JOURNAL OF COMPUTER AND SYSTEM SCIENCES **36**, 384–432 (1988)

The Topology of Provability in Complexity Theory*

KENNETH W. REGAN

Mathematical Sciences Institute, Cornell University, Ithaca, New York 14853

Received November 19, 1986; revised October 30, 1987

We present a general technique for showing that many properties of recursive languages are not provable. Here "provable" is taken with respect to a given sound, recursively axiomatized formal system $\mathfrak{F}$, such as Peano arithmetic. A representative application (Theorems 6.1–6.2) concerns the property of intractability, i.e., non-membership in the class **P**. It says that there exists a language E such that E is not in **P**, but the formal assertion 'E is not in **P**' is independent of $\mathfrak{F}$. Moreover, given any recursive language $A \notin \mathbf{P}$, we can construct E such that also $E \leqslant_m^p A$. Our techniques strengthen similar results in the literature and lead to several other applications pertaining to **P**-immune sets, oracle separations, and the Berman–Hartmanis conjecture. We explain the phenomenon of unprovability in terms of both recursive properties of the formal systems $\mathfrak{F}$ under consideration, and topological properties of complexity classes in a natural space which we call $\mathfrak{R}$. Provable properties correspond to closed sets of $\mathfrak{R}$. The topology provides geometric intuition for recognizing classes which are not closed in $\mathfrak{R}$, such as $\mathbf{NP}\backslash\mathbf{P}$ (unless it is empty). We show how independence results follow immediately for these classes. In conclusion we argue that the type of independence result presented here forms an obstacle for day-to-day work in complexity theory, but does not bear directly on the possible independence of the $\mathbf{P} \stackrel{?}{=} \mathbf{NP}$ question from Peano arithmetic or set theory. However, we believe our tools capable of measuring the link between the structure of a given language $\mathrm{E} \notin \mathbf{P}$ and the formal strength needed to prove the assertion '$\mathrm{E} \notin \mathbf{P}$.' Research in this direction has already been initiated by D. Joseph (*J. Comput. System Sci.* **25** (1983), 205–228).

1. INTRODUCTION

Which assertions in complexity theory are provable in the strong logical theories that underpin most mathematical work? Much interest has centered on whether the $\mathbf{P} \stackrel{?}{=} \mathbf{NP}$ question might be independent of Peano arithmetic (PA) or even stronger systems such as set theory, which for us means the Zermelo–Fraenkel system plus the axiom of choice (ZFC).

It is widely considered, however, that efforts so far at giving independence results for strong theories have fallen short of the sophistication needed to capture the issue of separating complexity classes. The first attempts [HaHo76, Har78] turned on properties of individual Turing machines, e.g., by constructing a TM M such

* This research was partly supported by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University.

0022-0000/88 $3.00

that $L(M) = \emptyset$ and yet neither '$\mathbf{NP}^{L(M)} = \mathbf{P}^{L(M)}$' nor '$\mathbf{NP}^{L(M)} \neq \mathbf{P}^{L(M)}$' is provable in ZFC. Although $\mathbf{NP}^{L(M)} = \mathbf{NP}$ and $\mathbf{P}^{L(M)} = \mathbf{P}$, this result does not really give evidence that '$\mathbf{NP} \stackrel{?}{=} \mathbf{P}$' is undecidable in ZFC, for two reasons:

(1) The TM M constructed is a "bad name" for the empty language. It has a subroutine which diagonalizes over proofs in ZFC. In fact, '$L(M) = \emptyset$' provably (in ZFC) implies the consistency of ZFC, so that the conclusion has much the same content as *Gödel's Second Incompleteness Theorem.*

(2) There is nothing special about ZFC in the result, which indeed was originally stated with reference to a general sound, recursively axiomatised (*r.a.* for short) formal system '$\mathfrak{F}$.' The construction works equally well for systems $\mathfrak{F}$ having '$\mathbf{NP} = \mathbf{P}$' or '$\mathbf{NP} \neq \mathbf{P}$' (whichever is true) as an axiom, even though such a system trivially decides the $\mathbf{NP} \stackrel{?}{=} \mathbf{P}$ question.

Among references raising these and related points we mention [Háj79, Phi79, Gra80, JoYg81, Lei82].

How may one meet these objections? Several authors following on from [Lip78, DL80] have delved into the workings of specific formal systems, but whether these are strong enough to be "acceptable for computation" has been much debated in the above references. The system studied in [Jos83] may have the most natural interest.

Recent attempts at a second brand of independence results for strong theories stem from work in [Sö82, Reg83b, Kow84, Har84]. While still not distinguishing among sound, r.a. systems $\mathfrak{F}$, they meet objection (1) by rendering certain assertions about languages E unprovable in $\mathfrak{F}$, almost regardless of how E is represented by a TM algorithm. For example, [Har84] shows how to construct a language $E \in \mathbf{DSPACE}[2^{2^n}]$ such that $\mathbf{NP}^E \neq \mathbf{P}^E$, and yet for no TM M accepting E (and obeying a double-exponential space bound) can $\mathfrak{F}$ prove '$\mathbf{NP}^{L(M)} \neq \mathbf{P}^{L(M)}$.' Comparing this with the first example, we have lost $L(M) = \emptyset$, but have gained that the unprovability of the assertion '$\mathbf{NP}^E \neq \mathbf{P}^E$' pertains to the language E itself.

The aim of this paper is to provide the sharpest possible rendition of these theorems, and of the technique called *delayed diagonalization* which underlies them. For example, we reduce the complexity of the language E above and remove the restriction to TMs obeying a double-exponential space bound. The idea is to make this brand of independence result easy to derive in a uniform manner.

We do this in the context of a topology $\mathfrak{R}$ on the class of recursive languages whose basic closed sets are intuitively recognizable, and which correspond to properties of languages that are provable in a very weak sense. Let $\mathscr{C}$ be a class which is not closed in $\mathfrak{R}$, for example $\mathscr{C} := \{L \in \mathbf{EXPTIME} \mid \mathbf{NP}^L \neq \mathbf{P}^L\}$. We then know immediately that for any sound, r.a. formal system $\mathfrak{F}$, there must be languages E in $\mathscr{C}$ whose membership in $\mathscr{C}$ is not provable in $\mathfrak{F}$.

One thing our results show is that as viewed by any sound, r.a. $\mathfrak{F}$, the "outer boundary" of a complexity class such as $\mathbf{P}$ is very "fuzzy." For example, with respect to $\mathfrak{F}$ (and a fixed definition of $\mathbf{P}$ over $\mathfrak{F}$), the class $\{L \in \mathbf{REC} \setminus \mathbf{P} \mid \mathfrak{F} \nvdash \text{'}L \notin \mathbf{P}\text{'}\}$ of *unprovably intractable* languages is nonempty. Thus it is consistent with $\mathfrak{F}$ to believe that **P** is larger than it is. In contrast, the main results of [Bak79] show that the class of languages which $\mathfrak{F}$ can prove to lie *inside* **P** is equal to **P** itself, so that the "inner boundary" of **P** is sharp.

There is still objection (2) to contend with, however. If anything our work makes it even more prominent, e.g., by answering negatively an open question of [Har84] about whether (**NP**-complete) languages which lie outside the p-isomorphism class of SAT, but not provably so, can be constructed with respect to some sound, r.a. formal systems $\mathfrak{F}$ but not to others. We shall not attempt to meet it in this paper, but do provide techniques in support of a speculative possibility growing out of [Jos83, PaHa77].

Briefly put, the idea is to measure the extent of languages unprovably having a property Π (such as intractability) in terms of the maximum growth rate of functions that $\mathfrak{F}$ can prove recursive, or figuratively the "speed limit" of $\mathfrak{F}$. [Jos83] gives evidence that SAT should be provably intractable in a (non-r.a.) system called **ET[Elem]** whose speed limit is observed only by elementary functions, because the distribution of "hard" vs "easy" instances to the satisfiability problem appears to be fairly uniform. The results of [Jos83] do not directly carry over to systems such as $\mathfrak{F} = \text{PA}$ because they use a backward-search property similar to the gist of the proof (in [Cu80]) that the elementary functions are precisely those computable in time bounded by an elementary function. Nevertheless, stronger techniques for analyzing the structures of **NP**-complete languages may yet resolve whether SAT can be proved intractable in (r.a.) systems $\mathfrak{F}$ having higher speed limits such as PA and ZFC.

We also raise several technical issues in representing and proving properties of languages which may be unsuspected by practitioners in complexity theory. The purpose going beyond this paper is to ask, "How much logical power will be needed to go after the really hard problems in complexity?"

1.1. *Overview*

Section 2 presents background material in complexity theory, logic, and topology. Section 3 introduces the topology $\mathfrak{R}$, preceding it by a larger space $\mathfrak{R}_0$ of which we regard $\mathfrak{R}$ as the effective part. It contains both constructions of general interest in computability theory, and examples of how different complexity classes are classified topologically.

Section 4 draws connections among subspaces which are closed in $\mathfrak{R}$, recursively presentable classes, and provable properties of languages. The characterizations we obtain lead us to promote $\mathfrak{R}$ as a natural object of study. Section 5 lays out the main technical tools, chiefly a refinement of delayed diagonalization, for showing certain classes not to be closed in $\mathfrak{R}$.

Section 6 combines the main theorem of Section 5 and the relationships established in Section 4 into several applications, which work with respect to any given sound, recursively axiomatized formal system $\mathfrak{F}$. Chief among them are: Section 6.1 that **NP**$\setminus$**P** contains unprovably intractable languages unless **NP** = **P**, Section 6.5 that there are recursive sets separating **NP** from **P** by oracle but not provably so, and Section 6.6 that if the Berman–Hartmanis conjecture fails, then there are **NP**-complete languages E which are not p-isomorphic to SAT, but such that the assertion 'E is not p-isomorphic to SAT' is independent of $\mathfrak{F}$. The last answers an open question of [Har84] for $\mathfrak{F} := \mathrm{PA}$, and we explain how the others strengthen previous work.

Contributing to this, we show in Section 6.2 that some "provable" properties of languages can be proved only under representations by very badly behaved TM algorithms. We show in Section 6.3 that **P**-*immunity* is not a provable property, and in Section 6.4 that languages which are not even provably infinite appear at many levels of complexity. We provide diagrams for these results, meanwhile coping with the fact that the underlying topology $\mathfrak{R}$ differs greatly from that of the Euclidean plane.

Section 7 explains how open problems such as $\mathbf{NP} \stackrel{?}{=} \mathbf{P}$ or $\mathbf{NP} \stackrel{?}{=} \mathbf{co\text{-}NP}$ reflect questions about the connectedness of certain classes as subspaces of $\mathfrak{R}$. In conclusion we assess the usefulness of $\mathfrak{R}$ in helping to resolve these problems.

Author's Note and Acknowledgments. This paper is an update of the similarly titled report [Reg86a] appearing in the "Proceedings of the 1st Structure in Complexity Theory Conference," Lecture Notes in Computer Science, Vol. 223, Springer-Verlag. Some remarks in the original have been turned into results, while some other material has been removed. Most of the work, including all lemmas stated without proof, comes from my recently completed doctoral dissertation [Reg86b].

Much of this research was conducted while I was at Merton College, Oxford University, and I thank the Fellows of Merton for their support. Support from the Cornell University Mathematical Sciences Institute, which is sponsored by the U.S. Army Research Office, on this final version has even extended to free off-hours access to their WP equipment. I am grateful to Professors Angus Macintyre and Uwe Schöning for helpful comments on earlier drafts and on my dissertation, and to Professor Dana Scott for bringing [Vis80] to my attention. I am also grateful for suberbly helpful and painstaking comments by an anonymous referee, whom I have also credited for answering two of my previously open questions in Section 3.3.

2. Concepts and Notation

Concepts and theorems known in the literature but not fully described here are set in *italics* at their first mention. General references are [BJ74] and [Rog67] for logic and recursion theory, [HU79] for automata and complexity theory, and [Dug66] for topology. We try to strike a compromise between stating logical formulas technically and the more readable but less rigorous expedient of referring to them by what they "mean." We use single quotes when we can point to an intended formal representation, and double, when talking about, e.g., whether $\mathfrak{F} \vdash$ "$\mathbf{NP} \neq \mathbf{P}$" in a general sense.

2.1. *Strings, Numbers, Functions, Sets, and Classes*

Taking $\Sigma := \{0, 1\}$, we identify Σ^* with $\mathbb{N}^+$ by associating to each string x the number $\mathbf{bin}(x)$ having binary representation $1x$. Under this correspondence the *empty string* λ maps to 1. From this we carry over properties of arithmetic and order from numbers to strings, e.g., by writing '$k \leqslant y$' when k is a number and y is a string. We regard the *characteristic function* χ_L of a language L as a map from Σ^* to $\{1, 2\}$, whereby for all x, $\chi_L(x) = 1 \Leftrightarrow x \in L$.

A *pairing function* is a bijection from $\Sigma^* \times \Sigma^*$ to Σ^*. We let $\langle \cdot, \cdot \rangle$ stand for any of the familiar recursive pairing functions found in, e.g., [Rog67] or [MY78], and π_1, π_2 for the associated *projection functions* satisfying $\pi_1(\langle x, y \rangle) = x$, $\pi_2(\langle x, y \rangle) = y$ for all $x, y \in \Sigma^*$.

The *join* of two languages A and B is written $A \oplus B$ and defined to be $\{x0 \mid x \in A\} \cup \{y1 \mid y \in B\}$. Note that we place the "decision bit" at the end rather than the beginning, as is also commonly practised. We compose $\oplus$ from left to right, so that, e.g., $A_1 \oplus A_2 \oplus A_3$ stands for $(A_1 \oplus A_2) \oplus A_3$.

We use $\backslash$ for set difference. For all $A, B \subseteq \Sigma^*$ the *symmetric difference* $A \triangle B$ is given by $(A \backslash B) \cup (B \backslash A)$. For complements we use a tilde $\sim$, so that $\sim A$ and $\tilde{A}$ both stand for $\Sigma^* \backslash A$. We write $A \equiv^f B$ if $A \triangle B$ is finite, and denote by A^f the equivalence class of A under $\equiv^f$.

We call any collection $\mathscr{C}$ of languages a class. Given $\mathscr{C}$, we define $\mathscr{C}^f$ to be $\bigcup_{A \in \mathscr{C}} A^f$. $\mathscr{C}$ is *closed under finite variations* (cfv) if $\mathscr{C} = \mathscr{C}^f$. $\mathscr{C}$ is *somewhere-cfv* (scfv) if $\mathscr{C} \supseteq A^f$ for some language A. What one expects on hearing "somewhere-" breaks down only in the case of the *empty class* $\varnothing$, which is cfv but not scfv. Last, a class $\mathscr{C}$ is a *recursive translation* of another class $\mathscr{D}$ if for some recursive language B, $\mathscr{C} = \{A \mid A \triangle B \in \mathscr{D}\}$.

2.2. *Turing Machines*

We fix a standard (cf. '*acceptable*' in [Rog67]) recursive enumeration $[M_i]_{i=1}^{\infty}$ of deterministic off-line multi-tape *oracle* Turing machine acceptors, stipulating that "real" TMs have empty oracle set. We write $L(M_i^A)$ for the language accepted by M_i with oracle set $A \subseteq \Sigma^*$, and just $L(M)$ for $L(M^{\varnothing})$.

For M_i (with empty oracle set) we define predicates $\mathrm{Halt}(i, x, m)$ and $\mathrm{Accept}(i, x, m)$ on $\mathbb{N}^+ \times \Sigma^* \times \mathbb{N}^+$ to hold when M_i on input x respectively halts or accepts within m steps. We write $M_i(x)\downarrow$ if $(\exists m)\,\mathrm{Halt}(x, i, m)$, and $M_i(x)\uparrow$ otherwise. M_i is *total* if $M_i(x)\downarrow$ for all x. We put $\mathrm{TOT} := \{i \in \mathbb{N}^+ \mid M_i \text{ is total}\}$.

The *full index set* of any class $\mathscr{C}$ of r.e. languages is $I_{\mathscr{C}} := \{i \in \mathbb{N}^+ \mid L(M_i) \in \mathscr{C}\}$. Note that $\sim I_{\mathscr{C}}$ equals $I_{\mathbf{RE} \backslash \mathscr{C}}$. We shall also use the related index set $J_{\mathscr{C}} := \mathrm{TOT} \cap \sim I_{\mathscr{C}}$, which equals $\{i \in \mathbb{N}^+ \mid M_i \text{ is total and } L(M_i) \notin \mathscr{C}\}$.

For each i, we denote the partial recursive *time* and *space functions* of M_i by $t_i(\cdot)$ and $s_i(\cdot)$, respectively. Following [HU79] we assume that M_i reads all of its input plus the blank tape cell marking its end, so that $t_i(n) \geqslant n + 1$ for all n. When M_i takes the minimum $|x| + 1$ steps on all inputs x, we say M_i runs in *real time.* In Section 2.4 we relax this notion slightly for TMs which compute functions rather than recognize languages.

2.3. *Reducibility Relations*

For any class $\mathscr{F}$ of functions, the associated *many–one* reductibility relation $\leqslant_m^{\mathscr{F}}$ is defined for all $A, B \subseteq \Sigma^*$ by: $A \leqslant_m^{\mathscr{F}} B$ iff for some $f \in \mathscr{F}$ and all $x \in \Sigma^*$, $x \in A \Leftrightarrow f(x) \in B$. If f is 1–1 and $\mathscr{F}$ contains a left inverse of f, then we write $A \leqslant_i^{\mathscr{F}} B$. If in addition f is *length-increasing*, i.e., if $|f(x)| > |x|$ for all $x \in \Sigma^*$, then we write $A \leqslant_{li}^{\mathscr{F}} B$. Finally, if $\mathscr{F}$ contains $\lambda x.x$ and is closed under composition then the subclass $G(\mathscr{F})$ of bijections $g\colon \Sigma^* \to \Sigma^*$ with $g, g^{-1} \in \mathscr{F}$ forms a group, and if $f \in G(\mathscr{F})$ we write $A \equiv_{\text{iso}}^{\mathscr{F}} B$ and say A and B are $\mathscr{F}$-isomorphic.

For example, when $\mathscr{F} := \mathbf{FP}$, namely the class of total functions which are computable in polynomial time, then this recipe gives the familiar notion of polynomial-time many–one reducibility $\leqslant_m^p$, and also $\leqslant_i^p$, $\leqslant_{li}^p$, and the notion of *p-isomorphism* $\equiv_{\text{iso}}^p$ from [BeHa77]. For any reducibility $\leqslant_r$ and $A, B \subseteq \Sigma^*$, we write $A \equiv_r B$ if $A \leqslant_r B$ and $B \leqslant_r A$. The "**P**-isomorphism theorem" of [BeHa77] then states that $\equiv_{li}^p$ equals $\equiv_{\text{iso}}^p$.

