1 Memoir 12.11.08

One of our goals in this section of the course is to understand the extent to which philosophical reflection on the nature of mathematical knowledge and mathematical truth has been shaped by the development of mathematical logic through the twentieth century. We began with a dramatic reading of the second problem which Hilbert propounded in his celebrated address to the International Congress of Mathematicians in Paris in 1900, reproduced below.

> When we are engaged in investigating the foundations of a science, we must set up a system of axioms which contains an exact and complete description of the relations subsisting between the elementary ideas of that science. The axioms so set up are at the same time the definitions of those elementary ideas; and no statement within the realm of the science whose foundation we are testing is held to be correct unless it can be derived from those axioms by means of a finite number of logical steps. Upon closer consideration the question arises: Whether, in any way, certain statements of single axioms depend upon one another, and whether the axioms may not therefore contain certain parts in common, which must be isolated if one wishes to arrive at a system of axioms that shall be altogether independent of one another.
>
> But above all I wish to designate the following as the most important among the numerous questions which can be asked with regard to the axioms: To prove that they are not contradictory, that is, that a definite number of logical steps based upon them can never lead to contradictory results.
>
> In geometry, the proof of the compatibility of the axioms can be effected by constructing a suitable field of numbers, such that analogous relations between the numbers of this field correspond to the geometrical axioms. Any contradiction in the deductions from the geometrical axioms must thereupon be recognizable in the arithmetic of this field of numbers. In this way the desired proof for the compatibility of the geometrical axioms is made to depend upon the theorem of the compatibility of the arithmetical axioms.
>
> On the other hand a direct method is needed for the proof of the compatibility of the arithmetical axioms. The axioms of arithmetic are essentially nothing else than the known rules of calculation, with the addition of the axiom of continuity. I recently collected them and in so doing replaced the axiom of continuity by two simpler axioms, namely, the well-known axiom of Archimedes, and a new axiom essentially as follows: that numbers form a system of things which is capable of no further extension, as long as all the other axioms hold (axiom of completeness). I am convinced that it must be possible to find a direct proof for the compatibility of the arithmetical axioms,

> by means of a careful study and suitable modification of the known methods of reasoning in the theory of irrational numbers.
>
> To show the significance of the problem from another point of view, I add the following observation: If contradictory attributes be assigned to a concept, I say, that mathematically the concept does not exist. So, for example, a real number whose square is -l does not exist mathematically. But if it can be proved that the attributes assigned to the concept can never lead to a contradiction by the application of a finite number of logical processes, I say that the mathematical existence of the concept (for example, of a number or a function which satisfies certain conditions) is thereby proved. In the case before us, where we are concerned with the axioms of real numbers in arithmetic, the proof of the compatibility of the axioms is at the same time the proof of the mathematical existence of the complete system of real numbers or of the continuum. Indeed, when the proof for the compatibility of the axioms shall be fully accomplished, the doubts which have been expressed occasionally as to the existence of the complete system of real numbers will become totally groundless. The totality of real numbers, i. e., the continuum according to the point of view just indicated, is not the totality of all possible series in decimal fractions, or of all possible laws according to which the elements of a fundamental sequence may proceed. It is rather a system of things whose mutual relations are governed by the axioms set up and for which all propositions, and only those, are true which can be derived from the axioms by a finite number of logical processes. In my opinion, the concept of the continuum is strictly logically tenable in this sense only. It seems to me, indeed, that this corresponds best also to what experience and intuition tell us. The concept of the continuum or even that of the system of all functions exists, then, in exactly the same sense as the system of integral, rational numbers, for example, or as Cantor's higher classes of numbers and cardinal numbers. For I am convinced that the existence of the latter, just as that of the continuum, can be proved in the sense I have described; unlike the system of all cardinal numbers or of all Cantor s alephs, for which, as may be shown, a system of axioms, compatible in my sense, cannot be set up. Either of these systems is, therefore, according to my terminology, mathematically non-existent.

