over the top to give handy ways of getting around the scheme. On this way of thinking there is no barrier to introducing functions from any (type of) thing to any (type of) thing, and Frege's restriction seems completely unmotivated. To motivate that restriction we have to turn this typical way of thinking on its head. For Frege the recognition of functions is not consequent on some already settled ontological scheme. Instead, it is by discerning functions of various kinds that the ontological scheme is set up in the first place.

Frege's way of thinking of functions is introduced in the section of ***Begriffsschrift*** that that encapsulates his contribution to logic.

> If, in an expression (whose content need not be assertable), a simple or a complex symbol occurs in one or more places and we imagine it as replaceable by another [symbol] (but the same one each time) at all or some of these places, then we call the part of the expression that shows itself invariant [under such replacement] a function and the replaceable part its argument. (1879, §9)

I don't think it would be too much of an exaggeration to hold that all of Frege's philosophy of logic should be traced to this passage. For current purposes, though, there are three main ideas to highlight, all of them consequences of the way the notion of a function is explained here, and all bound up with each other.

(i) First, the way functions are to be recognized on this model imposes a hierarchical ordering on them. We can extract a first-level function from a sentence (or judgement) only

---

in which, given Frege's extensionalism about concepts, we can take Φ to be the trivial mapping that takes any concept to itself.

if we already recognize some constant in the sentence as belonging to a range of alternatives by which it might be replaced. Taking that first-level function in turn to belong to a range of alternatives then allows the extraction of a second-level function. And so on. On this model there is simply no seeing what a higher-level function is supposed to be without seeing it as dependent on an already settled range of lower-level functions, this chain of dependence resting at bottom on the range of objects.

(i) Second, the ***incompleteness*** or ***unsaturatedness*** of functions is a matter of their having been extracted in this way from something complete, hence a matter of their needing to be ***completed by*** their arguments. It is for this reason that function and argument together cannot be taken to yield something that remains incomplete, or in other words why the values of functions have to be objects.[4]

(iii) Third, a way of analysing something into function and argument in accordance with this model is always one of many alternative analyses of it. For reasons I can only sketch here, I think it is this feature of his function–argument analysis that allowed Frege to hold that the logic he built on this model yields genuine knowledge of objects.[5]

On a pre-Fregean conception logic secures its generality by concerning itself only with the forms, and not the contents, of judgements (Kant, A54/B78). To that older way of thinking the horizontal interconnections between forms of judgements that interest logic are independent of the vertical connections to reality that supply real content to fill out those forms. This independence leaves space for a general metaphysical issue, of how any internal, rationally certifiable feature of thought can be sufficient to guarantee it any genuine content. One centrally important way of framing that issue exploits a contrast often drawn in discussions of Kant's philosophy between, on the one hand, a thin, logical notion of an object, as a subject of predication, an occupant of a distinguished place in the forms studied by logic, and on the other a rich, metaphysical notion of an object, as a common and therefore unifying focus of distinct representations, something that can be approached from different angles. It is then asked how thinking, whose internal logical features sustain only the thin notion, can be guaranteed to connect with objects in that richer sense.

---

[4] Dummett calls this the 'principle…of the completeness of the values of a function' (1993, p. 293). Elsewhere he gives a fuller explanation.

> Suppose one wants to construct an expression for a function from objects to first-level concepts. One will therefore start with an expression for a first-level concept, and remove from it (one or more occurrences of) some proper name. What one will be left with, however, is simply an expression with two argument places, which will stand for a first-level relation; there is no way to indicate that one of the argument places is to be filled first, and no point in doing so. (1998, p. 93)

Adapting the argument to the case in hand, consider a sentence in which there is supposed to occur a functional expression of the type $[\Phi_2(\boldsymbol{F}_1)]_1$, say 'ξ is the opposite of red' extracted from its application to a particular green traffic light in 'That light is the opposite of red'. If in this functional expression we imagine 'red' as replaceable, we reach the unequal-levelled relation, 'ξ is the opposite of φ'. Filling only the first argument place by 'that light' gives a second-level property true of any first-level property whose opposite is had by that light; filling only the second by 'red' gives a first-level property true of green things; but neither of these fillings completes, nor therefore yields the value of, 'ξ is the opposite of φ'.

[5] The following sketch is developed somewhat in §4.1 of Sullivan 2004.