Also studied in the literature are reducibilities arising from collections of oracle Turing machines. The collection of OTMs which run in polynomial time for all oracle sets defines polynomial-time Turing reducibility ($\leqslant_T^p$). As usual we write $\mathbf{P}^B$ for $\{A \subseteq \Sigma^* \mid A \leqslant_T^p B\}$, and define the relativized class $\mathbf{NP}^B$ similarly using nondeterministic OTMs (see [HU79]).

2.4. *Complexity Classes*

We name and define the complexity classes featured in this paper partly to suggest the mechanics of formalizing them. **L** is also called **LOGSPACE**. The language SAT of satisfiable Boolean formulas is not only known to be complete for **NP** under $\leqslant_m^p$, but also occupies the highest p-isomorphism equivalence class (ranked by $\leqslant_{li}^p$) among the **NP**-complete sets [BeHa77]. The *Berman–Hartmanis conjecture* asserts that this equivalence class is unique, i.e., that **NPC** = **NPI** as defined below. Otherwise we use no special properties of the language SAT:

$$\mathbf{RE} := \{L(M_i) \mid i \in \mathbb{N}^+\} \qquad \mathbf{P} := \{L(M_i) \mid (\exists k)[t_i \leqslant \lambda n.n^k + k]\}$$
$$\mathbf{REC} := \{L(M_i) \mid M_i \text{ is total}\} \qquad \mathbf{L} := \{L(M_i) \mid (\exists k)[s_i \leqslant \lambda n.k \cdot \log_2 n]\}$$
$$\mathbf{FIN} := \{L(M_i) \mid L(M_i) \text{ is finite}\} \qquad \mathbf{DLIN} := \{L(M_i) \mid (\exists k)[t_i \leqslant \lambda n.k \cdot n]\}$$
$$\mathbf{NP} := \{L \subseteq \Sigma^* \mid L \leqslant_m^p \text{SAT}\} \qquad \mathbf{NPC} := \{L \subseteq \Sigma^* \mid L \equiv_m^p \text{SAT}\}$$
$$\mathbf{NPI} := \{L \subseteq \Sigma^* \mid L \equiv_{\text{iso}}^p \text{SAT}\} \qquad \mathbf{EQ} := \{L \in \mathbf{REC} \mid \mathbf{NP}^L = \mathbf{P}^L\}$$
$$\mathbf{EXPTIME} := \bigcup_{c>0} \text{DTIME}[2^{cn}] \qquad \mathbf{EXP} := \bigcup_{c>0} \mathbf{DTIME}[2^{n^c}]$$
$$\mathbf{REG} := \{\textit{regular} \text{ languages}\} \qquad \mathbf{CFL} := \{\textit{context-free} \text{ languages}\}$$
$$\mathbf{DTISP}[n+1, \log_2 n] := \{L(M_i) \mid t_i = \lambda n.n + 1 \text{ and } (\exists k)[s_i \leqslant \lambda n.k \cdot \log_2 n]\}.$$
$$\mathbf{P\text{-}IMMUNE} := \{L \in \mathbf{REC} \backslash \mathbf{FIN} \mid (\forall A \subseteq \Sigma^*)\colon (A \in \mathbf{P} \wedge A \subseteq L) \Rightarrow A \in \mathbf{FIN}\}.$$

The levels of the *arithmetical hierarchy* are denoted by $\sum_k^0$ and $\prod_k^0$ ($k \geqslant 0$), where $\sum_0^0 = \prod_0^0 = \mathbf{REC}$, $\sum_1^0 = \mathbf{RE}$, etc. The analogous classes of the *polynomial hierarchy* receive a 'p' superscript, so that $\sum_1^p = \mathbf{NP}$, $\prod_1^p = \mathbf{co\text{-}NP}$, $\sum_2^p = \mathbf{NP}^{\mathbf{NP}}$, and so on. We write **PH** for $\bigcup_{k=1}^{\infty} \sum_k^p$.

For clarity we add an 'F' to the names of analogously defined classes of functions, e.g., **FREC**, **FP**, **FL**, $\mathbf{DLIN}_\mathbf{F}$, $\mathbf{DTISP}_\mathbf{F}[n+1, \log_2 n]$. The last of these is not closed under composition. To find a transitive reductibility relation which uses an intuitively minimal amount of time and space, we introduce:

DEFINITION 2.1. **RL** is the class of languages accepted by TM's M which run simultaneously in real time and log space, except that for some $l \in \mathbb{N}^+$ depending only on M, and all inputs x, M is allowed l "extra steps" after reaching the end of the input before it halts. $\mathbf{RL}_\mathbf{F}$ denotes the analogous class of functions. The associated many–one length-increasing-and-invertible reducibility relations are denoted by $\leqslant_m^{rl}$ and $\leqslant_{li}^{rl}$, respectively.

The initials stand for "real-time/log-space." The strict definition of real-time computation does not allow a transducer T to print after reaching the end of the input, but keeping to this would make simple reductions such as $\lambda x.x0$: $A \to A \oplus B$ impossible. We note that $\mathbf{RL}_\mathbf{F}$ is closed under composition (see [Reg86b]), so that the relations $\leqslant_m^{rl}$ and $\leqslant_{li}^{rl}$ are transitive.

LEMMA 2.1. (a) $\mathbf{RL} = \mathbf{DTISP}[n+1, \log_2 n]$.

(b) *For any* $h \in \mathbf{RL}_\mathbf{F}$ *and* $k \in \mathrm{Ran}(h)$, *the language* $h^{-1}(k)$ *is in* **RL**.

Part (a) follows from the construction of Theorem 2 in [HaSt65], as observed in [Ros67, p. 651]. Part (b) is virtually a special case of Theorem 4 on the same page of [Ros67]. Hence we rename '$\mathbf{DTISP}[n+1, \log_2 n]$' to '**RL**' in all that follows. On the other hand $\mathbf{DTISP}_\mathbf{F}[n+1, \log_2 n] \neq \mathbf{RL}_\mathbf{F}$ because the extra time taken to output $h(x)$ when $|h(x)| = |x| + l$ cannot be saved.

2.5. *Recursive and R.E. Presentations of Classes*

Under the pairing function $\langle \cdot, \cdot \rangle$ each language uniquely defines a class. For any $U \subseteq \Sigma^*$ and $k \in \mathbb{N}^+$ define the *kth projection* of U to be $U_k := \{x \mid \langle x, k \rangle \in U\}$, and set $\mathscr{P}\!r[U] := \{U_k \mid k \in \mathbb{N}^+\}$. If $\mathscr{C} = \mathscr{P}\!r[U]$ then U is a *universal language* for $\mathscr{C}$. A class $\mathscr{C}$ is *recursively presentable* (r.p.) if it has a recursive universal language, *r.e.-presentable* (or *presentable by r.e. indices*) if it has an r.e. universal language, and so on. The latter is equivalent to the more familiar stipulation that there be a recursive function σ: $\mathbb{N}^+ \to \mathbb{N}^+$ of TM indices such that $\mathscr{C} = \{L(M_{\sigma(i)}) \mid i \in \mathbb{N}^+\}$, and the former, to the same with $M_{\sigma(i)}$ total for each i. We similarly call a class $\mathscr{F}$ of total functions *recursively presentable* if there is a total recursive function $F(\cdot)$ such that $\mathscr{F} = \{\lambda x.F(\langle x, k \rangle) \mid k \in \mathbb{N}^+\}$.

If $\mathscr{C} \subseteq \mathscr{P}\!r[U]$ for some $U \in \mathbf{REC}$, then we say $\mathscr{C}$ is *bounded*; equivalently, $\mathscr{C} \subseteq \mathbf{DTIME}[t]$ for some recursive function t: $\mathbb{N} \to \mathbb{N}^+$. Otherwise, $\mathscr{C}$ is *unbounded.* A class $\mathscr{C}$ is *locally recursively presentable* (lrp) if $\mathscr{C} \cap \mathscr{D}$ is r.p. or empty for all classes $\mathscr{D}$ which are r.p. *and* cfv. **REC** and $\varnothing$ are not r.p., since one is unbounded and the other is empty, but both are lrp.

We appended "and cfv" to the definition of 'lrp' because the intersection of two r.p. classes $\mathscr{A}$ and $\mathscr{B}$, though nonempty, can fail to be r.p. or even r.e.-presentable [LaRo72]. As noted there, $\mathscr{A} \cap \mathscr{B}$ is r.p. *provided* it is scfv. [A-S84] restricts attention to lrp classes which are themselves cfv, and these shall be our basic objects of study. For all lrp cfv classes $\mathscr{A}$ and $\mathscr{B}$, both $\mathscr{A} \cup \mathscr{B}$ and $\mathscr{A} \cap \mathscr{B}$ are lrp and cfv.

We call a sequence $[\mathscr{C}_k]_{k \in K}$ (indexed in increasing order by the set $K \subseteq \mathbb{N}^+$) *recursively presented* if K is recursive, and there exists a recursive language U such that for all $k \in K$, the kth projection U_k is universal for $\mathscr{C}_k$.

LEMMA 2.2. *The family of r.p. classes is closed under* (a) *recursively presented unions,* (b) *recursive translations, and* (c) *the operation* $\mathscr{C} \mapsto \mathscr{C}^f$. *Likewise, if a given class* $\mathscr{C}$ *is r.e.-presentable then so is* $\mathscr{C}^f$.

Proof. (a) Given $[\mathscr{C}_k]_{k \in K}$ and U as above, and some language $A \in \bigcup_{k \in K} \mathscr{C}_k$, define $V := \{\langle x, \langle k, l \rangle \rangle \mid l \in \mathbb{N}^+$, and if $k \in K$ then $x \in (U_k)_l$, else $x \in A\}$. Then V is a recursive universal language for $\bigcup_{k \in K} \mathscr{C}_k$. The other parts are clear. ∎

The restatement of (a) for intersections fails, as we note in Section 3.3.3. The family of r.e.-presentable classes is not closed under recursive translations.

We extend the notion of recursive presentability to reducibility relations in the following manner.

DEFINITION 2.2. A relation $\leqslant_r$ on languages is an *effective reducibility relation* if there exists a recursive function σ of TM indices such that for all $A \in \mathbf{REC}$ and *total* TMs M_a accepting A, $M_{\sigma(a)}$ is total and accepts a universal language for $\{L \mid L \leqslant_r A\}$.

PROPOSITION 2.3. *Let* $\leqslant_r$ *be defined either by an r.p. function class* $\mathscr{F}$ *or a recursive enumeration* $[Q_k]_{k=1}^{\infty}$ *of total oracle Turing machines, as provided in Section* 2.3. *Then* $\leqslant_r$ *is an effective reducibility relation.*

Hence virtually all reducibilities that have been studied in the literature are effective, including some which lack the properties of the reflexivity and/or transitivity normally associated with the idea of a "reducibility relation."

2.6. *Formal Systems*

We let '$\mathfrak{F}$' stand for any formal system which is *sound, first-order*, and *recursively axiomatizable*. $\mathfrak{F}$ Proof($\cdot, \cdot$) stands for a recursive predicate such that for all *sentences* ψ over the language of $\mathfrak{F}$, $\mathfrak{F} \vdash \psi$ iff $(\exists d)$ $\mathfrak{F}$ Proof('ψ', d). Here $\mathfrak{F} \vdash \psi$ means "ψ is a theorem of $\mathfrak{F}$," and we suppose 'ψ' is ψ recoded over Σ^*. For short we say $\mathfrak{F}$ is *sound r.a.* We need $\mathfrak{F}$ to be sound more than merely *consistent* because we shall be using the truth of ψ, in cases where $\mathfrak{F}$ Proof('ψ', d) holds, in our constructions.

We distinguish between *positive* results of the form "Statement ψ is provable in system $\mathfrak{F}$," and *negative* results asserting the contrary. For the former, we make reasonable assumptions about the strength of $\mathfrak{F}$. A rock-bottom condition for $\mathfrak{F}$ to be "an acceptable formal theory of computation" (term and motivation from [Bak79]) is that the predicates Accept$(\cdot, \cdot, \cdot)$ and Halt$(\cdot, \cdot, \cdot)$ can be coded over the language of $\mathfrak{F}$ so that all true instances of them (and of their negations) are theorems of $\mathfrak{F}$. For short we then say $\mathfrak{F}$ *represents* T, since Accept and Halt are essentially subsumed by the *Kleene T-predicate*. PA and ZFC both represent T, as do sundry weak subsystems of them.

In stating "negative" results, we do not need to make any minimal-strength assumptions for $\mathfrak{F}$ at all. As remarked in [Háj79] we further need only suppose that $\mathfrak{F}$ is *sound over arithmetic*, i.e., that all sentences of $\mathfrak{F}$ referring only to natural numbers are true in the *standard model of arithmetic*, which suffices to encode TM computations for our purposes.

2.7. *Basic Concepts in Topology*

The notion of topology underlies many familiar concepts such as continuity, distance, separation, and connectedness. We shall interpret the delayed diagonalization technique (specifically, Theorem 5.1) using conditions on separating points by closed subspaces and on the connectedness of certain other subspaces. For, us the *points* are recursive languages and the *subspaces* are complexity classes, under a topology which arises from a generalization of recursive presentability. We use this topology to provide geometric intuition for the "shapes" of complexity classes, and seek an avenue for further results in complexity theory which draw on ideas developed in other branches of mathematics. Some background information:

- A *topological space* is a set $\mathcal{S}$ together with a collection $\mathfrak{O}$ of subsets of $\mathcal{S}$ which (i) includes $\varnothing$ and $\mathcal{S}$ itself, (ii) is closed under finite intersection, and (iii) is closed under arbitrary unions. The members of $\mathfrak{O}$ are the *open sets* of the topology. The complementary collection $\mathfrak{C} := \{\mathcal{S} \backslash \mathcal{O} \mid \mathcal{O} \in \mathfrak{O}\}$ is said to form a *closed topology* on $\mathcal{S}$; its members are the *closed sets.*

Although topologies are most commonly described via their open sets, we prefer to emphasize the closed sets, typically denoting spaces by $(\mathcal{S}, \mathfrak{C})$, or simply by $\mathcal{S}$ when $\mathfrak{C}$ is understood and vice versa. There is no loss in regarding $\mathfrak{C}$ and $\mathfrak{O}$ as the "same" topology on $\mathcal{S}$. To clarify references to the closed topology, we often say '$\mathscr{C}$ is closed in $\mathfrak{C}$' to mean the same thing as '$\mathscr{C} \in \mathfrak{C}$'. The *trivial* topologies on $\mathcal{S}$ are $(\mathcal{S}, \{\mathcal{S}, \varnothing\})$ and $(\mathcal{S}, \mathcal{P}(\mathcal{S}))$.

- A *closed basis* for $(\mathcal{S}, \mathfrak{C})$ is any collection $\mathfrak{B}$ of subsets of $\mathcal{S}$ whose closure under arbitrary intersections equals $\mathfrak{C}$. We also say $\mathfrak{B}$ *generates* $\mathfrak{C}$. A sufficient condition for $\mathfrak{B}$ to generate some closed topology on $\mathcal{S}$ is that $\mathfrak{B}$ be closed under finite $\bigcup$ and have $\varnothing$ and $\mathcal{S}$ as members. The complementary collection $\mathfrak{A} := \{\mathcal{S} \backslash \mathcal{B} \mid \mathcal{B} \in \mathfrak{B}\}$ is an *open basis* for the same topology.

A topology $\mathfrak{O}$ on $\mathscr{S}$ is *metrizable* if there is a *metric* ρ on $\mathscr{S}$ such that if one defines $\mathscr{A}_{X,\varepsilon} := \{Y \in \mathscr{S} \mid \rho(X, Y) < \varepsilon\}$ for all $X \in \mathscr{S}$ and $\varepsilon \geqslant 0$, then the collection $\{\mathscr{A}_{X,\varepsilon}\}$ is an open basis for $\mathfrak{O}$. We also say ρ *generates* $\mathfrak{O}$.

- For any $\mathscr{D} \subseteq \mathscr{S}$, the collection $\mathfrak{C} \upharpoonright \mathscr{D} := \{\mathscr{C} \cap \mathscr{D} \mid \mathscr{C} \in \mathfrak{C}\}$ forms a closed topology on $\mathscr{D}$. We variously call this the topology *induced* by $\mathfrak{C}$ onto $\mathscr{D}$, the *restriction* of $\mathfrak{C}$ to $\mathscr{D}$, or $\mathscr{D}$'s *own* topology. $\mathscr{D}$ is of course closed in $\mathfrak{C} \upharpoonright \mathscr{D}$, but need not be closed in $\mathfrak{C}$; its *closure in* $\mathfrak{C}$ equals $\bigcap \{\mathscr{C} \in \mathfrak{C} \mid \mathscr{C} \supseteq \mathscr{D}\}$. $\mathscr{D}$ is *dense* if its closure is the whole space $\mathscr{S}$. $\mathscr{D}$ is *nowhere-dense* if its closure contains no open set other than $\varnothing$, and *meager* if it is a countable union of nowhere-dense sets.

The space $(\mathscr{S}, \mathfrak{C})$ is *compact* if whenever $\mathfrak{C}' \subseteq \mathfrak{C}$ is a collection of closed sets haing empty intersection, some finite subcollection $\mathfrak{C}''$ of $\mathfrak{C}'$ also has empty intersection. (This is equivalent to the more common wording that every *open cover* of $\mathscr{S}$ has a finite subcover.) A subset $\mathscr{D} \subseteq \mathscr{S}$ is *compact* (*in* $\mathfrak{C}$) if the topological subspace $(\mathscr{D}, \mathfrak{C} \upharpoonright \mathscr{D})$ is compact.

Some Examples. (1) The familiar Euclidean metric generates the *usual topology* on $\mathbb{R}^n$ ($n \geqslant 1$). In the case of $\mathbb{R}$, the basic open sets are the open intervals, together with $\varnothing$. Since the "open balls" $\mathscr{A}_{X,\varepsilon}$ for rational values of ε and X suffice to generate this topology, it has a countable basis. A subspace $\mathscr{D} \subseteq \mathbb{R}^n$ is compact in the topology iff $\mathscr{D}$ is closed and bounded.

(2) The following metric on $\mathscr{P}(\Sigma^*)$ generates a topology $\mathfrak{I}$ which has been previously studied in complexity theory: For all $A, B \subseteq \Sigma^*$ put $\rho(A, B) := 2^{-n}$, where $\text{bin}^{-1}(n)$ equals the least string in $A \triangle B$; $\rho(A, B) = 0$ iff $A = B$. The space $(\mathscr{P}(\Sigma^*), \mathfrak{I})$ is compact, as is every closed subspace. The class of oracles A making $\mathbf{NP}^A = \mathbf{P}^A$ is meager in $\mathfrak{I}$, meaning that it is intuitively "small." For this and related results showing interest in $\mathfrak{I}$, see Section 4 of [BeGi 81], and also [Meh73; Dowd82; Lutz87]. Spaces related to $\mathfrak{I}$ have also been studied in recursion theory—see [Cu80, Lav77, Rog67; pp. 265–271, 339ff].

The topological properties we focus on are separation and connectedness. We itemize the definitions for future reference.