We began to discuss the role of axiomatization in mathematical proof and reviewed some of the proofs we'd seen, with an eye toward assessing the degree to which they conformed to the image of explicit logical deduction from clearly articulated axioms. We presented two proofs that the sum $1+2+\ldots+n$ equals $(n^2+n)/2$, both of which involved extensive use of "visualization". It was not obvious how to encapsulate either proof in a step-by-step deduction from axioms; nonetheless, both proofs seemed quite convincing. We suggested that

a prerequisite to achieving a step-by-step argument would be to provide a more explicit statement of the problem, since ellipses (...) are peculiarly difficult to handle in logical deductions. We provided the recursive definition of a function g whose value $g(n)$ is the sum of the first n numbers, as follows.

$$g(0) = 0;\ g(n+1) = g(n) + (n+1).$$

Armed with this definition, we will consider a third proof, by mathematical induction, that $g(n) = (n^2 + n)/2$ for all n. It was hoped that everyone would give some thought to this at the weekend, and to the following question as well, which was suggested by the second proof (David's proof) which involved considering an $n \times n$ grid cut in half along the diagonal. In particular, recall that Professors Ewald and Kazdan demonstrated that there is a one-one correspondence between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$ (here $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$). Can you explicitly define a function $f(m, n)$ which establishes such a correspondence?

Further food for thought: On the first exam, Professor Kazdan suggested that you make use of the following *lemma* to prove that the the square root of three is irrational: if a prime number p divides the product of a pair numbers a and b, then p divides a or p divides b. How would you go about proving the lemma? Why would you go about proving the lemma? Can you conjecture an answer to the question, "Which positive integers have rational square roots?" Can you use the lemma, or a generalization or corollary thereof, to establish your conjecture?

2 Memoir 12.11.13

We recalled the recursive definition of a function g whose value $g(n)$ is the sum of the first n numbers, as follows. g(0) = 0; g(n+1) = g(n) + (n+1). We proceeded to prove, by mathematical induction, that

$$\text{for all } n, \;\; g(n) = (n^2 + n)/2. \tag{1}$$

We began by stating the Principle of Mathematical Induction:

> if $P(0)$ and for every n, if $P(n)$, then $P(n+1)$, then for every n, $P(n)$.

This may be rendered in logical notation as follows:

$$(P(0) \wedge (\forall n)(P(n) \rightarrow P(n+1))) \rightarrow (\forall n)P(n).$$

Here we use symbols $\forall$ for the universal quantifier, read "for all" or "for every", $\wedge$ for the truth-functional connective conjunction, read "and", and $\rightarrow$ for the truth-functional connective material conditional, read "if , then ". Some additional symbols we will use below are $\exists$ for the existential quantifier, read "there is" or "there exists", $\vee$ for the truth-functional connective (inclusive) disjunction, read "or", and $\neg$ for truth-functional connective negation, read "not". As members of the class were quick to observe, the interpretation of the quantifiers involves specifying a domain of discourse over which the variables of quantification range. In the above example, we understand this domain of discourse to be the set of whole numbers (also known to Brooklynites of my generation as the natural numbers) $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

Now, to the proof of (1). For the purposes of applying the principle of mathematical induction, we let $P(n)$ be the property $g(n) = (n^2+n)/2$. By the definition of g we have $g(0) = 0$ thus verifying $P(0)$, the *basis* of our induction. Thus, in order to conclude our argument, it suffices to show that

$$(\forall n)(P(n) \rightarrow P(n+1)).$$

To this end, suppose that $P(n)$, that is, $g(n) = (n^2 + n)/2$ (our *induction hypothesis*). We must show that $P(n+1)$, that is, $g(n+1) = ((n+1)^2+(n+1))/2$. But, by the definition of g,

$$g(n+1) = g(n) + (n+1);$$

which, by the induction hypothesis,

$$= (n^2 + n)/2 + (n+1);$$

which, by elementary arithmetic (we suppressed explicit appeal to axioms of arithmetic to justify steps in the calculation)

$$= (n^2 + 2n + 1 + (n+1))/2 = ((n+1)^2 + (n+1))/2,$$

which concludes the proof.

We turned to justify the Principle of Mathematical Induction. We described how the principle would allow us to construct a proof for each instance $P(n)$ by iterated application of *modus ponens*, the rule of inference which licenses the conclusion B from premises A and if A, then B. Mike was quick to complain that this justification is, in some sense, circular, but in the end we all seemed to agree that it added some measure of confidence in conclusions drawn by the Principle of Mathematical Induction.