Now interpreters as different in other ways as Dummett and Ricketts[6] agree in holding that Frege's logical thought simply leaves no room for that kind of question, and that the reason for that is to be traced to his quantificational understanding of the generality of logic. Their view seems to be, though, that Frege worked exclusively with a logical notion of an object, and that the richer metaphysical notion of an object plays no role at all in his thinking. My own view is that Frege attempted a genuine fusion of these two notions,[7] and that his theory of the multiple analysability of judgements is the key to that attempt. On that theory any particular articulation of a judgement, so any way of seeing it as an instance of some generalization, is just one of many ways that judgement can be articulated. Recognizing the inferential powers of a judgement will in interesting cases involve recognizing more than one of those articulations. Frege's objects are the pivots around which these different articulations ultimately turn, and logical inference itself demands that they be recognized as the ***common*** pivots of ***different*** articulations. To take Frege's simplest instance, in articulating the judgement that Cato killed Cato in the various ways that Fregean analysis provides for we present the object Cato as that on which the distinct representations ***killer***, ***victim*** and ***suicide*** each bear; and recognizing Cato as the common focus of these representations is intrinsically involved in inferring from that judgement, in accordance with laws by which Frege's logic exceeds Kant's, that, for instance, Cato killed someone who killed himself. In that way the notion of an object demanded by Frege's polyadic logic does not supplant, but rather incorporates, the core of the metaphysical notion.

8

A plausible sketch can present all three of the above ideas as present in Frege's discussion of (***Directions***), and hence in his acceptance of both (***Numbers***) and (***Classes***). The first, hierarchical idea encouraged him to expect ***some*** further generalization of (***Directions***) than (***1st Level Abstraction***). The second idea, dictating that the values of functions are complete, would lead him to overlook (***Generalization 2***) as a candidate generalization in favour of (***Generalization 1***). And the third idea, that different articulations of a judgement ground logic's claim intrinsically to be dealing with objects, must have had a role in preventing him from looking askance at the result.

There is of course no question of turning this sketch into a compelling argument leading from the three ideas to the conclusion Frege reached. That is so simply because the three ideas are sound while Frege's conclusion was a disastrous mistake. More than that, though, there is probably no single best way of firming up the sketch even into an unsound argument that would enable us to identify precisely Frege's wrong move. The three ideas seemed to

---

[6] See e.g. Dummett 1991, Ch. 15; Ricketts 2004. This similarity between Dummett and Ricketts is discussed in Sullivan 2005.

[7] I think we need to attribute this more ambitious aim to Frege if we are to make sense of the way he explicitly opposes the Kantian way of thinking at the start of Part III of ***Begriffsschrift*** (1879, §§23-24), and, more generally, if we are to understand how logicism could ever have seemed to him a viable metaphysical option.

Frege collectively to point in the direction he took, and they can only even seem to do that so long as none of them is very sharply in focus.

Despite that, I have focused on the difference between (***Generalization 1***) and (***Generalization 2***) to bring out that the trouble must have lain in the way Frege's thinking connected the first and the second of these ideas, a conception of functions as hierarchically ordered and a commitment to their having a single type of value. These two ideas acted in concert to lead Frege to look favourably on (***Process***), when it should have been obvious to him that whatever impetus in that direction the first idea might seem to provide is undermined by the second. In other words, if the first idea invited him to look for a higher-order principle analogous to the sound principle exemplified by (***Directions***), then the second idea should have kicked in to make it plain to him that no higher-order principle he could recognize would be genuinely analogous.

The fact that it didn't sits alongside the observations of §3 as further evidence that Frege's own thinking had not fully adapted itself to what is most original in his conception of how logic relates to objects. Frege's view of the conceptual constituents of judgement as hierarchically ordered was not in itself a novelty. For Kant, too, concepts are 'predicates of possible judgements', and a judgement is the subsumption of a lower by a higher concept (A68-9/B93-4). That objects form the basis of this hierarchy is also an idea shared with Kant: 'Judgement is therefore the mediate knowledge of an object, that is, the representation of a representation of it. In every judgement there is a concept which holds of many representations, and among them of a given representation that is immediately related to an object' (ibid.). Where Frege crucially departs from Kant is in holding that a concept's location in the hierarchy, fixed by its relation to the absolute anchor objects provide, is internal or intrinsic to its nature, and therefore to the logical principles that hold of it.

For Kant the predicational structure of a judgment which brings it within the scope of logical principles is ***the same*** at every level of the hierarchy. It is for that reason that its eventual grounding in the intuition of objects, which must be secured if a judgement is to have any real content, is an extra-logical condition on the judgement, one to whose satisfaction or otherwise logic is constitutionally blind. For Frege, by contrast, distinct logical principles, distinguished precisely by their relation to the grounding objects provide, apply at different levels of the hierarchy. A determinate relation to and necessary bearing on objects is thus, in Frege's conception, intrinsic to judgements, insofar only as they exemplify the structures by which they are subject to his logical laws.

This radically new conception perhaps need not altogether exclude, as Frege's most stringent official thinking insists, the recognition of analogies between principles holding at different levels. But where there are analogies to recognize the correct expression of them must abstract ***completely*** from the anchoring to objects that distinguishes the several principles they embrace. It must, that is, take some such form as (***Generalization 2***), where the type of ***every*** component is identified only relatively to other components, without any having an absolute anchor to objects. It cannot, as (***Generalization 1***) attempts to do, abstract

partially from location in the hierarchy, assigning a merely relative location to some components but an absolute location to others.