Definition 2.3. (two of the "*classical separation conditions*"). $(\mathscr{S}, \mathfrak{C})$ is

(1) a T_1-*space* if for all distinct points $A_1, A_2 \in \mathscr{S}$, there exist closed sets $\mathscr{C}_1, \mathscr{C}_2$ such that $A_1 \notin \mathscr{C}_1$, $A_2 \notin \mathscr{C}_2$, while $A_1 \in \mathscr{C}_2$ and $A_2 \in \mathscr{C}_1$.

(2) a T_2-*space* (or *Hausdorff space*) if for all distinct points $A_1, A_2 \in \mathscr{S}$, there exist $\mathscr{C}_1, \mathscr{C}_2 \in \mathfrak{C}$ such that $\mathscr{C}_1 \cup \mathscr{C}_2 = \mathscr{S}$, $A_1 \notin \mathscr{C}_1$, and $A_2 \notin \mathscr{C}_2$.

Clearly every T_2-space is T_1. Equivalent conditions are (T_1) that for all $A \in \mathscr{S}$ the set $\{A\}$ is closed in $\mathfrak{C}$, and (T_2) that every distinct pair of points can be *separated* by disjoint open sets. The latter is a necessary condition for $(\mathscr{S}, \mathfrak{C})$ to be metrizable. Hence $\mathbb{R}$ and $\mathfrak{I}$ are T_2.

DEFINITION 2.4. (a) A topological space $(\mathscr{S}, \mathfrak{C})$ is *connected* if whenever $\mathscr{S} = \mathscr{C}_1 \cup \mathscr{C}_2$ for some disjoint closed sets $\mathscr{C}_1, \mathscr{C}_2$ then either $\mathscr{S} = \mathscr{C}_1$ or $\mathscr{S} = \mathscr{C}_2$.

(b) $(\mathscr{S}, \mathfrak{C})$ is *hyperconnected* if whenever $\mathscr{S} = \mathscr{C}_1 \cup \mathscr{C}_2$ for some closed sets $\mathscr{C}_1, \mathscr{C}_2$ then either $\mathscr{S} = \mathscr{C}_1$ or $\mathscr{S} = \mathscr{C}_2$.

(c) A subspace $\mathscr{D} \subseteq \mathscr{S}$ is connected or hyperconnected *in* $\mathfrak{C}$ according to whether $\mathscr{D}$ is connected or hyperconnected in the topology induced from $(\mathscr{S}, \mathfrak{C})$.

Put another way, a *hyper*connected space is one which cannot be written as a nontrivial union of closed subspaces. In algebraic geometry such a space is called *irreducible*. A simple example is a space whose closed sets comprise an infinite set $\mathscr{S}$ together with all of its finite subsets. The unfamiliarity of this concept largely owes to the fact that the only hyperconnected subsets of $\mathbb{R}^n$ $(n > 0)$ are single points. (The empty set is conventionally not counted.)

We collect without proof some helpful observations. Part (a) provides a useful shortcut in defining hyperconnected subspaces.

PROPOSITION 2.4. *Let $\mathfrak{C}$ form a closed topology on $\mathscr{S}$, and let $\mathscr{D} \subseteq \mathscr{S}$.*

(a) *$\mathscr{D}$ is hyperconnected in $\mathfrak{C} \Leftrightarrow$ whenever $\mathscr{D} \subseteq \mathscr{C}_1 \cup \mathscr{C}_2$ for some $\mathscr{C}_1, \mathscr{C}_2 \in \mathfrak{C}$, then either $\mathscr{D} \subseteq \mathscr{C}_1$ or $\mathscr{D} \subseteq \mathscr{C}_2$.*

(b) *It is* not *in general the case that $\mathscr{D}$ is connected in $\mathfrak{C}$ if whenever $\mathscr{D} \subseteq \mathscr{C}_1 \cup \mathscr{C}_2$ for some disjoint $\mathscr{C}_1, \mathscr{C}_2 \in \mathfrak{C}$, then either $\mathscr{D} \subseteq \mathscr{C}_1$ or $\mathscr{D} \subseteq \mathscr{C}_2$.*

(c) *If the space $(\mathscr{D}, \mathscr{S} \upharpoonright \mathscr{D})$ is hyperconnected then it is not T_2.*

(d) *Let $\mathfrak{C}_0$ form another closed topology on $\mathscr{S}$, where $\mathfrak{C}_0 \supseteq \mathfrak{C}$. If $\mathscr{D}$ is connected (resp. hyperconnected) in $\mathfrak{C}_0$, then $\mathscr{D}$ is connected (hyperconnected) in $\mathfrak{C}$.*

The topology $\mathfrak{R}$ studied in this paper is like none of the spaces referred to above. Classes such as **P** and **NP** are closed in $\mathfrak{R}$, but are not closed in $\mathfrak{I}$ (in fact, they are dense in $\mathfrak{I}$ simply because they contain all finite sets). The class $\mathbf{EQ} := \{A \in \mathbf{REC} \mid \mathbf{NP}^A = \mathbf{P}^A\}$ is also closed in $\mathfrak{R}$, and all closed classes other than **REC** are not only meager but also nowhere-dense in $\mathfrak{R}$. $\mathfrak{R}$ is not generated by any metric, to which we ascribe the inability of researchers to define a *norm* on recursive languages A which corresponds to any intuitive notion of the complexity of A. The only spaces we know to bear some kinship to $\mathfrak{R}$ are several which arise in denotational semantics; we discuss them in Section 7.

Before defining $\mathfrak{R}$ we find it convenient to introduce a larger topology $\mathfrak{R}_0$, which is generated by the collection of classes that are locally recursively presentable and closed under finite variations.

3. THE TOPOLOGIES $\mathfrak{R}$ AND $\mathfrak{R}_0$

Let $\mathfrak{B}_{\mathrm{lrp}}$ denote the collection of classes $\mathscr{C} \subseteq \mathbf{REC}$ which are locally recursively presentable and closed under finite variations. In Section 2.5 we observed that $\mathfrak{B}_{\mathrm{lrp}}$

is closed under finite $\bigcup$ and contains both **REC** and $\varnothing$. Hence it forms a closed basis for a topology on **REC**.

DEFINITION 3.1. $\mathfrak{R}_0$ denotes the closed topology on **REC** generated by $\mathfrak{B}_{\mathrm{lrp}}$.

We also regard $\mathfrak{R}_0$ as a topology on $\mathbf{REC}/\equiv^f$, since every member of $\mathfrak{R}_0$ is cfv. Since $\mathfrak{B}_{\mathrm{lrp}}$ is also closed under finite $\bigcap$, it is also an open basis for a topology on $\mathbf{REC}/\equiv^f$, but this latter topology is trivial. The reason is that the class A^f is r.p. for all $A \in \mathbf{REC}$ (so $A^f \in \mathfrak{B}_{\mathrm{lrp}}$), and every cfv class $\mathscr{C}$ equals the union of the classes A^f for all $A \in \mathscr{C}$. By the same token, $(\mathbf{REC}/\equiv^f, \mathfrak{R}_0)$ is a T_1-space. That it is not a T_2-space is implied by the following result, which is subsumed by Theorem 5.6 later on.

PROPOSITION 3.1. *The space* $(\mathbf{REC}, \mathfrak{R}_0)$ *is hyperconnected.*

The property of local recursive presentability is not effective. By "effective" we require that there should be a recursive procedure which, given a class $\mathscr{C} \in \mathfrak{B}_{\mathrm{lrp}}$, an r.p. class $\mathscr{D}$, and a language $A \in \mathscr{C} \cap \mathscr{D}$ "witnessing" that $\mathscr{C} \cap \mathscr{D}$ is nonempty, constructs a recursive presentation of $\mathscr{C} \cap \mathscr{D}$. We seek a replacement for $\mathfrak{B}_{\mathrm{lrp}}$ which does satisfy this requirement, and which provides a more acute generalization of recursive presentability for unbounded classes.

Moreover, $\mathfrak{B}_{\mathrm{lrp}}$ has cardinality $2^{\aleph_0}$, and we show below that $\mathfrak{R}_0$ has no countable basis whatsoever. We introduce a countable subcollection of $\mathfrak{B}_{\mathrm{lrp}}$ which generates a topology that we regard as "the effective part" of $\mathfrak{R}_0$, and which more closely captures the notion of provable properties of languages.

3.1. *The "R.R. Property" and the Definition of* $\mathfrak{R}$

The property in question was originally intended to generalize the concept of a witness function. The idea is that the "witness" to the assertion that two languages A and B are unequal need no longer be a string $z \in A \triangle B$ but can be any proof of '$A \neq B$' in a given sound r.a. formal system $\mathfrak{F}$.

All recursively presentable classes $\mathscr{C}$ give rise to recursive witness functions in the following sense: Let U be a recursive universal language for $\mathscr{C}$. For all i, j define $S(i, j)$ to be the assertion '$L(M_i) = U_j$'; then $L(M_i) \in \mathscr{C} \Leftrightarrow (\exists j)\, S(i, j)$ for all i. Whenever M_i is total and $L(M_i) \notin \mathscr{C}$, one can recursively compute from any j a witness z which refutes $S(i, j)$ by running M_i on successie inputs $z := \lambda, 0, 1, 00, \ldots$, and testing whether $z \in L(M_i) \triangle U_j$. Carrying out the above idea for the case of general predicates $S(\cdot, \cdot)$ and refutations of S in the formal system $\mathfrak{F}$ leads to

DEFINITION 3.2. A class $\mathscr{C}$ has the *recursive refutability* (r.r.) *property* with respect to a given formal system $\mathfrak{F}$ if for some predicate $S(\cdot, \cdot)$ over $\mathfrak{F}$,

$$\text{(a)} \quad \text{for all } i\text{: } L(M_i) \in \mathscr{C} \Leftrightarrow (\exists j)\, S(i, j), \tag{3.1}$$

$$\text{(b)} \quad \text{for all } i, j\text{: } M_i \text{ total} \wedge \neg S(i, j) \Rightarrow (\exists d)\, \mathfrak{F}\ \mathrm{Proof}(\neg S(i, j), d). \tag{3.2}$$

We call S an *r.r. predicate* for $\mathscr{C}$. Using the pairing function $\langle \cdot, \cdot \rangle$ we may extend the second argument to any number of variables. Although $\mathscr{C}$ can have many r.r. predicates, we denote some choice by $S_{\mathscr{C}}$ for emphasis. If $\mathscr{C} \subseteq \mathbf{RE}$ then $S_{\mathscr{C}}$ uniquely determinates $\mathscr{C}$. It is actually sensible to restrict the definition to classes $\mathscr{C}$ containing recursive languages only, but we find it technically more informative to stipulate '$\mathscr{C} \subseteq \mathbf{REC}$' only where needed.

Besides the motivation regarding witness functions, and the feeling that it gives the "right" generalization of recursive presentability for unbounded classes, the r.r. property has two features which we find most interesting, even surprising. First, the collection of classes $\mathscr{C} \subseteq \mathbf{REC}$ having the r.r. property with respect to $\mathfrak{F}$ is actually independent of $\mathfrak{F}$ so long as $\mathfrak{F}$ is sound, recursively axiomatized, and meets the very minimal conditions outlined in Section 2.6. Second, it is not a trivial matter to show that the class of recursive languages has the r.r. property, and this is stronger than the well-known result that **REC** "has an r.e. basis", i.e., is r.e.-presentable. The standard definition of **REC** via '$L(M_i)$ is recursive' $\leftrightarrow (\exists j)$['$M_j$ is total' and '$L(M_i) = L(M_j)$'] does not yield an r.r. predicate $S(i, j)$. We collect the results from [Reg 83b] which show these facts, and which help set up the topology $\mathfrak{R}$. Recall the definition of $J_{\mathscr{C}}$ as $\{i \mid M_i \text{ is total and } L(M_i) \notin \mathscr{C}\}$ in Section 2.2.

THEOREM 3.2. *With respect to any sound r.a. $\mathfrak{F}$ which represents T,*

(a) **REC** *has the r.r. property.*

(b) *For any class $\mathscr{C} \subseteq \mathbf{REC}$, $\mathscr{C}$ has the r.r. property $\Leftrightarrow J_{\mathscr{C}} \in \prod_2^0$.*

(c) *If $\mathscr{C}$ and $\mathscr{D}$ have the r.r. property, then so do $\mathscr{C} \cup \mathscr{D}$ and $\mathscr{C} \cap \mathscr{D}$.*

Proof. (sketch). (a) For all $i, j, c \in \mathbb{N}^+$ define the following predicates. (Here 'TB' stands for "time bound," $x, y \in \Sigma^*$ refer to input strings, and $m, n \in \mathbb{N}^+$ are numbers of steps in the computations of M_i on x and M_j on y.)

$$\mathrm{TB}(i, j, c) := (\forall x, m, n)\colon \left[\neg \mathrm{Halt}(x, j, m) \;\&\; \bigwedge_{y \leqslant x} \mathrm{Halt}(i, y, n) \right] \rightarrow m + 1 \leqslant n + c.$$

$$S_{\mathbf{REC}}(i, j, c) := \text{`}L(M_i) = L(M_j)\text{'} \wedge \text{`}M_j \text{ is total'} \wedge \mathrm{TB}(i, j, c).$$

To verify that $S_{\mathbf{REC}}(\cdot, \cdot, \cdot)$ is an r.r.-predicate for **REC**, one may check the following claims (i)–(v). Full details, including the way to formalize $S_{\mathbf{REC}}$ over $\mathfrak{F}$ using the T-predicate, may be found in [Reg86b]; cf. also [Reg83a, b].

Claims (left to reader). (i) If M_i is total and $\mathrm{TB}(i, j, c)$ holds for some j and c, then M_j is total, and its running time on all inputs is bounded by that of M_i up to the additive constant c.

(ii) If M_i not total, then for any j such that M_j is total, $\mathrm{TB}(i, j, c)$ can be satisfied by choosing c large enough.

(iii) For all i, $L(M_i) \in \mathbf{REC} \Leftrightarrow (\exists j, c)\, S_{\mathbf{REC}}(i, j, c)$.

(iv) For all $i, j, c \in \mathbb{N}^+$, $\neg \mathrm{TB}(i, j, c) \Rightarrow (\exists d)\, \mathfrak{F}\, \mathrm{Proof}(\neg \mathrm{TB}(i, j, c), d)$.

(v) If M_i is total, $\mathrm{TB}(i, j, c)$ holds, but $S_{\mathbf{REC}}(i, j, c)$ fails, then $(\exists d)\, \mathfrak{F}\, \mathrm{Proof}(\text{'}L(M_i) \neq L(M_j)\text{'}, d)$.

Here (iii) follows from (i) and (ii), and satisfies (3.1) in Definition 3.2. Parts (iv) and (v) together satisfy (3.2). This finishes (a).

(b) Suppose $\mathscr{C}$ has the r.r. property with respect to $\mathfrak{F}$. Let $S_{\mathscr{C}}(\cdot, \cdot)$ be an r.r. predicate for $\mathscr{C}$, and for short put $\mathrm{TOT} := \{i \in \mathbb{N}^+ \mid M_i \text{ is total}\}$. Then for all i,

$$i \in J_{\mathscr{C}} \Leftrightarrow i \in \mathrm{TOT} \wedge (\forall j)(\exists d)\, \mathfrak{F}\, \mathrm{Proof}(\neg S_{\mathscr{C}}(i, j), d), \tag{3.3}$$

since $\mathscr{C} \subseteq \mathbf{REC}$. Since $\mathrm{TOT} \in \prod_2^0$ and $\mathfrak{F}\, \mathrm{Proof}(\cdot, \cdot)$ is recursive, $J_{\mathscr{C}} \in \prod_2^0$.

Conversely, suppose $J_{\mathscr{C}} \in \prod_2^0$. This means that there exists a recursive predicate $R(\cdot, \cdot, \cdot)$ such that for all i, $i \in J_{\mathscr{C}} \Leftrightarrow (\forall a)(\exists b)\, R(i, a, b)$. (Formally we can encode this as a decidable predicate over $\mathfrak{F}$ using $\mathrm{Accept}(\langle \cdot, \langle \cdot, \cdot \rangle \rangle, r, \cdot)$ for some total TM M_r computing R.) Then define for all $i, j, c, a \in \mathbb{N}^+$:

$$S_{\mathscr{C}}(i, j, c, a) := S_{\mathbf{REC}}(i, j, c) \wedge (\forall b) \neg R(i, a, b). \tag{3.4}$$

An argument similar to (a) shows that $S_{\mathscr{C}}$ is an r.r. predicate for $\mathscr{C}$.

(c) If $S_{\mathscr{C}}(i, j)$ and $S_{\mathscr{D}}(i, k)$ are r.r. predicates for $\mathscr{C}$ and $\mathscr{D}$, respectively, then $\lambda(i, j, k).S_{\mathscr{C}}(i, j) \vee S_{\mathscr{D}}(i, k)$ is an r.r. predicate for $\mathscr{C} \cup \mathscr{D}$, and similarly $\lambda(i, j, k).S_{\mathscr{C}}(i, j) \wedge S_{\mathscr{D}}(i, k)$ is an r.r. predicate for $\mathscr{C} \cap \mathscr{D}$. ▌

Since the definitions of $J_{\mathscr{C}}$ and the class $\prod_2^0$ do not involve $\mathfrak{F}$, the r.r. property is independent of $\mathfrak{F}$ when we restrict attention to classes $\mathscr{C} \subseteq \mathbf{REC}$. We henceforth drop the clause "with respect to $\mathfrak{F}$." The next result shows that the r.r. property is an effective form of local recursive presentability.

THEOREM 3.3 [Reg83b]. (a) *Let $\mathscr{C}$ be an scfv class contained in* **REC**. *Then $\mathscr{C}$ is recursively presentable $\Leftrightarrow$ $\mathscr{C}$ is bounded and has the r.r. property.*

(b) *Moreover, given an r.r. predicate $S_{\mathscr{C}}$ for $\mathscr{C}$, a recursive language D, and a language A such that $A^f \subseteq \mathscr{C} \cap \mathscr{P}\imath[D]$, one can uniformly and effectively obtain a recursive universal language for $\mathscr{C} \cap \mathscr{P}\imath[D]$.*

Proof (sketch). (a) That $\mathscr{C}$ is scfv means that there is some $A \in \mathbf{REC}$ such that $A^f \subseteq \mathscr{C}$, and that $\mathscr{C}$ is bounded means that $\mathscr{C} \subseteq \mathscr{P}r[V]$ for some $V \in \mathbf{REC}$. For $(\Leftarrow)$, let $R(\cdot, \cdot, \cdot)$ be a recursive predicate such that for all i, $i \in J_{\mathscr{C}} \Leftrightarrow (\forall a)(\exists b)\, R(i, a, b)$. As remarked in Section 2.5 one can find a recursive function σ: $\mathbb{N}^+ \rightarrow \mathbb{N}^+$ such that for all i, $M_{\sigma(i)}$ is total and accepts V_i. Then define

$$\begin{aligned} U := \{\langle x, \langle i, a \rangle \rangle \mid &(\forall b \leqslant x)[\neg R(\sigma(i), a, b) \wedge x \in V_i] \\ &\vee (\exists b \leqslant x)[R(\sigma(i), a, b) \wedge x \in A]\}. \end{aligned} \tag{3.5}$$

Then U is a recursive universal language for $\mathscr{C}$.