We then undertook to axiomatize a tiny fragment of the arithmetic of whole numbers, that involving only zero and the successor function (that is, the function s which maps each whole number n to $n + 1$). We looked for inspiration from Hilbert's characterization of the axiomatic method:

> When we are engaged in investigating the foundations of a science, we must set up a system of axioms which contains an exact and complete description of the relations subsisting between the elementary ideas of that science. The axioms so set up are at the same time the definitions of those elementary ideas; and no statement within the realm of the science whose foundation we are testing is held to be correct unless it can be derived from those axioms by means of a finite number of logical steps.

In light of this, we searched for properties of the structure $\mathsf{N} = \langle \mathbb{N}, s, 0 \rangle$ which would characterize this structure uniquely, that is, any other structure which satisfied these properties would be *isomorphic* to N. In order to state such properties, we set up a language to describe N. the language included a symbol S to refer to the successor function and a symbol $\mathbf{0}$ to denote 0. After considerable effort, we came up with the following infinite list of properties.

1. $(\forall n)\neg(S(n) = \mathbf{0})$. (Zero has no predecessor.)
2. $(\forall n)(\neg(n = \mathbf{0}) \rightarrow (\exists m)(S(m) = n))$. (Every number besides zero has a predecessor.)
3. $(\forall m)(\forall n)(S(m) = S(n) \rightarrow m = n)$. (The successor function is one-to-one.)

$3 + i$. $(\forall n)\neg(S^i(n) = n)$. (The successor function has no cycle of length i. S^i stands for i occurrences of the symbol S. This is an infinite list of conditions, one for each whole number $i > 0$.)

We observed that any structure A which satisfies all these properties contains a part that is isomorphic to N, but we left open the question whether A itself is isomorphic to N. We will return to this question in our next class. At the end of class, Neil pressed the interesting question whether the fact that we have introduced infinitely many axioms might conflict with Hilbert's dictum that "no statement within the realm of the science whose foundation we are testing is held to be correct unless it can be derived from those axioms by means of a finite number of logical steps" – another question to which we will return.

3 Memoir 12.11.15

We began by taking up the question whether our attempted axiomatization of the theory of the whole numbers with zero and successor, that is, the theory of N, succeeded in characterizing N uniquely, that is, up to isomorphism. First, Ashleigh noted that there are structures other than N which satisfy our axioms, for example the structure $\mathsf{M} = \langle \mathsf{N}, s', 0\rangle$ which is identical to N except that $s'(3) = 5, s'(5) = 4$, and $s'(4) = 6$. Though distinct from N, this structure is, nonetheless, isomorphic to N (the map which transposes 4 and 5 and is the identity elsewhere is an isomorphism from N onto M). After much discussion, we realized that the are many (pairwise) non-isomorphic structures which satisfy our axioms. They may be conveniently described as follows. Let $\mathbb{Z}$ be the set of integers, that is, $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$ and let Y be an arbitrary set. We define a structure $A_Y = \langle \mathbb{N} \cup (Y \times \mathbb{Z}), s^*, 0\rangle$ where the function s^* is defined as follows.

1. for every $n \in \mathbb{N}$, $s^*(n) = n + 1$.
2. for every $\langle a, p\rangle \in Y \times \mathbb{Z}$, $s^*(\langle a, p\rangle) = \langle a, p + 1\rangle$.

It is easy to verify that for every set Y, A_Y satisfies the all the axioms (1)-(3+i) we recorded in Memoir 2. Note also that for any two sets Y and Y', A_Y is isomorphic to $A_{Y'}$ if and only if Y is equipollent to Y'. It follows at once that there are countably many pairwise non-isomorphic countable structures which satisfy the axioms, and for each uncountable infinite cardinality, there is exactly one structure of that cardinality up to isomorphism which satisfies the axioms. We began to inquire into the question how this great variety of interpretations might bear on the success of our axiomatization from Hilbert's point of view. In particular, is it possible that with all these different interpretations (and with the additional feature that we have recorded infinitely many axioms) that nonetheless every correct statement (that can formulated in the language we described) may be derived via finitely many logical steps from our axioms. In order to pursue this question, we introduced the notion of logical consequence. Let X be a collection of statements and let D be a statement. D is a *logical consequence of* X if and only if for every interpretation A, if A satisfies every statement in X, then A satisfies D. Now, in the case of our axioms, since they have so many different satisfying interpretations, it may seem likely that there is a statement D which is true in one and false in another. In this case, neither D nor $\neg D$ would be a logical consequence of our axioms, and hence our axioms would fail of allowing the derivation of every correct statement about $\mathbb{N}$ (on the assumption that if a statement is derivable from our axioms, then it is a logical consequence of them). We will inquire further into the matter next time.