In summary of this we could say:

> The principles (***1st Level Abstraction***) and (***2nd Level Abstraction***) are genuinely Fregean principles – one of them false, of course – in which every type is anchored.
>
> (***Generalization 2***), in which no type is anchored, is a principle that properly belongs to a pre-Fregean way of thinking, though one which might, relaxing his official strictures, be innocently incorporated into a Fregean scheme.
>
> (***Generalization 1***), in which some types are anchored while others float free, represents an incoherent mixture of Fregean and pre-Fregean understandings of the nature of logical principles.

Sadly, it was (***Generalization 1***) that Frege set out to defend. Frege fell into the contradiction, then, because even he did not keep a grip on what makes his hierarchy of concepts distinctively Fregean.

(***Classes***) exacted the penalty. (***Numbers***) did not, keeping Fregean logicism running for another nineteen years and its neo-Fregean relative for a century more. But on the perspective advanced here, the consistency of (***Numbers***), as advanced by Frege, was sheer luck.

9

This conclusion can be seen as supplementing a suggestion of Michael Dummett's (1991a), that Frege opened himself to the contradiction by running together two quite different models of re-articulating or re-carving the content of a judgement: on the one hand, the model illustrated by different analyses of a sentence like 'Cato killed Cato' in ***Begriffsschrift*** §9; and, on the other, the model supposed to be illustrated by re-carving an equivalence as an identity in ***Grundlagen*** §64. With hindsight the essential difference between these models stands out starkly. Both are supposed to be ways of recognizing new entities as extractable from the content of a judgement already understood. But while the ***Begriffsschrift*** §9 model always pushes ***up*** the hierarchy, in that the new entities recognized are always of higher level than those presupposed by their recognition, the ***Grundlagen*** §64 model is supposed to loop back down to justify recognition of logical objects. I have tried to add to Dummett's suggestion something that I think is needed to make it genuinely explanatory. Conflation of the two models could have made the second's way of yielding objects seem innocent to Frege only if he already thought of the first as having an important role in sustaining logic's claim to provide genuine knowledge of objects. When we hold in focus both that point of connection and the difference just noted the proper conclusion seems to be this: on the one hand, the ***Begriffsschrift*** §9 model ***does*** justify Frege's contention that the notion of an object is a logical one, and that logic is intrinsically engaged with objects; on the other, the ***Grundlagen*** §64 model falls disastrously short of justifying the recognition of logical objects.[8]

University of Stirling
Stirling FK9 4LA
Scotland
p.m.sullivan@stir.ac.uk

[8] I should like to thank: audiences at the American Philosophical Association (Eastern Division) and the Moral Sciences Club, Cambridge, for helpful suggestions when ancestors of this paper were presented; Michael Potter, for many discussions of the topic; the Arts and Humanities Research Council, for Research Leave during which one round of reworking occurred.

REFERENCES

Boolos, George (1996). 'Whence the contradiction?'. In M. Schirn, ed., ***Frege: Importance and Legacy*** (Berlin: De Gruyter), pp. 234-252. .

Dummett, Michael (1988). 'Reply to "Dummett's dig", by Baker and Hacker', ***Phliosophical Quartertly*** 38, pp. 87-103.

—— (1991). ***Frege: Philosophy of Mathematics*** (London: Duckworth).

—— (1991a). 'More about thoughts'. In his ***Frege and Other Philosophers*** (Oxford: Clarendon Press), pp. 289-314.

—— (1993). 'Existence'. In his ***The Seas of Language*** (Oxford: Clarendon Press), pp. 277-307.

—— (1996). 'Reply to Boolos'. In M. Schirn, ed., ***Frege: Importance and Legacy*** (Berlin: De Gruyter), pp. 253-260.

Frege, Gottlob (1879). ***Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens***. Translated by T. W. Bynum in ***Conceptual Notation and related articles*** (Oxford: Clarendon Press, 1972).

—— (1884). ***Die Gundlagen der Arithmetik***. Reprinted with a translation by J. L. Austin as ***The Foundations of Arithmetic*** (Oxford: Blackwell, 1978).

—— (1893). ***Grundgesetze der Arithmetik***, vol. 1. Part translated in ***The Basic Laws of Arithmetic***, ed. M. Furth (Berkeley and Los Angeles: University of California Press,1964).

—— (1979). ***Posthumous Writings***, trans. P. Long and R. White (Oxford: Blackwell).

Kant, Immanuel (1933). ***Critique of Pure Reason***, trans. N. Kemp Smith (London: Macmillan). (References in standard A/B format.)

Ricketts, Thomas (2004). 'Frege, Carnap, and Quine: continuities and discontinuities'. In S. Awodey and C. Klein, eds., ***Carnap Brought Home: The View from Jena*** (New York: Open Court), pp. 181-202.

Sullivan, Peter M. (2004). 'Frege's logic'. In D. M. Gabbay and J. Woods, eds., ***Handbook of the History of Logic***, vol. 3 (Amsterdam: Elsevier), pp. 659-750.

—— (2005). 'Metaperspectives and internalism in Frege'. In M. Beaney & E. Reck, eds., ***Frege: Critical Assessments*** (London: Routledge).