For ($\Rightarrow$), let U be a recursive universal language for $\mathscr{C}$, and define $S(i, j) :=$ '$L(M_i) = U_j$' for all i and j. As observed at the outset of this subsection, $S(\cdot, \cdot)$ is an r.r. predicate for $\mathscr{C}$, This finishes (a).

(b) This uses the idea of (a), verifying that all of the steps can be performed recursively. For full details see [Reg86b]. ∎

DEFINITION 3.3. $\mathfrak{B}_{rr}$ is the collection of classes $\mathscr{C} \subseteq \mathbf{REC}$ which are r.r. and cfv. $\mathfrak{R}$ is the collection of classes which are arbitrary intersections of members of $\mathfrak{B}_{rr}$.

PROPOSITION 3.4. *$\mathfrak{R}$ forms a closed topology on* **REC**. *$\mathfrak{B}_{rr}$ is a countable closed basis for this topology.*

Proof. Since there are only countably many (r.r.) predicates, $\mathfrak{B}_{rr}$ is countable. From the definitions in Section 2.7, it suffices to verify (i) that $\mathfrak{B}_{rr}$ is closed under finite $\bigcup$, (ii) that $\mathfrak{B}_{rr}$ contains **REC**, and (iii) that $\mathfrak{B}_{rr}$ contains the empty class $\varnothing$. The first follows from Theorem 3.2(c), while (ii) is Theorem 3.2(a). For (iii), define $S_{\varnothing}(i, j)$ to be '$0 = 1$' regardless of i and j. Then $S_{\varnothing}$ is an r.r. predicate for the empty class. ∎

We also regard $\mathfrak{R}$ as a topology on the space $\mathbf{REC}/\equiv^f$, as with $\mathfrak{R}_0$. Similarly, $(\mathbf{REC}/\equiv^f, \mathfrak{R})$ is a T_1-space which is not a T_2-space, and the *open* topology generated by $\mathfrak{B}_{rr}$ has no interest because it is trivial.

3.2. *Comparing $\mathfrak{R}$ and $\mathfrak{R}_0$*

We collect some results which justify both our calling $\mathfrak{R}$ the effective part of the topology $\mathfrak{R}_0$, and our shifting attention to $\mathfrak{R}$ in what follows. The noneffective diagonalization used in Theorem 3.6 may hold special interest for recursion theorists.

PROPOSITION 3.5. (a) *Every class which is closed in the topology* (**REC**, $\mathfrak{R}$) *is closed in* (**REC**, $\mathfrak{R}_0$). *That is,* $\mathfrak{R} \subseteq \mathfrak{R}_0$.

(b) *For all bounded classes $\mathscr{D}$, $\mathfrak{R}_0 \upharpoonright \mathscr{D} = \mathfrak{R} \upharpoonright \mathscr{D}$.*

(c) *Every class which is connected* (*resp. hyperconnected*) *in $\mathfrak{R}_0$ is connected* (*hyperconnected*) *in $\mathfrak{R}$.*

Proof. (a) By Theorem 3.3(a), the basis $\mathfrak{B}_{rr}$ for $\mathfrak{R}$ is contained in $\mathfrak{B}_{\mathrm{lrp}}$. Hence $\mathfrak{R} \subseteq \mathfrak{R}_0$.

(b) Since $\mathscr{D}$ is bounded, $\mathscr{D}$ is contained in some r.p. cfv class $\mathscr{E}$. Let $\mathscr{A}$ be closed in the topology induced by $\mathfrak{R}_0$ on $\mathscr{D}$, i.e., $\mathscr{A} = \mathscr{D} \cap \mathscr{C}$ for some $\mathscr{C} \in \mathfrak{R}_0$. There exists a (possibly uncountable) indexed collection $\{\mathscr{C}_\alpha\}$ of lrp cfv classes such that $\mathscr{C} = \bigcap_\alpha \mathscr{C}_\alpha$. Since $\mathscr{D} \subseteq \mathscr{E}$, $\mathscr{A}$ also equals $\mathscr{D} \cap (\bigcap_\alpha (\mathscr{C}_\alpha \cap \mathscr{E}))$. For any index α, $\mathscr{C}_\alpha \cap \mathscr{E}$

is either r.p. or empty, and hence closed in $\mathfrak{R}$. Thus also $\bigcap_\alpha (\mathscr{C}_\alpha \cap \mathscr{E}) \in \mathfrak{R}$. This makes $\mathscr{A}$ closed in $\mathfrak{R}\restriction\mathscr{D}$. Since $\mathscr{A}$ is arbitrary we conclude $(\mathfrak{R}_0\restriction\mathscr{D}) \subseteq (\mathfrak{R}\restriction\mathscr{D})$. The reverse inclusion follows from (a).

(c) This follows from (a) and Proposition 2.4(d). ∎

THEOREM 3.6. (a) *The space* (**REC**, $\mathfrak{R}_0$) *has no countable basis.*

(b) *There are* $2^{\aleph_0}$ *classes which are closed in* (**REC**, $\mathfrak{R}_0$) *but not in* (**REC**, $\mathfrak{R}$).

(c) *There are* $2^{\aleph_0}$ *classes which are closed in* (**REC**, $\mathfrak{R}$).

Proof. (a) Let $\mathfrak{B}$ be any countable collection of classes which are closed in (**REC**, $\mathfrak{R}_0$). To prove the theorem, we construct an lrp cfv class $\mathscr{A} \neq$ **REC** which cannot be written as an intersection of members of $\mathfrak{B}$. W.l.o.g. we may assume that **REC** itself is not a member of $\mathfrak{B}$.

Let $[f_i]_{i=1}^\infty$ enumerate **FREC** (necessarily nonrecursively), and let $[\mathscr{B}_j]_{j=1}^\infty$ enumerate $\mathfrak{B}$. For all k and n, define $g_k(n) := \max\{f_i(n) \mid i \leqslant k\}$. Then each g_k is total recursive, so **DTIME**$[g_k]$ is r.p., and hence bounded and closed in $\mathfrak{R}_0$. Also **DTIME**$[g_k] \subseteq$ **DTIME**$[g_l]$ whenever $k \leqslant l$, and **REC** equals $\bigcup_{k=1}^\infty$ **DTIME**$[g_k]$.

Since **REC** is hyperconnected in $\mathfrak{R}_0$, **REC** cannot be written as a finite union of proper subclasses which are closed in $\mathfrak{R}_0$. Hence for any k, there exists a language A_k in **REC**$\backslash(\bigcup_{j=1}^k \mathscr{B}_j \cup$ **DTIME**$[g_k])$. Define $\mathscr{A}_k := A_k^f$ for each k, and finally define $\mathscr{A} := \bigcup_{k=1}^\infty \mathscr{A}_k$. Clearly $\mathscr{A}$ is cfv and $\mathscr{A} \neq$ **REC**.

To see that $\mathscr{A}$ is closed in $\mathfrak{R}_0$, consider that any r.p. cfv class $\mathscr{D}$ is contained in **DTIME**$[g_k]$ for some k. Since $\mathscr{A} \cap$ **DTIME**$[g_k]$ contains only the languages $A_1, \ldots, A_k$ up to $\equiv^f$, $\mathscr{A} \cap \mathscr{D}$ is r.p.-or-empty for the simple reason of being finite modulo $\equiv^f$. So $\mathscr{A}$ is in fact lrp. For all k, however, A_k is a language in $\mathscr{A}\backslash\mathscr{B}_k$. Thus $\mathscr{A}$ is not an intersection of members of $\mathfrak{B}$.

(b) With reference to (a), take $\mathfrak{B}$ to be $\mathfrak{B}_{rr}$ minus $\{$**REC**$\}$. At each step in the construction of $\mathscr{A}$, there are at least two ways to choose A_k which are not $\equiv^f$-equivalent (in fact, infinitely many ways). To avoid duplicating choices, we need only ensure that A_k is in **REC**$\backslash(\bigcup_{j=1}^k \mathscr{B}_j \cup$ **DTIME**$[g_k] \cup \bigcup_{j=1}^{k-1} \mathscr{A}_j)$. So there are $2^{\aleph_0}$ different ways to choose a class $\mathscr{A}$ which is not closed in $\mathfrak{R}$.

(c) We claim that there are $2^{\aleph_0}$ ways to choose a sequence $C_1, C_2, C_3, \ldots$ of members of **REC** such that (i) $C_i \not\equiv^f C_j$ whenever $i \neq j$, and (ii) there is a sequence $\mathscr{B}_1, \mathscr{B}_2, \mathscr{B}_3, \ldots$ of classes closed in (**REC**, $\mathfrak{R}$) such that each $\mathscr{B}_k$ contains $\{C_l \mid l \geqslant k\}$, but $\bigcap_{k=1}^\infty \mathscr{B}_k = \varnothing$. If we then put $\mathscr{C}_k := C_k^f$ and $\mathscr{E}_k := \mathscr{B}_k \cup \bigcup_{j=1}^{k-1} \mathscr{C}_k$ for each k, then each $\mathscr{E}_k$ is closed in $\mathfrak{R}$, and $\bigcap_{k=1}^\infty \mathscr{E}_k$ equals just the union of the classes $\mathscr{C}_k$. Hence with $\mathscr{C} := \bigcup_{k=1}^\infty \mathscr{C}_k$ we have that $\mathscr{C}$ is closed in $\mathfrak{R}$.

With reference to the construction in (a), define $C_1 := A_1$, and for $k > 1$, $C_k := A_1 \oplus \cdots \oplus A_k$. By the argument of (b) there are $2^{\aleph_0}$-many different cfv classes $\mathscr{C} := \bigcup_{k=1}^\infty C_k^f$, each contained in **REC**, which can be constructed in this manner. Also define $\mathscr{B}_k := \{L \in$ **REC** $\mid C_k \leqslant_T^p L\}$ for each k. Theorem 3.9 (below) implies that each $\mathscr{B}_k$ has the r.r. property (otherwise, our use of '$\leqslant_T^p$' is arbitrary). Since we have $l \geqslant k \Rightarrow C_k \leqslant_T^p C_l$ for all $k, l \in \mathbb{N}^+$, each $\mathscr{B}_k$ contains $\{C_l \mid l \geqslant k\}$, and since $\bigcap_{k=1}^{\infty} \mathscr{B}_k = \varnothing$, the conclusion follows. ∎

Figure 3.1 show intuitively why the sequence $[C_k]_{k=1}^{\infty}$ yields a class $\mathscr{C}$ which is closed in $\mathfrak{R}$, whereas $[A_k]_{k=1}^{\infty}$ does not. We inquire, without the *continuum hypothesis*, whether $\mathfrak{R}_0$ has any basis of cardinality strictly less than $2^{\aleph_0}$. The construction in (c) extends to show that no complexity class containing a small subclass of the regular languages is compact in the topology $\mathfrak{R}$.

THEOREM 3.7. *Let $\mathscr{D}$ be any cfv class which contains (a recursive translation of) $\{\mathbf{0^*1^k} \mid k \geqslant 0\} \cup \{\mathbf{1^*0^k} \mid k \geqslant 0\}$. Then* (a) $\mathrm{card}(\mathfrak{R} \upharpoonright \mathscr{D}) = 2^{\aleph_0}$, *and* (b) *$\mathscr{D}$ is not compact in $\mathfrak{R}$. The same hold for $\mathfrak{R}_0$ in place of $\mathfrak{R}$.*

Proof (sketch). With reference to the proof of Theorem 3.6(c), take $\mathscr{B}_k := \{\mathbf{0^*1^l} \mid l \geqslant k\}^f \cup \{\mathbf{1^*0^l} \mid l \geqslant k\}^f$, and for each k choose C_k to be either $\mathbf{0^*1^k}$ or $\mathbf{1^*0^k}$. The choices yield $2^{\aleph_0}$ closed classes of $\mathfrak{R}$ contained in $\mathscr{D}$, while (b) holds because $\bigcap_{k=1}^{\infty} \mathscr{B}_k = \varnothing$ while no finite subintersection is empty. The rest follows because $\mathfrak{B}_{\mathrm{rr}}$ and $\mathfrak{B}_{\mathrm{lrp}}$ are closed under recursive translations. ∎

For example, **NPI** is not compact in $\mathfrak{R}$. Thus the analogy "closed + bounded = compact" with the usual topology on $\mathbb{R}$ disappointingly fails.

In sum, it may be surprising that $\mathfrak{R}$ and $\mathfrak{R}_0$ are the same on all bounded classes, and yet $\mathfrak{R}$ has a countable basis (namely $\mathfrak{B}_{rr}$) while $\mathfrak{R}_0$ does not. A countable basis for an uncountable space often gives key insight into its structure. This leads us to investigate $\mathfrak{R}$ further in this and the next section. Since the applications to come in Section 6 deal with bounded classes, we shall not be concerned about the difference between $\mathfrak{R}$ and $\mathfrak{R}_0$ in what follows.

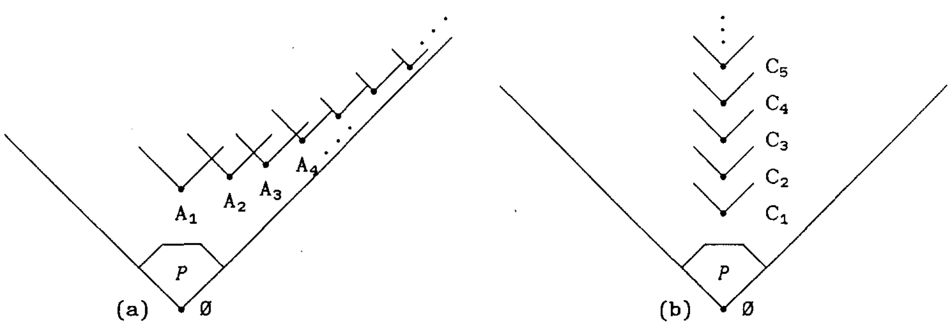

FIG. 3.1. Two sequences $[A_k]_{k=1}^{\infty}$ and $[C_k]_{k=1}^{\infty}$ of recursive languages whose complexities increase without bound. (Languages making an angle steeper than 45° are said to be comparable under $\leqslant_T^p$). Let $\mathscr{A} := \bigcup_{k=1}^{\infty} \{A_k\}^f$, $\mathscr{C} := \bigcup_{k=1}^{\infty} \{C_k\}^f$. The sequence $[A_k]_{k=1}^{\infty}$ diagonalizes over all "upward Turing-reducibility cones" of the form $\{L \in \mathbf{REC} \mid B \leqslant_T^p L\}$ for $B \notin \mathbf{P}$, and the class $\mathscr{A}$ in Theorem 3.6 is not closed in the topology $\mathfrak{R}$. The sequence $[C_k]_{k=1}^{\infty}$ stays within successive cones, and it follows that $\mathscr{C}$ *is* closed in $\mathfrak{R}$.

3.3. *Classifying Complexity Classes Topologically*

First and foremost we appeal to the topological distinction between classes which are closed in $\mathfrak{R}$ and those which are not. With reference to the distinguished basis $\mathfrak{B}_{rr}$ and Theorem 3.3, we can break this down further into four spanning, mutually exclusive categories of cfv classes (other than $\varnothing$):

(1) Recursively presentable classes.

(2) Unbounded r.r. classes.

(3) Classes which are closed in $\mathfrak{R}$, but not in the basis $\mathfrak{B}_{rr}$.

(4) Classes which are not closed in $\mathfrak{R}$.

Before offering examples for each category, we present some results which help in identifying them.

Part (a) of the next result is well known in the form that $\Phi[r]$ is recursively presentable *if* it is scfv. Part (b) parallels Theorem 1 of [LaRo72], except for the latter's use of "recursively presentable" for what we call "r.e.-presentable." (It also virtually follows from Proposition 5.2.2 of [MY78], under appropriate choices of the functions in that result.)

PROPOSITION 3.8. (a) *For any Blum complexity measure Φ and total recursive function $r\colon \Sigma^* \to \mathbb{N}^+$, $\Phi[r]$ has the r.r. property.*

(b) *For any recursively presentable class $\mathscr{C}$ and total recursive function r: $\Sigma^* \to \mathbb{N}^+$, there exists a Blum complexity measure Φ such that $\mathscr{C} = \Phi[r]$.*

This says that among classes which are scfv, the theory of Blum complexity classes is essentially the same as that of r.p. classes.

THEOREM 3.9 [Reg86b]. (a) *For any effective reducibility relation $\leqslant_r$ and language $A \in \mathbf{REC}$, the class $\{L \in \mathbf{REC} \mid A \leqslant_r L\}$ has the r.r. property.*

(b) *If $A \equiv^f B \Rightarrow A \equiv_r B$ for all recursive languages A and B, then $\equiv_r$ is also an effective reducibility relation.*

Proof (sketch). (a) From Definition 2.2 in Section 2.5 there exists a recursive function $\sigma\colon \mathbb{N}^+ \to \mathbb{N}^+$ such that for all j with M_j total, the class of languages $\leqslant_r$-reducible to $L(M_j)$ equals $\mathscr{P}\!\iota[L(M_{\sigma(j)})]$. For all $i, j, c \in \mathbb{N}^+$ define $S_A(i, j, c) :=$ $S_{\mathbf{REC}}(i, j, c) \wedge$ '$A \in \mathscr{P}\!\iota[L(M_{\sigma(j)})]$'. This can be formally expanded into an r.r. predicate for $\{L \in \mathbf{REC} \mid A \leqslant_r L\}$. (Note that using $S_{\mathbf{REC}}$ lets one replace a given M_i accepting a recursive language by a total machine M_j accepting the same language.)

(b) This can be shown using Theorem 3.3(b), wherein the condition $\equiv_r \supseteq \equiv^f$ applies, and the uniform, effective dependence of S_A on A in part (a). ∎

EXAMPLES. $\equiv_T^p$, $\equiv_T^{\log}$, $\equiv_T^{\text{lin}}$ and $\equiv_m^p$, $\equiv_m^{\log}$, $\equiv_m^{\text{lin}}$ are effective reductibility relations. (The latter three need attention to the degenerate cases $\{\varnothing\}$ and $\{\Sigma^*\}$.)

That $\equiv_{\mathrm{iso}}^{p}$ is effective follows from Proposition 2.3 and the fact that the group of p-isomorphisms is recursively presentable [Reg83a]. Left in the middle are $\equiv_1^p$ and $\equiv_i^p$ (i.e., equivalence under polynomial-time 1–1 and invertible reductions; recall that $\equiv_{li}^p$ equals $\equiv_{\mathrm{iso}}^p$). These do not contain $\equiv^f$, and we ask,

Open Question. Are $\equiv_1^p$ and $\equiv_i^p$ effective reductibility relations?