4 Memoir 12.11.20

We continued our discussion of our attempted axiomatization of the theory of $\mathbb{N}$ which we compiled during our second class meeting (see Memoir 2). We named our collection of axioms $T_{\mathbb{N}}$. We recalled the notion of logical consequence:

> a statement D is a logical consequence of a set of statements T if and only if every interpretation which satisfies all the statements in T also satisfies D.

We introduced the notions of *categoricity* and *completeness*: a set of statements is categorical if and only if there is exactly one interpretation, up to isomorphism, which satisfies it; a set of statements is complete if and only for every statement D (in the relevant language) either D or its negation is a logical consequence of T. We noted that if a set of statements T is categorical, then T is complete. We noted that the converse is not true, in general. In particular, no set of sentences of first-order logic which is satisfied by some infinite interpretation is categorical (this is a corollary to the Compactness Theorem for first-order logic discussed below). On the other hand, many such sets are complete (given any interpretation I, consider the set of all first-order sentences satisfied by I). In particular, we asserted (but did not prove) that $T_{\mathbb{N}}$ is complete. It follows at once that all the structures A_Y (see Memoir 3) are indistinguishable by first-order sentences, that is, they are *elementarily equivalent*. (Structures A and B are elementarily equivalent if and only if for every first order sentence D, A satisfies D if and only if B satisfies D.) Despite the fact that $T_{\mathbb{N}}$ is *not* categorical, it nonetheless has as consequences all correct statements about $\mathbb{N}$ which can be formulated in first-order logic!

We went on to consider possibilities for providing categorical descriptions of $\mathbb{N}$ in more expressive logical languages. Pursuing a suggestion of David's, we discussed an axiomatization which adds the infinitary sentence:

$$(\forall(n)(n = \mathbf{0} \vee n = S(\mathbf{0}) \vee n = S(S(\mathbf{0})) \vee \ldots).$$

We agreed that this sentence, along with axioms 1 and 3 of $T_{\mathbb{N}}$ provides a categorical axiomatization of $\mathbb{N}$. We were leery of the infinitary character of the new axiom, but waited on further reflection to mount any definite critique. We recalled the Principle of Mathematical Induction:

$$(\forall P)[(P(0) \wedge (\forall n)(P(n) \rightarrow P(n+1))) \rightarrow (\forall n)P(n)].$$

We discussed the fact that this principle involves a second-order quantifier $(\forall P)$ which ranges over all subsets of the universe of discourse of an interpretation, as opposed to first-order statements which involve quantification only over members of the universe of discourse. We showed that this principle, together with axioms 1 and 2 of $T_{\mathbb{N}}$, is also a categorical axiomatization of $\mathbb{N}$. We dub this axiomatization PA_2 – second-order Peano arithmetic.

We then considered why we'd labored so long over the question of a complete axiomatization of the theory of $\mathbb{N}$, since PA_2 provides such and had essentially

been suggested by Eric during our second class meeting. In order to explain this, we recurred to the question whether this axiomatization might still suffer inadequacy *vis-à-vis* Hilbert's demand that our axiomatization allow finite derivation of all consequences. If this is to be the case, the consequence relation (for the relevant logic) must satisfy the following *Compactness Principle*: for any statement D and set of statements T, if D is a logical consequence of T, then D is a logical consequence of some finite subset of T. We showed that the consequence relation for any logical language which provides a categorical axiomatization of $\mathbb{N}$ is not compact. The argument is as follows. Suppose T is such a categorical axiomatization. Introduce a new predicate symbol Q and add to T the following infinite list of statements: $(\exists n)Q(n), \neg Q(\mathbf{0}), \neg Q(S(\mathbf{0})), \neg Q(S(S(\mathbf{0}))), \ldots$. We showed that this set of sentences is unsatisfiable, but that every finite subset of this set is satisfiable. It follows at once that $\neg(\mathbf{0} = \mathbf{0})$ is a consequence of the entire set, but of no finite subset. Thus, our second-order axiomatization, and the infinitary axiomatization, fail to respect Hilbert's call for axioms which allow for finite derivability of all correct statements. We noted, but did not prove, that the consequence relation for first-order logic is compact.