In particular we ask whether $\{L \mid L \equiv_1^p A\}$ and $\{L \mid L \equiv_i^p A\}$ fail to be recursively presentable for some sets $A \in \mathbf{REC}$ which are not *polynomial cylinders.*

The lists of classes which follow are meant to be suggestive rather than exhaustive. Many entries are known in the literature; see e.g., [Sö82]. Others can be verified using Proposition 3.8 and Theorem 3.9, or by direct construction of presentations by appropriate automata (as in the case of **CFL** and **DCFL**). We also appeal to the closure of the family of r.p. classes under recursively presented unions (Lemma 2.2a), e.g., in the case of $\mathbf{PH} = \bigcup_{k=1}^{\infty} \sum_k^p$.

3.3.1. *Recursively Presentable Classes*

(a) **P**, **NP**, **co-NP**, $\sum_k^p$, $\sum_k^{\mathrm{lin}}$ $(k \geqslant 0)$, **PSPACE**, **EXPTIME**, **EXP**, **DLIN**, **NLIN**, **L**, **NL**, **DSPACE**$[n]$, **REG**, **CFL**, **DCFL**, **CSL** $(=\mathbf{NSPACE}[n])$, **DSPACE**$[n]$, **RL**, $\mathbf{NC}_1$, $\mathbf{NC}_2$, ...

(b) **PH**, the boolean hierarchy **BH**, the linear hierarchy **LINH**, **NC**, **ELEM** $(=\bigcup_{k=1}^{\infty} \mathbf{DTIME}[\exp^k(n)])$, the levels of the *Grzegorczyk hierarchy*, ...

(c) $\mathbf{NP} \cap \mathbf{co\text{-}NP}$, $\mathbf{NP} \cup \mathbf{co\text{-}NP}$, **RP**, $\mathbf{RP} \cap \mathbf{co\text{-}RP}$, **BPP**, **PP**, **R-NC**, **AM**, **MA**, ...

(d) The classes of languages which are complete for **NP**, **PP**, **PSPACE**, **EXPTIME**, ... under any of the effective reductibilities $\leqslant_T^p$, $\leqslant_m^p$, $\leqslant_m^{\log}$, $\leqslant_m^{rl}$, ...

In fact, the complete sets for each of the above classes $\mathscr{C}$ form an r.p. class, if $\mathscr{C}$ has any complete sets at all under the given reducibility. **NPI** is also r.p., as is $\{L \mid L \equiv_{\mathrm{iso}}^p A\}$ for any $A \in \mathbf{REC}$, since $\equiv_{\mathrm{iso}}^p$ is effective.

3.3.2. *Unbounded R.R. Classes*

(a) The class of **NP**-hard sets, i.e., $\{L \mid \mathrm{SAT} \leqslant_T^p L\}$, under the implicit restriction $L \in \mathbf{REC}$. Also the classes of **PSPACE**-hard sets, $\sum_k^p$-hard sets $(k \geqslant 0)$, **PP**-hard sets, and so forth. Similarly for $\leqslant_m^p$, $\leqslant_T^{\log}$, $\leqslant_m^{\log}$, etc., building on any class which has a complete member under the reducibility.

(b) **P-SPARSE**, **P**/poly, the class of sets having *exponential density*, ...

(c) **EQ**, i.e., $\{A \in \mathbf{REC} \mid \mathbf{NP}^A = \mathbf{P}^A\}$. ([Háj79] shows that $J_{\mathbf{EQ}} \in \prod_2^0$.)

(d) $\{L \mid L \subseteq A\}$, $\{L \mid L \supseteq A\}$, $\{L \mid L \nsubseteq A\}$, $\{L \mid L \nsupseteq A\}$, where A is a fixed recursive language.

The classes in (d) are not cfv for $A \notin \{\varnothing, \Sigma^*\}$, but their closures under finite variations also have the r.r. property. (Generally, $\mathscr{C}$ r.r. $\Rightarrow \mathscr{C}^f$ is r.r.; see [Reg86b].) A point to notice is that r.r. predicates for these classes depend uniformly and effectively on A. Therefore the class $\{L \oplus A \mid A$ is finite or $L \not\geq A$ $(A, L \in \mathbf{REC})\}$ also has the r.r. property.

The prevalence of r.r. classes among those considered attractive objects of study in the literature likely owes to the fact that these classes have the simplest possible definitions in arithmetic, as shown by Theorem 3.2(b).

3.3.3. *Classes Which Are Closed in* $\mathfrak{R}$

(a) $\bigcap_{\varepsilon>0} \mathbf{DTIME}[n^{1+\varepsilon}]$, $\bigcap_{\varepsilon>0} \mathbf{DTIME}[2^{n^\varepsilon}]$, ...

(b) The class of **PH**-hard sets (with respect to either $\leqslant_T^p$ or $\leqslant_m^p$). This equals $\bigcap_{k=1}^{\infty}\{\Sigma_k^p\text{-hard sets}\}$. Similarly $\{\mathbf{BPP}\text{-hard sets}\}$, $\{\mathbf{RP}\text{-hard sets}\}$, $\{\mathbf{NP} \cap \mathbf{co\text{-}NP}\text{-}$hard sets, ...

(c) **P-IMMUNE** $\cup$ **FIN**. This equals the intersection of $\{L \mid A \not\subseteq L\}$ over all infinite languages $A \in P$. The class of languages which are *bi-immune* to **P** is also closed in $\mathfrak{R}$, since it equals **P-IMMUNE** $\cup$ **FIN** $\cap$ **co-(P-IMMUNE** $\cup$ **FIN)**.

Open Question. Do any of these classes have the r.r. property?

The intersections in (a) and (b) are recursively presented, but the last is not. However, we can rewrite **P-IMMUNE** $\cup$ **FIN** as $\bigcap_{A \in P}\{L \mid \text{if } A \text{ is infinite then } L \not\geq A\}$. By remarks in Section 3.3.2(d) and the fact that **P** is r.r., this is a recursively presented intersection of r.r. classes. We show in Section 6.3 that **P-IMMUNE** itself is neither closed in $\mathfrak{R}$ nor r.e.-presentable. (I am grateful to the anonymous referee for the **P-IMMUNE** $\cup$ **FIN** example, and for a demonstration that the r.r. property is not preserved under recursively presented intersections, which I have expanded as follows:)

PROPOSITION 3.10. *There is a recursively presented sequence* $[\mathscr{C}_k]_{k=1}^{\infty}$ *of r.p. cfv classes whose intersection* $\mathscr{C}_\infty$ *contains* **FIN**, *but is not r.p.*

Thus (d) $\mathscr{C}_\infty$ is unconditionally a class which is closed in $\mathfrak{R}$, but is not in the basis $\mathfrak{B}_{rr}$. We have already shown in Theorem 3.6 that $2^{\aleph_0}$-many such classes exist, but the concrete example may have its own interest.

Proof (sketch). Consider the recursive "universal halting language" $U := \{\langle i, \langle x, n\rangle\rangle \mid \text{Halt}(i, x, n)\}$. For each k, define $\mathscr{C}_k := \{U_i \mid M_i$ fails to halt on at least k different inputs$\}^f$. Using the fact that each $\mathscr{C}_k$ contains **FIN** (consider TMs M_i which halt on no inputs), one can uniformly and effectively obtain a recursive presentation of $\mathscr{C}_k$ in terms of k.

Now consider $\mathscr{C}_\infty := \bigcap_{k=1}^{\infty} \mathscr{C}_k$. Define COFIN $:= \{i \mid M_i$ halts on all but finitely many inputs$\}$. Since $U \in \mathbf{REC}$, there is a recursive function σ: $\mathbb{N}^+ \to \mathbb{N}^+$ such that for all i, $M_{\sigma(i)}$ is total and accepts U_i. Then for all i, $i \in \text{COFIN} \Leftrightarrow M_{\sigma(i)}$ is total and $L(M_{\sigma(i)}) \notin \mathscr{C}_\infty$. Hence $\sigma(\cdot)$ many–one reduces COFIN to $J_{\mathscr{C}_\infty}$. Since COFIN is known to be Σ_3^0-complete [Rog67, p. 328], $J_{\mathscr{C}_\infty} \notin \Pi_2^0$. Thus $\mathscr{C}_\infty$ lacks the r.r. property, so it is not r.p. ∎

Checking that COFIN $\leqslant_m J_{\mathscr{C}_\infty}$ is easier if one skips the closure under finite variations in defining $\mathscr{C}_k$, though this step is needed to show that neither $\mathfrak{B}_{rr}$ nor $\mathbf{B}_{\text{lrp}}$ is closed under recursively presented intersection. We ask whether a similar argument can be used to reduce COFIN to $J_{\mathscr{C}}$ with $\mathscr{C} := \mathbf{P}\text{-}\mathbf{IMMUNE} \cup \mathbf{FIN}$.

3.3.4. *Classes Which Are Not Closed in* $\mathfrak{R}$

L. Landweber, R. Lipton, and E. Robertson [LLR81] showed that $\mathbf{NP}\backslash\mathbf{P}$ is not recursively presentable. Technically this is so even if $\mathbf{NP}\backslash\mathbf{P} = \emptyset$. As we expound upon in Section 5, one can extend their techniques to show that $\mathbf{NP}\backslash\mathbf{P}$ cannot be written as an intersection of lrp cfv classes, unless $\mathbf{NP}\backslash\mathbf{P} = \emptyset$. Then $\mathbf{NP}\backslash\mathbf{P}$ is not closed in $\mathfrak{R}$. (Equivalently, $\mathbf{NP}\backslash\mathbf{P}$ is not closed in $\mathfrak{R}_0$, since it is bounded.) The techniques apply for many other "difference classes" of the form $\mathscr{E} := \mathscr{D}\backslash\mathscr{C}$, where $\mathscr{C}$ is r.p. and $\mathscr{D}$ is closed downward under $\leqslant_m^p$. We shall see in Section 5 that the latter is a sufficient condition for $\mathscr{D}$ to be connected in the topology $\mathfrak{R}$, from which the non-closedness of $\mathscr{E}$ follows by definition (unless $\mathscr{D} = \mathscr{C}$). We have already promised that **REC** is hyperconnected in $\mathfrak{R}$; this suffices to verify the first few cases itemized below.

PROPOSITION 3.11. *The following classes are not closed in* $\mathfrak{R}$:

(a) **REC\FIN**, **REC\P**, **EXPTIME\P**, **EXP\NP**, **PSPACE\LOGSPACE**, **NP\P** (*unless* **NP** = **P**; *ditto* **PH\P**, **NP\NPC**, etc.), **NLIN\DLIN**, **P\NC** (*unless* **P** = **NC**), **CSL\RUD** (i.e., $\mathbf{NSPACE}[n]\backslash\mathbf{LINH}$), **RL\FIN**, ...

(b) $\{L \in \mathbf{REC} \mid L$ *is not* $\mathbf{NP}$*-hard*$\}$, $\{L \in \mathbf{EXP} \mid L$ *is not* $\mathbf{PH}$*-hard*$\}$, *etc.*

(c) *The class of recursive oracles separating* **P** *from* **NP** (*i.e.*, **REC\EQ**).

(d) *The class of recursive languages which are neither p-sparse nor of exponential density.*

(e) *The class of languages in* **EXPTIME** *which are not bi-immune to P.*

(f) $\mathbf{P} \triangle \mathbf{POLYLOGSPACE}$, $\mathbf{P} \triangle \mathbf{CSL}$, $\mathbf{NP} \triangle \mathbf{CSL}$, $\mathbf{NP} \triangle \mathbf{co\text{-}NP}$ (*unless* **NP** = **co-NP**), $\mathbf{P} \triangle \mathbf{DSPACE}[n^r]$ (*for any* $r \geqslant 1$; *cf. results of R. Book cited in* [HU79]), $(\mathbf{PSPACE}\backslash\mathbf{LOGSPACE}) \cup \mathbf{FIN}$....

The following results help relate these examples to other entities in this paper, and will be used in Sections 4 and 6.

THEOREM 3.12 [Reg86b]. *For any classes* $\mathscr{C} \subseteq \mathbf{REC}$ *and* $\mathscr{D} \subseteq \mathbf{RE}$,

(a) *If* $\mathscr{C} \supseteq A^f$ *for some infinite language* A, *then* $I_{\mathscr{C}}$ *is* Σ_3^0*-hard under* $\leqslant_m$.

(b) *If* $\mathscr{C}$ *has the r.r. property, then* $I_{\mathscr{C}} \in \Sigma_3^0$.

(c) *If* $\mathscr{C}$ *is r.e.-presentable then* $I_{\mathscr{C}} \in \Sigma_3^0$.

(d) *If* $I_{\mathscr{C}} \in \Sigma_3^0$ *and* $\mathscr{C} \supseteq \mathbf{FIN}$, *then* $\mathscr{C}$ *is r.e.-presentable.*

(e) *If* $\mathscr{C}$ *has the r.r. property and* $I_{\mathscr{D}} \in \Sigma_3^0$, *then* $I_{\mathscr{D}\backslash\mathscr{C}} \in \Sigma_3^0$.

Proof (sketch). (a) Let A be recursive and infinite. It is possible to construct a recursive function σ: $\mathbb{N}^+ \to \mathbb{N}^+$ such that for all i, $i \in \mathbf{COFIN} \Rightarrow L(M_{\sigma(i)}) \equiv^f A$, and $i \notin \mathbf{COFIN} \Rightarrow L(M_{\sigma(i)})$ is not recursive.

(b) This follows from $J_{\mathscr{C}} \in \prod_2^0$.

(c) Let $U \in \mathbf{RE}$ be universal for $\mathscr{C}$. Then for all i, $L(M_i) \in \mathscr{C} \Leftrightarrow (\exists k)$: $(\forall x)$ [M_i accepts $x \Leftrightarrow x \in U_k$]. The clause in [...] equals [$\exists m R(m) \Leftrightarrow \exists n S(n)$] for some recursive predicates $R(\cdot)$ and $S(\cdot)$, and can be placed into $\prod_2$ prenex form. Hence the whole definition is $\sum_3$. (This part does not need $\mathscr{C} \subseteq \mathbf{REC}$.)

(d) By $I_{\mathscr{C}} \in \sum_3^0$ there exists a recursive predicate $R(\cdot, \cdot, \cdot, \cdot)$ such that for all i, $i \in I_{\mathscr{C}} \Leftrightarrow (\exists a)(\forall b)(\exists c)\, R(i, a, b, c)$. If one then defines

$$U := \{\langle x, \langle i, a\rangle\rangle \mid (\forall b \leqslant x)(\exists c, m)\text{: } R(i, a, b, c) \text{ and Accept}(x, i, m)\}, \tag{3.6}$$

then U is an r.e. universal language for $\mathscr{C}$.

(e) For all i, $i \in I_{\mathscr{D}\setminus\mathscr{C}} \Leftrightarrow i \in I_{\mathscr{D}} \wedge (\exists j)$ [M_j is total $\wedge$ $j \in J_{\mathscr{C}} \wedge L(M_i) = L(M_j)$]. The quantifier schema on recursive predicates for the right-hand side is '$\exists\forall\exists \wedge \exists[\forall\exists \wedge \forall\exists \wedge \forall(\exists \leftrightarrow \exists)]$.' This reduces to '$\exists\forall\exists$' as above. ∎

From (a), (b), (e), and *Myhill's theorem* (see [Rog67], pp. 85–88, 315–316), there is a recursive isomorphism which maps the indices for TMs accepting languages in **P** onto those for TMs accepting languages not in **P**. Similarly, $I_{\mathbf{P}} \equiv_{\text{iso}} I_{\mathbf{REC}\setminus\mathbf{P}} \equiv_{\text{iso}} I_{\mathbf{REC}}$, and $I_{\mathbf{P}} \equiv_{\text{iso}} I_{\mathbf{NP}\setminus\mathbf{P}} \equiv_{\text{iso}} I_{\mathbf{NP}}$ unless **NP** = **P**, etc. Thus it is arithmetically as simple to define membership of $L(M_i)$ in $\mathscr{C}$ when $\mathscr{C}$ is a "difference class" as when $\mathscr{C}$ is r.r. or r.e.-presentable. Moreover, by (a), no cfv classes other than $\varnothing$ and **FIN** have simpler definitions. Thus the r.r. property and the index sets '$J_{\mathscr{C}}$' yield finer distinctions than the sets '$I_{\mathscr{C}}$.'

By (b) and (d), any r.r. class which contains **FIN** is r.e.-presentable. We do not know the answer to the following, which (in tandem with a statement involving "time-uniform" formulas) was incorrectly claimed in [Reg86a].

Open Question. If a class $\mathscr{C} \subseteq \mathbf{REC}$ is r.r. and scfv, then is $\mathscr{C}$ r.e.-presentable?

By way of comparison, the class '$\mathscr{C}_\infty$' in Section 3.3.3 is an example of an scfv class closed in $\mathfrak{R}$ which is not r.e.-presentable. This because $J_{\mathscr{C}_\infty}$, being $\sum_3^0$-hard under $\leqslant_m$, is not in $\prod_3^0$, whereas $I_{\mathscr{C}} \in \sum_3^0 \Rightarrow J_{\mathscr{C}} \in \prod_3^0$ for all $\mathscr{C} \subseteq \mathbf{RE}$. Conversely, $(\mathbf{REC}\setminus\mathbf{P}) \cup \mathbf{FIN}$ is an example of an r.e.-presentable class which is not closed in either $\mathfrak{R}$ or $\mathfrak{R}_0$; ditto $(\mathbf{PSPACE}\setminus\mathbf{LOGSPACE}) \cup \mathbf{FIN}$ and so on. This follows from (e), (d), and the closure of $\sum_3^0$ under finite $\bigcup$.

All these examples show that the topology $\mathfrak{R}$ is a natural object of study. We further this end in the next section by linking $\mathfrak{R}$ in another way to proofs of assertions about algorithms.

4. Provable Properties and $\mathfrak{R}$

Every r.e. language A can be represented by a finite object, namely the index i of a Turing machine M_i accepting A. Thus it is natural (though by no means mandatory) to formalize properties of languages in terms of TMs accepting them, particularly when the formal system $\mathfrak{F}$ allows quantification over integers but not over sets. In the doing so one encounters the problem of determining which TM is the "correct" one to define A by. Given any consistent, r.a. $\mathfrak{F}$ and r.a. language A, one can find TMs M_i and M_j accepting A such that '$L(M_i) = L(M_j)$' is independent of $\mathfrak{F}$ (cf. [HaHo76]). This means that some properties of A are provable under one representation but not the other.

A frequently adopted criterion for proving properties Π of r.e. languages evades this problem in a simple way. It stipulates that Π is *provable* (in $\mathfrak{F}$) if for all languages A having property Π, there exists some TM M_i accepting A such that the sentence '$L(M_i)$ has Π' is provable in $\mathfrak{F}$. [Bak79] uses it in asserting that the property of membership in P is provable in any reasonably strong $\mathfrak{F}$. The reason given is that for every $A \in P$ there exists an M_i accepting A which is *clocked* to run in time n^k for some k. Then $\mathfrak{F}$ proves '$L(M_i) \in P$' by virtue of proving 'M_i runs in polynomial time.'