In order to form a more concrete idea of the notion of derivation, we discussed the following example, a discussion we may continue next time.

1. All Philadelphians are either lawyers or doctors.
2. No doctors drink beer.
3. All lawyers drink wine.
4. Some Philadelphians drink beer.
   THEREFORE
5. Some Philadelphians drink wine.

$(\exists x)(Px \wedge Wx)$ is deducible from $(\forall x)(Px \rightarrow (Lx \vee Dx))$, $(\forall x)(Dx \rightarrow \neg Bx)$, $(\forall x)(Lx \rightarrow Wx)$, and $(\exists x)(Px \wedge Bx)$

$\{1\}$	(1) $(\exists x)(Px \wedge Bx)$	P
$\{1,2\}$	(2) $Px \wedge Bx$	(1)x EII
$\{3\}$	(3) $(\forall x)(Px \rightarrow (Lx \vee Dx))$	P
$\{3\}$	(4) $Px \rightarrow (Lx \vee Dx)$	(3) UI
$\{5\}$	(5) $(\forall x)(Dx \rightarrow \neg Bx)$	P
$\{5\}$	(6) $Dx \rightarrow \neg Bx$	(5) UI
$\{7\}$	(7) $(\forall x)(Lx \rightarrow Wx)$	P
$\{7\}$	(8) $Lx \rightarrow Wx$	(7) UI
$\{1,2,3,5,7\}$	(9) $Px \wedge Wx$	$(2,4,6,8)$ TF
$\{1,3,5,7\}$	(10) $(\exists x)(Px \wedge Wx)$	EG; $\{2\}$EIE

5 Memoir 12.11.27

We began by looking at the derivation we discussed at the end of our last class meeting. We focussed on the feature that allowed us to pass in one step to a statement which is a truth-functional consequence of earlier statements in the derivation. We defined the notion of truth-functional consequence: a statement S is a truth functional consequence of a set of statements T if and only if S is satisfied by every truth assignment to its truth functionally prime constituents which satisfies every statement in T. We observed that there are 2^n truth assignments to n prime constituents and, as a result, a derivation involving a one step inference of this kind might well lack transparency, to both human and machine, if the derivation contained statements with many prime constituents. We noted that a derivation supplemented with the truth table, though in general extremely long, could be easily recognized, at least by a machine. We noted that it is an open problem whether easily recognizable proofs of tautological consequence could, in general, be provided which are significantly shorter than truth-table verifications. In the course of our discussion, the comparison of the rates of growth of powers of n with exponentials in n arose. In particular, the question, how do you know that for every k, there is an m such that for every n, if $m < n$, then $n^k < 2^n$, seemed worthy of attention. We further discussed the notion of derivation and arrived at the following formulation. A formal system F defines a mechanically decidable relation $\mathsf{D}er(D, S)$ between sequences of (formal) statements D and (formal) statements S, read D is a derivation of S in F. S is derivable in F if and only if there is a D such that $\mathsf{D}er(D, S)$.

We recalled the notion of logical validity, a special case of logical consequence where the set of premisses is empty: a statement S is valid if and only if it is satisfied by every interpretation. We noted that a statement S is a logical consequence of a finite set of statements $T = \{S_1, \ldots, S_n\}$ if and only if the conditional statement $(S_1 \wedge \ldots S_n) \rightarrow S$ is valid. We stated Gödel's Completeness Theorem: there is a formal system F such that for every first-order statement S,

- if S is derivable in F, then S is valid (soundness of F-derivability), and
- if S is valid, then S is derivable in F (completeness of F-derivability).

The completeness theorem establishes that the collection of valid first-order statements is semi-decidable, as follows. We say a set of statements X is semi-decidable if and only if there is a mechanical procedure M such that for every statement S, M outputs "yes" with input S if and only if S is a member of X. The semi-decision procedure for validity may be described as follows: the procedure M with input S enumerates all sequences of statements $D_1, D_2, \ldots$(potential derivations) and successively tests whether $\mathsf{D}er(D_1, S)$ holds, whether $\mathsf{D}er(D_2, S)$ holds, and so on. M outputs "yes", if it reaches a stage n where the test $\mathsf{D}er(D_n, S)$ is positive. The Completeness Theorem guarantees that M thus defined is a semi-decision procedure for validity. Note that M runs on forever and gives no output in case S is not derivable. Next time, we

will discuss whether this result can be improved to establish the decidability of first-order validity. Suggested reading: Alan Turing, On computable numbers, with an application to the *Entsheidungsproblem*, sections 1 and 9.

6 Memoir 12.11.29

We reminded ourselves that the set of valid statements of first-order logic is semi-decidable. We went on to consider the question: is the set of valid statements of first-order logic decidable? In order to engage the possibility of a negative answer, we need a precise notion of mechanical computability. Turing, in his 1936 paper, On computable numbers, with an application to the Entscheidungsproblem, presented such a notion, now known as Turing computability. In section 9 of this paper, Turing provided an analysis of the notion of computation which provides strong evidence that every algorithm can be rendered via a program for one of his machines. This is the strongest argument for the Church-Turing thesis: every mechanically computable function is Turing computable – a thesis that is generally accepted today.

We gave an example of a Turing machine M which provides a decision procedure for the set of binary strings which contain an even number of occurrences of 1. The instructions for the machine consist of the following quintuples:

1. $\langle p, 0, B, R, p\rangle$
2. $\langle p, 1, B, R, q\rangle$
3. $\langle p, B, Y, S, h\rangle$
4. $\langle q, 0, B, R, q\rangle$
5. $\langle q, 1, B, R, p\rangle$
6. $\langle q, B, N, S, h\rangle$

The machine has a one way infinite tape, a reading head which scans one square of the tape at a time, a facility for printing on the currently scanned square of the tape, and a means for shifting the tape one square left or right. The machine has finitely many internal states and the program determines the action of the machine in a given internal state when scanning a given symbol. For example, instruction 1 above specifies that M, when in internal state p and scanning a 0 prints a blank, moves the tape one square to the right and remains in state p. We showed, by example, that when M is started in state p on the leftmost square of a tape containing a binary string occupying some finite initial segment of the tape, it will terminate with a tape whose only nonblank square is inscribed with Y if and only if the binary string contains an even number of occurrences of 1 (resp. N if and only if the binary string contains an odd number of occurrences of 1).

We discussed various possible enrichments to the notion of a Turing machine, for example, additional tapes, complex states, and compound symbols. We indicated that each of these enrichments can be accommodated within the spare resources of a Turing machine. Eric wondered if one wouldn't need some additional resource to program a decision procedure for the set of binary strings which contain the same number of 0's and 1's. We agreed that one of the exam

problems would be to program a Turing machine which provides such a decision procedure.

7 Memoir 12.12.04

Let V be the set of valid sentences of first-order logic. We showed that:

Theorem 1 (Church-Turing) *V is undecidable.*

The proof of Theorem 1 uses some of the central ideas of the theory of Turing computable functions. We make implicit use of the Church-Turing Thesis throughout the proof. First, observe that a set X is semi-decidable if and only if there is a Turing machine M such that for every a, if $a \in X$, then the computation executed by M with input a terminates (halts) and if $a \notin X$ then the computation executed by M with input a does not terminate. Next, note that there are countably many Turing machines. In particular,

Theorem 2 (Turing) *There is an indexing M_e, $e \in N$, of the Turing machines by whole numbers such that the binary partial function U defined as follows is Turing computable. For all e and x, $U(e, x)$ is the result of executing a computation via M_e with input x.*

This is Turing's Universal Function Theorem which is the conceptual progenitor of the programmable digital computer (note how the existence of U effaces the distinction between program and data). Now, for every whole number e, let W_e be the set of x such that $U(e, x)$ is defined, that is, W_e is the semi-decidable set determined by M_e. We are now in a position to show that,

Theorem 3 (Church-Turing) *There is a set which is semi-decidable, but undecidable.*

Proof: Note that $\{W_e \mid e \in N\}$ is the set of all semi-decidable sets. Let $K = \{e \mid e \in W_e\}$. It follows from Theorem 2 that K is semi-decidable. Let $\overline{K} = N - K = \{e \mid e \notin W_e\}$. As in Cantor's diagonal argument, it is easy to see that for every e, $\overline{K} \neq W_e$. Hence, $\overline{K}$ is not semi-decidable. But a set is decidable if and only if both it and its complement are semi-decidable. It follows at once that K is a semi-decidable set which is undecidable. ■

In order to make use of Theorem 3, to prove Theorem 1, we introduced the notion of reducibility. We say that a set X is reducible to a set Y if and only if there is a computable function f such that for all a, $a \in X$ if and only if $f(a) \in Y$. It follows that if X is reducible to Y and Y is decidable, then X is decidable. Theorem 1 is thus a corollary to Theorem 3 and the following result.