This criterion is weak; indeed we contend that it is the weakest "sensible" notion of a provable property of languages. We adopt it because its negation gives a strong notion of an *unprovable* property, and we shall be primarily concerned with unprovability results.

To make the criterion more formal, we consider formulas $\phi(\cdot)$ over $\mathfrak{F}$ having one free variable i ranging over $\mathbb{N}^+$; i is intended to refer to the indexing $[M_i]_{i=1}^{\infty}$ of TM acceptors fixed in Section 2.2. Then ϕ is *provable* if for all $i \in \mathbb{N}^+$, $\phi(i) \Rightarrow (\exists d \in \Sigma^*)\ \mathfrak{F}\ \mathrm{Proof}(\phi(i), d)$. Technical remark: we do not require $\mathfrak{F}$ to prove this implication. The formula ϕ is *decidable* if ϕ and $\neg\phi$ are both provable.

Definition 4.1. A class $\mathscr{C} \subseteq \mathbf{RE}$ is *represented by* ϕ if $\mathscr{C} = \{L(M_i) \mid \phi(i)\}$.

Lemma 4.1. *With reference to any sound r.a. formal system $\mathfrak{F}$ which represents T, a property Π of r.e. languages is provable if and only if $\mathscr{C}_\Pi := \{L \mid L \text{ has } \Pi\}$ is representable by a provable formula $\phi(\cdot)$.*

By the remarks about formalizing properties of languages we may presume that Π is represented to begin with by a formula $\psi(\cdot)$ expressing '$L(M_i)$ has Π' over $\mathfrak{F}$. The proof is not quite so simple as defining $\phi(i) := (\exists d)\ \mathfrak{F}\ \mathrm{Proof}(\text{'}\psi(i)\text{'}, d)$ for each i. For this $\phi(\cdot)$ to be provable, $\mathfrak{F}$ must be able to "proof" its own proofs, i.e., $\mathfrak{F}\ \mathrm{Proof}(\text{'}\psi\text{'}, d) \Rightarrow (\exists e)\ \mathfrak{F}\ \mathrm{Proof}(\text{'}\mathfrak{F}\ \mathrm{Proof}(\psi, d)\text{'}, e)$ must hold for all sentences ψ and derivations d. This is the first of the (*Hilbert–Bernays*) *derivability conditions* (see [BJ74]), and meeting it may not be possible for certain very weak formal systems. We make a slight change so that only the "recursive properties" of $\mathfrak{F}$ matter.

Proof. Let $\psi(\cdot)$ represent Π over $\mathfrak{F}$, and let a_0 be the index of a fixed total TM accepting $\{\langle `\eta', d\rangle \mid \mathfrak{F}\,\mathrm{Proof}(`\eta', d)\}$. Define $\phi(\cdot)$ for all i by $\phi(i) := (\exists d, m)\,\mathrm{Accept}(\langle `\psi(i)', d\rangle, a_0, m)$. Then ϕ is a provable formula representing $\mathscr{C}_\Pi$, since $\mathfrak{F}$ proves all true instances of $\mathrm{Accept}(\cdot, \cdot, \cdot)$. ∎

Thus we identify provable properties of r.e. languages with classes $\mathscr{C}$ that are representable by provable formulas of $\phi(\cdot)$ of TM indices. Before proceeding we remark that any such $\phi(\cdot)$ can be effectively replaced by a *decidable* formula $\phi'(\cdot)$ which not only represents $\mathscr{C}$, but also holds for essentially the same TMs. This is because a proof of $\phi(i)$ can be encoded by adding *dummy states* to M_i; the effect is analogous to *Craig's lemma* in logic. For simplicity we continue to refer to "provable" formulas.

We measure the quality of a formula $\phi(\cdot)$ representing a class $\mathscr{C} \subseteq \mathbf{REC}$ by the degree to which $\phi(i)$ holds for "good" algorithms M_i accepting A, for all $A \in \mathscr{C}$. Two conditions reflecting this concern are

DEFINITION 4.2. A formula $\phi(\cdot)$ of TM indices *hits total TMs* if for all i making $\phi(i)$ hold, there exists some j such that $\phi(j)$ holds and $L(M_i) = L(M_j)$. The formula *only holds for total TMs* if for all i, $\phi(i) \Rightarrow M_i$ is total.

As before, when we speak of a provable formula which only holds for total TMs, we do not require $\mathfrak{F}$ to prove the last implication.

LEMMA 4.2. *With reference to any sound r.a. formal system $\mathfrak{F}$ which represents T, and for any nonempty class $\mathscr{C} \subseteq$ **RE**,*

(a) *$\mathscr{C}$ is r.e.-presentable $\Leftrightarrow$ $\mathscr{C}$ is representable by a provable formula.*

(b) *If **FIN** $\subseteq \mathscr{C} \subseteq$ **REC**, then $I_{\mathscr{C}} \in \Sigma_3^0 \Leftrightarrow \mathscr{C}$ is r.e.-presentable $\Leftrightarrow \mathscr{C}$ is representable by a provable formula which hits total TMs.*

(c) *$\mathscr{C}$ is recursively presentable $\Leftrightarrow$ $\mathscr{C}$ is representable by a provable formula which only holds for total TMs.*

Proof. (a) If $\mathscr{C}$ is r.e.-presentable, then as remarked in Section 2.5 there exists a recursive function σ: $\mathbb{N}^+ \to \mathbb{N}^+$ such that $\mathscr{C} = \{L(M_{\sigma(i)}) \mid i \in \mathbb{N}^+\}$. For all j, define $\phi(j) := `j \in \mathrm{Ran}(\sigma)'$, which is formalizable as an instance of Accept for some TM which enumerates the values of σ. Then $\phi(\cdot)$ is a provable formula representing $\mathscr{C}$.

Conversely, given such $\phi(\cdot)$ define $U := \{\langle x, \langle j, d\rangle\rangle \mid [\mathfrak{F}\,\mathrm{Proof}(`\phi(j)', d)$ and $x \in L(M_j)]$ or $[\neg\mathfrak{F}\,\mathrm{Proof}(`\phi(j)', d)$ and $x \in A]\}$, where A is some fixed language in $\mathscr{C}$. By the soundness of $\mathfrak{F}$, U is an r.e. universal language for $\mathscr{C}$.

(b) ($\Rightarrow$) If $I_{\mathscr{C}} \in \Sigma_3^0$, then there exists a recursive predicate $R(\cdot, \cdot, \cdot, \cdot)$ such that $L(M_i) \in \mathscr{C} \Leftrightarrow (\exists a)(\forall b)(\exists c)\, R(i, a, b, c)$ for all i. One can construct a recursive function τ: $\mathbb{N}^+ \times \mathbb{N}^+ \to \mathbb{N}^+$ such that for all $i, a \in \mathbb{N}^+$ and $x \in \Sigma^*$, $M_{\tau(i,a)}$ on input x first searches on successive $b \leqslant x$ for c_b satisfying $R(i, a, b, c_b)$. If c_b is found for all b then $M_{\tau(i,a)}$ simulates M_i on all inputs $y \leqslant x$, and if this process halts $M_{\tau(i,a)}$ finally simulates M_i on input x and accepts iff M_i accepts. So either $L(M_{\tau(i,a)})$ is finite or M_i is total and $(\forall b)(\exists c)\, R(i, a, b, c)$ holds, whence $L(M_i) \in \mathscr{C}$. Then the formula defined for all j by $\phi(j) := `j \in \mathrm{Ran}(\tau)'$ is provable, represents $\mathscr{C}$, and hits total TMs.

Part (a) and Theorem 3.12(c, d) establish the other implications; the $(\Leftarrow)$ parts do not require $\mathscr{C} \supseteq \mathbf{FIN}$.

(c) When $\mathscr{C}$ is r.p., or conversely when $\phi(\cdot)$ is a provable formula which only holds for total TMs, the corresponding construction of (a) yields the desired conclusion. ∎

The upshot is that classes $\mathscr{C}$ which lack the indicated structural properties contain languages whose membership in $\mathscr{C}$ is unprovable in correspondingly strong senses.

THEOREM 4.3. *For all nonempty classes $\mathscr{C} \subseteq$* **RE**, *and with respect to any sound r.a. formal system $\mathfrak{F}$:*

(a) *If $\mathscr{C}$ is not r.e.-presentable, then for any formula $\psi(\cdot)$ representing $\mathscr{C}$, there exists a language $E \in \mathscr{C}$ such that $\psi(i)$ is not provable for* any TM *M_i accepting E.*

(b) *If $\mathscr{C}$ is not representable by a provable formula which hits total TMs, then for any $\psi(\cdot)$ representing $\mathscr{C}$, there exists a language $E \in \mathscr{C}$ such that $\psi(i)$ is not provable for any* total *TM M_i accepting E.*

(c) *If $\mathscr{C}$ is not recursively presentable, then for any formula $\psi(\cdot)$ representing $\mathscr{C}$, there exists a language $E \in \mathscr{C}$ such that $\psi(i)$ is not provable for any* provably *total TM M_i accepting E.*

Proof. (a) Suppose not, i.e., $(\forall E \in \mathscr{C})(\exists i, d)[E = L(M_i) \wedge \mathfrak{F}\ \mathrm{Proof}(\psi(i), d)]$. Let M_{a_0} be a total TM computing $\mathfrak{F}$ Proof$(\cdot, \cdot)$ as in the proof of Lemma 4.1. Then for all i define $\phi(i) := (\exists d, m)\, \mathrm{Accept}(\langle `\psi(i)', d\rangle, a_0, m)$. By our supposition and the soundness of $\mathfrak{F}$, ϕ is a provable formula and represents $\mathscr{C}$. By Lemma 4.2(a), this makes $\mathscr{C}$ r.e.-presentable, yielding a contradiction.

(b) Let $\psi(\cdot)$ be the formula in (a). If for all $E \in \mathscr{C}$ there were some total M_i accepting E such that $\mathfrak{F} \vdash \psi(i)$, then the derived formula ϕ would be a provable formula representing $\mathscr{C}$ which hits total TMs.

(c) With a_0 as in (a), define $\phi(i) := (\exists d, e, m, n)[\mathrm{Accept}(\langle `\psi(i)', d\rangle, a_0, m) \wedge \mathrm{Accept}(\langle `M_i \text{ is total}', e\rangle, a_0, n)]$ for all i instead. Since $\mathfrak{F}$ is sound, the contrary supposition about ψ implies that the provable formula $\phi(\cdot)$ represents $\mathscr{C}$ and only holds for total TMs, contradicting Lemma 4.2(c). ∎

Let $\mathfrak{F}$ and some "standard" definition $\psi(\cdot)$ of $\mathscr{C}$ over $\mathfrak{F}$ be fixed. Then we conclude that the unprovability of '$E \in \mathscr{C}$' does not depend on the TM *name* chosen for E, within the provisos of each case. So it must be traced to the structure of the language E itself.

Part (c) is technically meaningful only for classes of languages accepted by provably total machines, which by Lemma 4.2(c) are automatically bounded and hence contained in **DTIME**$[u]$ for some recursive function u dependent only on $\mathfrak{F}$. (See also [Gor79].) For larger classes we can show,

THEOREM 4.4. *Suppose $\mathscr{C} \subseteq$ **REC** is scfv and not recursively presentable. Then for any sound r.a. formal system $\mathfrak{F}$, formula $\psi(\cdot)$ representing $\mathscr{C}$ over $\mathfrak{F}$, and recursive function u: $\mathbb{N} \to \mathbb{N}^+$, there exists a laguage $E \in \mathscr{C}$ such that for all TMs M_i accepting E which run in time $u(n)$, $\mathfrak{F} \nvdash \psi(i)$.*

Proof. Let $\mathscr{C} \supseteq A^f$ for some $A \in$ **REC**. For all $i, d \in \mathbb{N}^+$ and $x \in \Sigma^*$ let $M_{\tau(i,d)}$ on input x first test whether $\mathfrak{F}$ Proof('$\psi(i)$', d) holds. If the test fails, then $M_{\tau(i,d)}$ accepts x iff $x \in A$. If it passes, $M_{\tau(i,d)}$ then tests whether M_i on input y halts within $u(|y|)$ steps, for all $y \leqslant x$. If the second test fails for some y, $M_{\tau(i,d)}$ again accepts x iff $x \in A$. Otherwise $M_{\tau(i,d)}$ accepts x iff M_i accepts x. The function τ can be computed recursively.

For all i and d, if $\mathfrak{F}$ Proof('$\psi(i)$', d) and M_i runs in time $u(n)$, then $L(M_{\tau(i,d)}) = L(M_i)$. If $\mathfrak{F}$ Proof('$\psi(i)$', d) fails then $L(M_i) = A$, while if it holds but M_i fails to run in time $u(n)$, then $L(M_i) \equiv^f A$. In all cases $L(M_{\tau(i,d)}) \in \mathscr{C}$. If one denies the conclusion in the statement of this theorem then one obtains $\mathscr{C} = \{L(M_{\tau(i,d)}) \mid i, d \in \mathbb{N}^+\}$. Since $M_{\tau(i,d)}$ is total for all i and d, this gives the contradiction that $\mathscr{C}$ is recursively presentable. ▌

We do not know whether the condition that $\mathscr{C}$ be scfv is needed here. We give in Section 6.2 an example of a nonempty bounded cfv class which is not r.p. but is representable by a provable formula that hits total TMs. Therefore the conclusion of Theorem 4.3(b) is not simply a "limiting case" of Theorem 4.4 as the recursive function u: $\mathbb{N} \to \mathbb{N}^+$ increases without bound.

The upshot is that if $\mathscr{C}$ is not closed in $\mathfrak{R}$, then membership in $\mathscr{C}$ is not a provable property of languages; nor is it even an infinite conjunction of provable properties. The next section gives us a method for identifying natural classes which are not closed in $\mathfrak{R}$.

5. UNIFORM DIAGONALIZATION THEOREMS

U. Schöning [Sö81, Sö82] and P. Chew and M. Machtey [CM81] found general means of applying the technique of *delayed diagonalization*, whose use in complexity theory goes back to [Lad75, BCH70]. The main theorem of this section sharpens their results through its use of real-time/log-space reductions, and brings out the symmetry between the classes $\{\mathscr{C}_k\}$ and the languages $\{A_k\}$. Since **FLIN**, **FL**, and **FP** all contain $\mathbf{RL_F}$, its statement also holds with $\leqslant_{li}^{\text{lin}}$, $\leqslant_{li}^{\log}$, $\leqslant_{li}^{p}$, and *a fortiori* $\leqslant_m^{\text{lin}}$, $\leqslant_m^{\log}$, and the $\leqslant_m^p$ of [Sö82], in place of $\leqslant_{li}^{rl}$. The main theorem of [Sö82] has $m = 2$, also assumes $A_1 \in \mathbf{P}$, $A_2 \notin \{\varnothing, \Sigma^*\}$, and concludes $E \leqslant_m^p A_2$, as follows directly from our statement.

THEOREM 5.1. (a) *Let* $\mathscr{C}_1, \ldots, \mathscr{C}_m$ *be r.p. c.f.v. classes, and let* $A_1, \ldots, A_m$ *be recursive languages such that for each k,* $A_k \notin \mathscr{C}_k$. *Then there exists a recursive language E such that* $E \notin \mathscr{C}_1 \cup \cdots \cup \mathscr{C}_m$, *and yet* $E \leqslant_{li}^{rl} A_1 \oplus \cdots \oplus A_m$.

(b) *The conclusion of* (a) *holds even when some* $\mathscr{C}_k$ *are only r.e.-representable instead of r.p., so long as the corresponding languages* A_k *are finite. If* A_k *is the only nonempty language in the list* $A_1, \ldots, A_m$, *then also* $E \subseteq A_k$.

The theorem is in fact constructive; i.e., there is an effective procedure which, given indices for total TMs accepting the languages $A_1, \ldots, A_m$ and universal languages for the classes $\mathscr{C}_1, \ldots, \mathscr{C}_m$, produces the index of a total TM accepting E so long as the hypotheses hold. With reference to (b) the procedure works even if the TMs for the indicated classes $\mathscr{C}_k$ are not total.

EXAMPLE ("Ladner's theorem" [Lad75]). With reference to Fig. 5.1, take

$$\mathscr{C}_1 := \mathbf{NPC}, \qquad A_1 := \varnothing$$
$$\mathscr{C}_2 := \mathbf{P}, \qquad A_2 := \mathrm{SAT}.$$

Assuming $\mathbf{NP} \neq \mathbf{P}$, these choices satisfy the hypotheses of Theorem 5.1 with $m = 2$. The conclusion yields a language E such that $E \notin \mathbf{P} \cup \mathbf{NPC}$ and yet $E \leqslant_m^{rl} \mathrm{SAT} \oplus \varnothing$, so $E \in \mathbf{NP}$. (Also by (b), $E \subseteq \mathrm{SAT}$.) Hence there are languages in $\mathbf{NP}$ which are neither $\mathbf{NP}$-complete nor in $\mathbf{P}$.

[Sö82] gives a similar example with

$$\mathscr{C}_1 := \{L \subseteq \Sigma^* \mid L \equiv_T^p \mathrm{QBF}\} \qquad A_1 := \varnothing,$$
$$\mathscr{C}_2 := \mathbf{PH} \qquad A_2 := \mathrm{QBF},$$

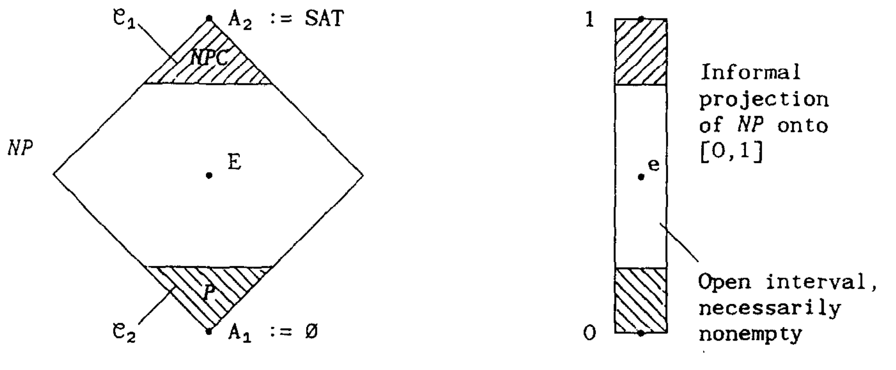

FIG. 5.1. Illustration of Ladner's theorem. *Assumption*: $\mathbf{NP} \neq \mathbf{P}$. *Conclusion*: For some language $E \in \mathbf{NP}$, $A_1 <_m^p E <_m^p A_2$ (so $E \notin \mathbf{P}$, $E \notin \mathbf{NPC}$). By the assumption, $\mathscr{C}_1 := \mathbf{NPC}$ and $\mathscr{C}_2 := \mathbf{P}$ are disjoint proper subsets of $\mathbf{NP}$. Since both are r.p. cfv, and classes such as $\mathbf{NP}$ cannot be written as the disjoint union of r.p. cfv classes, we conclude that $\mathbf{NP} \setminus (\mathbf{NPC} \cup \mathbf{P})$ is nonempty. An analogy with the usual topology on the reals is sketched at right.

where QBF is the **PSPACE**-complete language of satisfiable Boolean formulas. The conclusion is that if **PSPACE** $\neq$ **PH**, then there are languages in **PSPACE** which are neither **PSPACE**-complete (under $\leqslant_T^p$) nor in the polynomial hierarchy.