Theorem 4 (Church-Turing) *For every set X, if X is semi-decidable, then X is reducible to V.*

We sketched a proof of this theorem as follows. Let X be semi-decidable. Then there is a Turing machine M such that for every a, M halts with input a if and only if $a \in X$. We showed how to construct, for each a, a sentence $f(a)$ of first-order logic which describes the computation executed by M with input a and which is valid if and only if that computation terminates.

8 Addendum to Memoir 12.12.04

At the very end of our last class we proved the Gödel incompleteness theorem, alas, in a such a great rush, that the details may have passed by in a blur. Here is a brief exposition. In accord with our discussion, we define a formal system to be a semi-decidable set of sentences closed under logical consequence. We say a formal system F *numeral-wise represents* a relation $R \subseteq \mathbb{N}^k$ if and only if there is a formula $A(y_1, \ldots, y_k)$ of F such that for every k-tuple of numbers $\langle n_1, \ldots, n_k \rangle$, if $\langle n_1, \ldots, n_k \rangle \in R$, then $A(\mathbf{n_1}, \ldots, \mathbf{n_k}) \in F$, and if $\langle n_1, \ldots, n_k \rangle \notin R$, then $\neg A(\mathbf{n_1}, \ldots, \mathbf{n_k}) \in F$. We say a formal system F is computationally adequate if and only if F is consistent and every decidable relation on $\mathbb{N}$ is numeral-wise representable in F. If the formula $A(y)$ numeral-wise represents a decidable set X in a formal system F we call $(\forall y)A(y)$ a Π^0_1-*sentence*. We call a computationally adequate formal system *1-consistent* if and only if it does not contain the negation of any true Π^0_1-sentence.

Theorem 5 (Gödel) *If F is a computationally adequate formal system, then there is a true Π^0_1-sentence which is not a member of F. If, moreover, F is 1-consistent, then F is incomplete, in particular, there is a Π^0_1-sentence such that neither it nor its negation is a member of F.*

Proof: Let Z be the complement of a semi-decidable, but undecidable, set of numbers (we proved that such sets exist, see Memoir 7). Let M be a Turing machine which computes a semi-decision procedure for the complement of Z. The relation "M does not halt in at most n steps on input k" is a decidable relation and thus we can effectively construct, for each k, a formula $A_k(y)$ which numeral-wise represents, in F, the set of n such that "M does not halt in at most n steps on input k". Observe that for all k, $k \in Z$ if and only if the Π^0_1-sentence $(\forall y)A_k(y)$ is true. It follows at once from the consistency of F and the fact that $A_k(y)$ numeral-wise represents, in F, the set of n such that "M halts in at most n steps on input k", that no false sentence of the form $(\forall y)A_k(y)$ is a member of F. Hence, if every true Π^0_1-sentence were a member of F, then the function which maps k to the sentence $(\forall y)A_k(y)$ would be an effective reduction of Z to F. But this is impossible, since F is semi-decidable and Z is not. Hence, there is a true Π^0_1-sentence which is not a member of F, thereby establishing the first claim of the theorem. The "if, moreover," claim now follows directly from the definition of 1-consistency. ■

It is a corollary of the Gödel Completeness Theorem for first-order logic, that the set of first-order logical consequences of a semi-decidable set of first-order sentences, is itself semi-decidable. A first-order theory is called axiomatizable if it consists of the logical consequences of a decidable set of sentences. Thus Theorem 5 applies to any axiomatizable, computationally adequate first-order theory. There are first-order theories of arithmetic with finitely many axioms which are computationally adequate; the logical consequences of the first three axioms for successor given in Memoir 2, together with "recursive definitions" of addition and multiplication, is such a theory.