Here, if one also supposes **NP** $\neq$ **P** then one can add (Fig. 5.2):

$$\mathscr{C}_3 := \{L \in \mathbf{PSPACE} \mid \mathrm{SAT} \leqslant_T^p L\} \qquad A_3 := \varnothing.$$

Then one obtains a language $E \in$ **PSPACE** which is not in the polynomial hierarchy and not **NP**-hard either. That $\mathscr{C}_1$, $\mathscr{C}_2$, and $\mathscr{C}_3$ are r.p. is well known, and follows from the classification in Section 3.3.

In Fig. 5.1 we have diagramed **P** and NPC as closed regions in the plane, and formally think of **NP**, which equals $\{L \mid \varnothing \leqslant_m^p L \leqslant_m^p \mathrm{SAT}\}$, as a connected "reducibility interval" projected into $\mathbb{R}^2$. A further projection onto the real line reflects the fact that an interval of $\mathbb{R}$ cannot be decomposed into disjoint closed sets. In Fig. 5.2, however, the "closed sets" $\mathscr{C}_2$ and $\mathscr{C}_3$ are not disjoint, and the analogy with $\mathbb{R}$ fails. To obtain conclusions similar to that of Fig. 5.1 in these cases, we shall appeal to the fact that reducibility intervals are actually hyperconnected in the underlying topology, namely $\mathfrak{R}$.

5.1. *Outline of the proof of Theorem* 5.1

Our proof of Theorem 5.1, which also extends it to the case of infinitely many classes $\mathscr{C}_k$ and languages A_k, refines methods which have previously appeared in the literature. Full details may be found in [Reg86b] and will form the subject of another paper. We give the lemmas which outline the proof, lending them more general forms which may bear independent interest. In place of $\{1, ..., m\}$ we use an

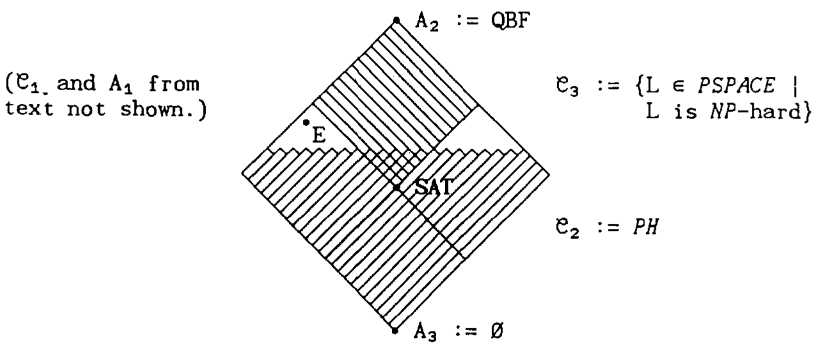

FIG. 5.2. An example where r.p. cfv classes overlap. *Assumptions*: **PSPACE** $\neq$ **PH**, **NP** $\neq$ **P**. *Conclusion*: There exists $E \in$ **PSPACE** such that E is neither **NP**-hard nor in the polynomial hierarchy. Note that the projection onto $\mathbb{R}$ used in Fig. 5.1 fails. Whereas any closed interval in $\mathbb{R}$ that is not a single point equals some nontrivial union of closed subintervals, complexity classes such as **PSPACE**, **NP**, and **P** cannot be written as a nontrivial union of r.p. cfv classes. The topological property this reflects, namely *hyperconnectedness*, does not crop up in the usual topology on $\mathbb{R}^n$.

abstract set K to index the classes $\mathscr{C}_k$ and A_k, technically stipulating that $1 \in K$ to make domains of the form $\{k \in K \mid k \leqslant x\}$ in the statements below nonempty even when $x = \lambda$. We also only care that the functions 'f' and 'h' below are recursive; closer analysis of their growth rates in specific instances may provide the power needed for the project relating to [Jos83] and [PaHa77] that was proposed in the Introduction.

In the first two definitions and the first lemma, none of the entities defined is assumed to be recursive. Note: $(U_k)_l$ equal $\{z \mid \langle\langle z, l\rangle, k\rangle \in U\}$.

DEFINITION 5.1. Let $[U_k]_{k \in K}$ and $[A_k]_{k \in K}$ be sequences of languages, where $1 \in K \subseteq \mathbb{N}^+$. Define $\mathscr{C}_k := \mathscr{P}\!\ell[U_k]$ for each $k \in K$. Call $f\colon \Sigma^* \to \Sigma^*$ a *witness-ranging function for* $[A_k]_{k \in K}$ *with respect to* $[\mathscr{C}_k]_{k \in K}$ if for all $x \in \Sigma^*$ and $k, l \leqslant x$ ($k \in K$, $l \in \mathbb{N}^+$), there exists $y \in A_k \triangle (U_k)_l$ such that $x < y \leqslant f(x)$.

DEFINITION 5.2. Given a strictly increasing function $f\colon \Sigma^* \to \Sigma^*$ and a function h defined on Σ^*, say h *out-runs* f if for each $k \in \operatorname{Ran}(h)$, there are infinitely many x such that h takes the constant value k on the half-open interval $(x, f(x)]$. Say a language L *out-runs* f iff χ_L outruns f.

When $\operatorname{Ran}(h)$ has cardinality at least 2 we can rephrase this by saying that for each $k \in \operatorname{Ran}(h)$, the function h_k mapping any string $x \in \Sigma^*$ to the least $z > x$ such that $h(z) \neq k$ is infinitely often greater than f. This accounts for our saying h "out-runs" f. The first lemma is a noneffective result which isolates the combinatorial twist in the diagonalization mechanism.

LEMMA 5.2. *Let* $[U_k]_{k \in K}$ *and* $[A_k]_{k \in K}$ *be collections of languages indexed by the set* $K \subseteq \mathbb{N}^+$. *Suppose* $f\colon \Sigma^* \to \Sigma^*$ *is a witness-ranging function for* $[A_k]_{k \in K}$ *with respect to* $[\mathscr{C}_k]_{k \in K}$, *where* $\mathscr{C}_k := \mathscr{P}\!\ell[U_k]$ *for all k. Then for any function h from* Σ^* *onto K which out-runs f, the language*

$$E := \bigcup_{k \in K} (A_k \cap h^{-1}(k)) \tag{5.1}$$

is not in $\mathscr{C}_k$ *for any* $k \in K$.

The next lemma works toward showing that when we work with recursive presentations of classes $\mathscr{C}_k$ and languages A_k, the construction of E in Lemma 5.2 is effective. We introduce J in (b) to address Theorem 5.1(b), whereby J is considered to be a subset of $\{1, ..., m\}$.

LEMMA 5.3. (a) *Let* $K \subseteq \mathbb{N}^+$ *be recursive, and let U and A be recursive languages such that for all* $k \in K$, $A_k^{\ell} \cap \mathscr{C}_k = \varnothing$, *where* $\mathscr{C}_k := \mathscr{P}\!\ell[U_k]$ *as before. Then there is a recursive witness-ranging function f for* $[A_k]_{k \in K}$ *with respect to* $[\mathscr{C}_k]_{k \in K}$. *One can compute f uniformly and effectively from J, K, U, and A.*

(b) *The same is true if U is r.e. so long as for some recursive set* $J \subseteq K$ *and all* $k \in K$: $k \in J \Rightarrow A_k$ *is finite, and* $k \notin J \Rightarrow U_k$ *is recursive.*

The languages $\{A_k\}$ and the function h from Σ^* onto K used in Lemma 5.2 place a bound on the complexity of E, because E reduces to the language $A_K := \{xk \mid k \in K, x \in A_k\}$ via the mapping $x \mapsto xh(x)$. In the case of Theorem 5.1 where $K = \{1, ..., m\}$, E reduces to the language $A_1 \oplus \cdots \oplus A_m$ via much the same reduction, up to details of finite coding which we leave to the reader. The nub is that if $h \in \mathbf{RL}_\mathbf{F}$, then the reduction is also a function in $\mathbf{RL}_\mathbf{F}$ since $\mathrm{Ran}(h)$ is finite. The result we need is thus:

LEMMA 5.4. *For any total recursive function $f\colon \Sigma^* \to \Sigma^*$, and any $m \geqslant 2$, one can obtain a function $h\colon \Sigma^* \to \{1, ..., m\}$ which outruns f, where $h \in \mathbf{RL}_\mathbf{F}$.*

For stronger results, the former for the case $m = 2$ and the latter for $m = \infty$, see [Sdt85] and [Reg86b]. The phenomenon of unprovability in our applications in Section 6 arises from the ability to find an $h \in \mathbf{RL}_\mathbf{F}$ which out-runs every function that $\mathfrak{F}$ can prove to be recursive.

Proof of Theorem 5.1 (sketch). Let $U_1, ..., U_m$ be recursive universal languages for $\mathscr{C}_1, ..., \mathscr{C}_m$. Taking $K := \{1, ..., m\}$ in Lemma 5.3 yields a recursive witness-ranging function f for $A_1, ..., A_m$ with respect to $\mathscr{C}_1, ..., \mathscr{C}_m$. Lemma 5.4 finds a function $h \in \mathbf{RL}_\mathbf{F}$ which out-runs f. Define $E := \bigcup_{k=1}^m (A_k \cap h^{-1}(k))$. By Lemma 5.2, $E \notin \mathscr{C}_1 \cup \cdots \cup \mathscr{C}_m$. To finish the proof one need only check that the available 1–1 length-increasing reduction from E to $A_1 \oplus \cdots \oplus A_m$, which essentialy has the form $x \mapsto xh(x)$, can be inverted as well as computed in real time and log space. ∎

5.2. *Interpreting Theorem* 5.1 *as a Statement about Connectedness*

For convenience we restate Theorem 5.1 under a slight but suggestive weakening of the hypotheses. The fact that we are dealing with only finitely many classes $\mathscr{C}_k$ and languages A_k comes into the proof, which is no longer effective. We call a class $\mathscr{C}$ "locally r.e.-representable" if $\mathscr{C} \cap \mathscr{D}$ is r.e.-presentable for all r.p. cfv classes $\mathscr{D}$ such that $\mathscr{C} \cap \mathscr{D}$ is nonempty.

THEOREM 5.5. (a) *Let $\mathscr{C}_1, ..., \mathscr{C}_m$ be arbitrary intersections of lrp cfv classes, and let $A_1, ..., A_m$ be recursive languages such that for each k, $A_k \notin \mathscr{C}_k$. Then there exists a recursive language E such that $E \notin \mathscr{C}_1 \cup \cdots \cup \mathscr{C}_m$ and yet $E \leqslant_{li}^{rl} A_1 \oplus \cdots \oplus A_m$.*

(b) *The conclusion of* (a) *holds even when some of the classes $\mathscr{C}_k$ are intersections of locally r.e.-presentable cfv classes, so long as the corresponding languages A_k are finite. If A_k is the only nonempty language in the list $A_1, ..., A_m$, then also $E \subseteq A_k$.*

Proof. (a) For each $k \in \{1, ..., m\}$, there exists some lrp cfv class $\mathscr{C}'_k$ such that $\mathscr{C}'_k \supseteq \mathscr{C}_k$ and $A_k \notin \mathscr{C}_k$. Define $\mathscr{C}''_k := \mathscr{C}'_k \cap \{L \mid L \leqslant_T^p A_1 \oplus \cdots \oplus A_m\}$. Since $\leqslant_T^p$ (which is chosen somewhat arbitrarily) is an effective reducibility relation, the class $\{L \mid L \leqslant_T^p A_1 \oplus \cdots \oplus A_m\}$ is r.p. It is also cfv. Thus $\mathscr{C}''_k$ is cfv, and either r.p. or empty. Applying Theorem 5.1(a) for the classes $\mathscr{C}''_1, ..., \mathscr{C}''_m$ (disregarding ones that are empty, if any) then yields E such that $E \leqslant_{li}^{rl} A_1 \oplus \cdots \oplus A_m$ and $E \notin \bigcup_{k=1}^m \mathscr{C}''_k$.

Since $\leqslant_T^p \supseteq \leqslant_{li}^{rl}$, we have $E \notin \bigcup_{k=1}^m \mathscr{C}'_k$, and hence $E \notin \bigcup_{k=1}^m \mathscr{C}_m$. Part (b) is similar to (a). ∎

The condition on the classes $\mathscr{C}_1, ..., \mathscr{C}_m$ in (a) above is precisely that they be closed in the topology (**REC**, $\mathfrak{R}_0$). With $m := 2$ the conditions $A_1 \notin \mathscr{C}_1$, $A_2 \notin \mathscr{C}_2$ recall the definition of a T_2-space in Section 2.7. However, the clause $E \notin \mathscr{C}_1 \cup \mathscr{C}_2$ implies that $\mathscr{S} = \mathscr{C}_1 \cup \mathscr{C}_2$ and $E \in \mathscr{S}$ cannot both hold. When $\mathscr{S}$ is closed downward under $\leqslant_{li}^{rl}$ and $\oplus$ (and $A_1, A_2 \in \mathscr{S}$), $E \in \mathscr{S}$ holds, and so $\mathscr{S} \nsubseteq \mathscr{C}_1 \cup \mathscr{C}_2$.

Now Proposition 2.4(a) in Section 2.7 tells us that a class $\mathscr{D}$ is hyperconnected in $\mathfrak{R}_0$ iff there are no classes $\mathscr{C}_1, \mathscr{C}_2$ closed in $\mathfrak{R}_0$ such that $\mathscr{D} \subseteq \mathscr{C}_1 \cup \mathscr{C}_2$ while $\mathscr{D} \nsubseteq \mathscr{C}_1$ and $\mathscr{D} \nsubseteq \mathscr{C}_2$. $\mathscr{D}$ is connected in $\mathfrak{R}_0$ iff there are no classes $\mathscr{C}_1, \mathscr{C}_2$ closed in $\mathfrak{R}_0$ such that $\mathscr{D}$ equals the *disjoint* union of $\mathscr{C}_1 \cap \mathscr{D}$ and $\mathscr{C}_2 \cap \mathscr{D}$. The last two statements hold similarly for $\mathfrak{R}$.

The following definition and result are motivated by the form of E in Lemma 5.2, and by the fact that all convex subsets in $\mathbb{R}^n$ are connected in the usual topology. (A subset $D \subseteq \mathbb{R}^n$ is *convex* if for all $a, b \in D$ and $l \in [0, 1]$, $al + b(1-l) \in D$.)

DEFINITION 5.3. For any languages A, B, and L, the *splice* of A, B by L is defined to be $(A \cap L) \cup (B \cap \tilde{L})$. A class $\mathscr{D}$ is *closed under splices by* sets in another class $\mathscr{L}$ if for all $A, B \in \mathscr{D}$ and $L \in \mathscr{L}$, $(A \cap L) \cup (B \cap \tilde{L}) \in \mathscr{D}$.

If $L \in$ **RL**, then L essentially adds nothing to the time/space complexity of the language $(A \cap L) \cup (B \cap \tilde{L})$. Thus we think of classes which are closed under such "easy" splices as convex in terms of complexity.

THEOREM 5.6. *Let $\mathscr{D}$ be any cfv class of recursive languages.*

(a) *If $\mathscr{D}$ is closed under splices by sets in* **RL** *then $\mathscr{D}$ is hyperconnected in $\mathfrak{R}_0$, and hence hyperconnected in $\mathfrak{R}$.*

(b) *If $\mathscr{D}$ is closed under $\bigcap$ with sets in* **RL**, *then $\mathscr{D}$ is connected in $\mathfrak{R}_0$ (and $\mathfrak{R}$).*

This makes hyperconnectedness in $\mathfrak{R}$ analogous to connectedness (of convex sets) in $\mathbb{R}^n$. The splicing condition does not characterize hyperconnected classes in either $\mathfrak{R}$ or $\mathfrak{R}_0$, though we suspect it can be modified to do so for $\mathfrak{R}$.

Proof. (a) If $\mathscr{D}$ is not hyperconnected in $\mathfrak{R}_0$, then there are classes $\mathscr{C}_1, \mathscr{C}_2$ closed in $\mathfrak{R}_0$ whose union contains $\mathscr{D}$, and languages $A_1 \in \mathscr{D} \backslash \mathscr{C}_1$, $A_2 \in \mathscr{D} \backslash \mathscr{C}_2$. Then Theorem 5.5(a) yields a set E of the form $(A_1 \cap h^{-1}(1)) \cup (A_2 \cap h^{-1}(2))$ for some $h \in \mathbf{RL_F}$, where $h: \Sigma^* \to \{1, 2\}$ and $E \notin \mathscr{C}_1 \cup \mathscr{C}_2$. Since $h^{-1}(2) = {\sim} h^{-1}(1)$, E equals the splice of A_1, A_2 by $h^{-1}(1)$. By Lemma 2.1, $h^{-1}(1) \in$ **RL**, and the splicing condition then gives $E \in \mathscr{D}$, contradicting $\mathscr{D} \subseteq \mathscr{C}_1 \cup \mathscr{C}_2$.

(b) By the hypothesis on $\mathscr{D}$, $\mathscr{D}$ contains $\varnothing$. If $\mathscr{D}$ is not connected in $\mathfrak{R}_0$, then $\mathscr{D}$ can be partitioned into classes $\mathscr{C}_1, \mathscr{C}_2$ closed in $\mathfrak{R}_0$ such that w.l.o.g. $\varnothing \in \mathscr{C}_2$. Take $A_1 \in \mathscr{D} \backslash \mathscr{C}_1$ and $A_2 := \varnothing$; then for some $L \in$ **RL**, Theorem 5.5(a) gives the language $E := (A_1 \cap L) \cup (\varnothing \cap \tilde{L}) = A_1 \cap L$ such that $E \notin \mathscr{C}_1 \cup \mathscr{C}_2$. This yields a similar contradiction. Proposition 2.4(d) makes (a, b) hold for $\mathfrak{R}$. ∎

An immediate corollary is Proposition 3.1, namely that **REC** is hyperconnected in $\mathfrak{R}_0$. So are, **P**, **NP**, **NP** $\cap$ **co-NP**, **RP**, **BPP**, **PP**, **PH**, **PSPACE**, **L**, **NL**, **DLIN**, **NLIN**, **EXPTIME**, **EXP**, and sundry other classes. **NP** $\cup$ **co-NP** is connected in $\mathfrak{R}_0$, and hyperconnected in $\mathfrak{R}_0$ iff **NP** = **co-NP**.

Although the conclusions are weaker, we state similar results in terms of $\mathfrak{R}$ rather than $\mathfrak{R}_0$ in what follows, for convenience in referring to Sections 3 and 4. Since all applications in Section 6 involve bounded classes, there is no difference.

THEOREM 5.7. *Let $\mathscr{D}$ be closed under $\bigcap$ with sets in* **RL**, *and let $\mathscr{C}$ be a class closed in $\mathfrak{R}$ which does not contain $\mathscr{D}$. Let $\mathscr{E}$ be any r.e.-representable class such that $\mathscr{E} \cap$* **FIN** *$= \emptyset$. Then $\mathscr{E}$ does not contain $\mathscr{D} \backslash \mathscr{C}$.*

We do not know whether this holds under the weaker stipulation that $\mathscr{D}$ be connected in $\mathfrak{R}$, or even if $\mathscr{D}$ is hyperconnected in $\mathfrak{R}_0$. There is also a statement paralleling Theorem 5.6(a) and Theorem 5.5(b) when $\mathscr{D}$ is closed under splices by sets in **RL**, whose formulation we leave to the interested reader. All of this points up the distinguished role of **FIN**.

Proof. Suppose not; then Theorem 5.5(b) applies with $\mathscr{C}_1 := \mathscr{E}^f$, $A_1 := \emptyset$, $\mathscr{C}_2 := \mathscr{C}$, and $A_2 \in \mathscr{D} \backslash \mathscr{C}$. This yields E such that $E \notin \mathscr{C}_1 \cup \mathscr{C}_2$ and $E = A_2 \cap L$ for some $L \in$ **RL**. However, the condition on $\mathscr{D}$ gives $E \in \mathscr{D}$, a contradiction. ∎

EXAMPLE. **NP** is closed under $\bigcap$ with sets in **P**, and **P** is closed in $\mathfrak{R}$. *Now assume* **NP** $\neq$ **P**. By the definition of connectedness for **NP**, **NP**$\backslash$**P** is not closed in $\mathscr{R}$, so it is not r.p. [LLR81]. In particular, **NP**$\backslash$**P** is not equal to any r.p. class such as **NPC**, giving Ladner's theorem again. Since **P** $\supseteq$ **FIN**, Theorem 5.7 shows that **NP**$\backslash$**P** is not r.e.-presentable either [CM81]. Using Theorem 4.3(a), we conclude that membership in **NP**$\backslash$**P** is an unprovable property of languages, in a very strong sense which we elaborate in Section 6.1.

The next proposition gives sufficient conditions for (hyper)connectedness in $\mathfrak{R}$ in the more familiar terms of downward closure under reducibilities. Similar statements hold for any effective reducibility $\leqslant_r \supseteq \leqslant_{li}^{rl}$ and its associated *zero-degree*, such as $\leqslant_{li}^{p}$ and **P**.

PROPOSITION 5.8. (a) *If $\mathscr{D}$ is closed under $\oplus$ and downward under $\leqslant_{li}^{rl}$, then $\mathscr{D}$ is closed under splices by sets in* **RL**, *and thus hyperconnected in $\mathfrak{R}$.*

(b) *Let $\mathscr{D}$ be closed under joins with $\emptyset$ and downward under $\leqslant_{li}^{rl}$. Alternatively, let $\mathscr{D}$ be closed downward under $\leqslant_m^{\text{lin}}$ and contain some set other than $\emptyset$ or Σ^*. Then $\mathscr{D}$ is closed under $\bigcap$ with sets in* **RL**, *and thus connected in $\mathfrak{R}$.*

Proof. (a) Given $A, B \in \mathscr{D}$ and $L \in$ **RL**, take $E := (A \cap L) \cup (B \cap \tilde{L})$. Then $E \leqslant_{rl}^{li} A \oplus B$ by the map $x \mapsto xb$, where b is '0' if $x \in L$ and '1' otherwise. Thus $E \in \mathscr{D}$.

(b) Given $A, B \in \mathscr{D}$ and $L \in$ **RL**, and taking $E := A \cap L$, we have $E \leqslant_{li}^{rl} A \oplus \emptyset$ as before. The first condition on $\mathscr{D}$ then implies $E \in \mathscr{D}$. Under the second, we observe that $A \oplus \emptyset \leqslant_m^{\text{lin}} A$ for all $A \neq \Sigma^*$. In the case $A = \Sigma^*$ we have $E = L$, and then $E \in \mathscr{D}$ follows because the second condition implies $\mathbf{RL} \subseteq \mathscr{D}$. ∎

Some Remarks. (1) We do not know whether (b) holds with '$\leqslant_m^{rl}$' in place of '$\leqslant_m^{\text{lin}}$'. The problem is that $A \cap L \leqslant_m^{rl} A$ (or alternatively $A \oplus \emptyset \leqslant_m^{rl} A$) may fail because the "knowledge" that the input x should not be copied onto the output tape can come too late for a real-time reduction. This also highlights the effects of our defining '$\oplus$' with the "decision bit" on the right, and may reflect an anomaly of the usual model of a multitape TM transducer.

(2) There are classes which meet the sufficient conditions for hperconnectedness in $\mathfrak{R}$ given in Theorem 5.7 and Proposition 5.8, but do not meet the corresponding conditions for connectedness. An example is $\{A \oplus L \mid L \in \mathbf{P}\}$, for any recursive language $A \neq \emptyset$. We find the given criteria easy to apply, however, and ask openly for weaker sufficient conditions on connectedness.

(3) D. Schmidt [Sdt85] gives results analogous to Theorems 5.6 and 5.7, saying that classes $\mathscr{D}$ meeting certain conditions cannot be written as (a) disjoint or (b) nontrivial unions of r.p. cfv classes. These conditions are that $\mathscr{D}$ be *recursive gap closed* (somewhat stronger than saying every recursive function can be out-run by some $D \in \mathscr{D}$), and for (a) that $\mathscr{D}$ be closed under finite $\bigcap$, while for (b) that $\mathscr{D}$ be closed under both finite $\bigcap$ and finite $\bigcup$. However, **PP** is an example of a natural class which is not known to be closed under finite $\bigcup$ or finite $\bigcap$, but is closed under splices by sets in **P**, and is thus hyperconnected in $\mathfrak{R}$. In fact, **PP** is closed under $\bigcap$ *with* sets in **P** and under *disjoint* $\bigcup$. (To be fair, the italicized modifiers could be added to the results in [Sdt85] as they stand.)

We conclude by giving several more examples of hyperconnected classes. **P-SPARSE** and **P/poly** respectively denote the classes of languages which are polynomially *sparse* or which have *small circuits.*

PROPOSITION 5.9. *The following classes are all hyperconnected in* $\mathfrak{R}$:

(a) **P-SPARSE**

(b) **P/poly**,

(c) **P-IMMUNE** $\cup$ **FIN**

(d) $(\mathbf{P\text{-}IMMUNE} \cup \mathbf{FIN}) \cap \{L \mid L \leqslant_m^{\text{lin}} A\}$, *for any* $A \in \mathbf{REC}$.

Proof. Each of (a)–(c) is closed under splices by sets in **P**. To see this for (c), let $A, B \in \mathbf{P\text{-}IMMUNE} \cup \mathbf{FIN}$, let $D \in \mathbf{P}$, and put $E := (A \cap D) \cup (B \cap \tilde{D})$. Suppose that E is neither finite nor **P**-immune. Then there exists an infinite language $C \in \mathbf{P}$ such that $C \subseteq E$. Then $C \cap D$ and $C \cap \tilde{D}$ are both in **P**, and are respectively contained in A and B. But at least one of $C \cap D$ and $C \cap \tilde{D}$ is infinite, so at least one of A and B is infinite and not **P**-immune. This contradiction finishes (c). Part (d) follows from (c) and Proposition 5.8(b) if $A \notin \{\emptyset, \Sigma^*\}$, and holds trivially if $A = \emptyset$ or $A = \Sigma^*$. ∎

In Section 6 we discuss the possibility of strengthening Theorems 5.6 and 5.7 so that they apply to classes $\mathscr{D}$ such as $\mathbf{NPC} := \{L \mid L \equiv_m^p \mathrm{SAT}\}$ and $\mathbf{NPI} := \{L \mid L \equiv_{\mathrm{iso}}^p \mathrm{SAT}\}$. Neither of these is closed under splices by the single language $\varnothing \oplus \Sigma^*$, even if $\mathbf{P} = \mathbf{NP}$.

6. Some Applications

In Section 4 we showed that provable properties Π of languages correspond to r.e.-presentable classes, and under slight strenghtenings of the condition for Π to be provable, to r.r. and recursively presentable classes. In Section 5 we developed a method for showing that certain classes are not r.p. or r.r., and in some cases not r.e.-presentable. Combining this work gives us a general technique for generating independence results. We apply it for properties which have attracted much attention in the literature. In all applications, however, we highlight the simple recursive machinery which makes them work, as opposed to elements which are more specific to, say, polynomial-time complexity.

Another motivation is to promote an intuitive feel for the "shapes" of complexity classes. We use an analogy between $\mathfrak{R}$ and the usual topology of the plane in diagramming them. The standard basis for $\mathbb{R}^2$ gives all basic closed sets smooth boundaries, and this is how we represent classes having the r.r. property. All of the hyperconnected classes we refer to, such as **P** and **NP**, are closed under splices by sets in **RL**, and so intuitively correspond to convex regions in the plane. The insight that the latter regions are connected in $\mathbb{R}^2$, and thus cannot be decomposed into closed sets, helps illustrate the workings of the technique. In the diagrams we use shading effects to highlight classes which figuratively "contain their boundary," as does a closed region in the plane. The classes on the "open" side of the boundary stand for the unprovable properties. The formal backing provided by $\mathfrak{R}$ for this visual element may help lead to new results, and may be even better for helping researchers recognize wrong paths quickly.

Throughout this section '$\mathfrak{F}$' stands for any fixed sound, recursively axiomatized formal system, and "provable" means "a theorem of $\mathfrak{F}$." For the sake of definiteness the reader may think of $\mathfrak{F}$ as PA or ZFC.

6.1. *Unprovably Intractable Languages*

U. Schöning [Sö82] showed the basic idea that there can be no general procedure for proving that languages in $\mathbf{NP} \setminus \mathbf{P}$ do not belong to **P**. The result extends to say that "no nontrivial property of languages in a class closed under $\leqslant_m^p$ is decidable," a statement reminiscent of *Rice's theorem* in recursion theory. [Reg83b] offered the slight improvement (still assuming $\mathbf{NP} \neq \mathbf{P}$) of showing that particular languages E exist in $\mathbf{NP} \setminus \mathbf{P}$ such that $\mathfrak{F}$ cannot prove '$E \notin \mathbf{P}$,' and stronger still, such that every witness function f for E with respect to (a given universal language defining) **P** grows too fast to be provably recursive in $\mathfrak{F}$. (This is so even without the requirement that f be provably a witness function.) The present result is also mentioned in [Kow84] and cited in [Har84], though these only conclude that '$L(M_k) \notin \mathbf{P}$' is unprovable when M_k belongs to a fixed, provably recursive presentation of **NP** by polynomial-time bounded NDTMs.

THEOREM 6.1. *If* $\mathbf{NP} \neq \mathbf{P}$, *then for any formula* $\psi(\cdot)$ *repremsenting* $\mathbf{NP} \setminus \mathbf{P}$ *over* $\mathfrak{F}$, *there is a language* $\mathbf{E} \in \mathbf{NP} \setminus \mathbf{P}$ *such that for all TMs* M_e *accepting* E, $\mathfrak{F}$ *fails to prove* $\psi(e)$.

Proof. **NP** is closed under $\leqslant_m^p$, and $\mathbf{P} \supseteq \mathbf{FIN}$. By Proposition 5.8(b), one can apply Theorem 5.7 with $\mathscr{D} := \mathbf{NP}$. Since $(\mathbf{NP} \setminus \mathbf{P}) \cap \mathbf{FIN} = \varnothing$, $\mathbf{NP} \setminus \mathbf{P}$ is not r.e.-presentable. The existence of E then follows from Theorem 4.3(a). ∎

Call such a language E *unprovably intractable* (with respect to $\mathfrak{F}$ and the definition ψ). Figure 6.1(a) illustrates the situation. Note that the unprovability of '$L(M_e) \notin \mathbf{P}$' does not depend on the choice of M_e accepting E. Without the assumption $\mathbf{NP} \neq \mathbf{P}$ we obtain the following result:

THEOREM 6.2. *Let A be any language not in* $\mathbf{P}$, *and let* $\mathcal{C}$ *be the class represented by some formula* $\phi(\cdot)$ *of TM indices over* $\mathfrak{F}$.

(a) *If* $\mathbf{P} \nsubseteq \mathcal{C}^f$, *then one can find a language E such that* $E \notin \mathbf{P}$, $E \leqslant_m^p A$, *and for all provably total TMs* M_e *accepting* E, $\mathfrak{F} \nvdash \phi(e)$.

(b) *If* $\mathbf{FIN} \nsubseteq \mathcal{C}^f$, *then one can find E such that* $E \notin \mathbf{P}$, $E \leqslant_m^{\text{lin}} A$, *and for all TMs* M_e *accepting* E, $\mathfrak{F} \nvdash \phi(e)$.

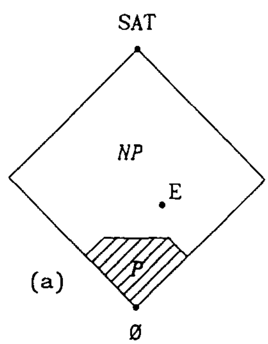

FIG. 6.1a. *Assumption*: $\mathbf{NP} \neq \mathbf{P}$. *Conclusion*: There exist unprovably intractable languages E in $\mathbf{NP} \setminus \mathbf{P}$. **NP** is closed downward under $\leqslant_m^p$, so it is connected in $\mathfrak{R}$. **P** is closed in $\mathfrak{R}$, and if $\mathbf{NP} \neq \mathbf{P}$, forms a nonempty proper subset of **NP**. Then $\mathbf{NP} \setminus \mathbf{P}$ is not closed in $\mathfrak{R}$. Hence membership in $\mathbf{NP} \setminus \mathbf{P}$ is not a provable property of languages, even with regard to arbitrarily strong sound, recursively axiomatized formal systems $\mathfrak{F}$, and under a very weak notion of provable properties. In fact, since $\mathbf{NP} \setminus \mathbf{P}$ contains no finite sets, it is not r.e.-presentable, and so not even the weakest such notion considered in Section 4 holds for $\mathbf{NP} \setminus \mathbf{P}$.

Proof. (a) For all i, define $\psi(i) := \phi(i) \wedge$ 'M_i is total,' and define $\mathscr{E} := \{L(M_i) \mid \mathfrak{F} \vdash \psi(i)\}$. Then $\mathscr{E} \subseteq \mathcal{A}$. If $\mathscr{E}$ is not scfv then some finite variation of A lies outside $\mathscr{E}$ and so meets the conclusions. Else $\mathscr{E}$ is r.p. by Lemma 4.2(c), and so $\mathscr{E}^f$ is r.p. Thus $\mathscr{E}^f$ and $\mathbf{P}$ are closed in $\mathfrak{R}$. Now put $\mathscr{D} := \{L \mid L \leqslant_m^p A\}$. Then $\mathscr{D}$ is closed under $\leqslant_m^p$ and $\oplus$, so by Proposition 5.8(a) it is hyperconnected in $\mathfrak{R}$. This means that $\mathscr{D} \subseteq \mathscr{E}^f \cup \mathbf{P}$ iff $\mathscr{D} \subseteq \mathscr{E}^f$ or $\mathscr{D} \subseteq \mathbf{P}$.

Now $\mathscr{D} \nsubseteq \mathbf{P}$ because $A \notin \mathbf{P}$, and $\mathscr{D} \nsubseteq \mathscr{E}^f$ because $\mathbf{P} \subseteq \mathscr{D}$ and $\mathbf{P} \nsubseteq \mathscr{E}^f$. (Note. $\mathbf{P} \subseteq \mathscr{D}$ because $\mathbf{P}$ is contained in the zero-degree of $\leqslant_m^p$; $\mathbf{P} \nsubseteq \mathscr{E}^f$ because $\mathfrak{F}$ is sound.) Hence $\mathscr{D} \setminus (\mathscr{E}^f \cup \mathbf{P})$ is nonempty, and any language E in the difference has the required properties.

(b) Take $\mathscr{D} := \{L \mid L \leqslant_m^{\text{lin}} A\}$, $\mathscr{C} := \mathbf{P}$, and $\mathscr{E} := \{L(M_i) \mid \mathfrak{F} \vdash \phi(i)\}$. Then apply Theorem 5.7 to conclude that $\mathscr{E}^f \nsupseteq (\mathscr{D} \setminus \mathscr{C})$. ∎

Figure 6.1(b) diagrams the situation in Theorem 6.2(b). One can also apply Theorem 5.5 to obtain E directly, and in (b) conclude also that $E \leqslant_{li}^{rl} A \oplus \varnothing$.

Taking $\mathcal{A}$ to be $\mathbf{RE} \setminus \mathbf{P}$ shows that there are unprovably intractable languages which reduce to A. Since there is nothing really special about $\mathbf{P}$ and $\mathbf{NP}$ in the above, we can state unconditionally (and somewhat less formally):

THEOREM 6.3. **EXPTIME\P** *contains unprovably intractable languages.* **PSPACE\LOGSPACE** *contains languages E such that '$E \in$* **LOGSPACE**' *is consistent with $\mathfrak{F}$.* **NLIN\DLIN** *contains languages E which do not provably require more than linear time on a deterministic TM.*

The proof of each follows quickly from the downward closure of the larger class

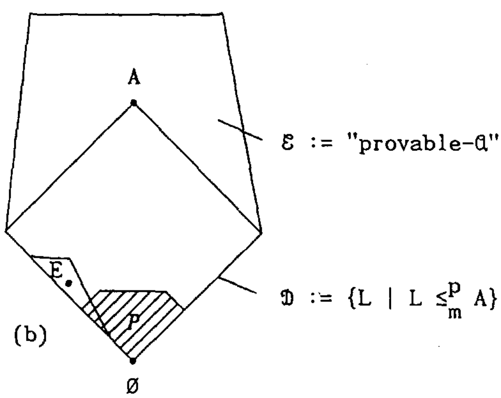

FIG. 6.1b. *Assumptions*: $A \in \mathbf{REC} \setminus \mathbf{P}$, $\mathbf{P} \nsubseteq \mathcal{A}$. (Consider the case $\mathcal{A} := \mathbf{REC} \setminus \mathbf{P}$.) *Conclusions*: $(\exists E \subseteq \Sigma^*)$: $E \leqslant_m^p A$, $E \notin \mathbf{P}$, and $\mathfrak{F}$ fails to prove '$E \in \mathcal{A}$.' (In this case, $\mathfrak{F}$ fails to prove '$E \notin \mathbf{P}$.') Given a formula $\phi(\cdot)$ representing $\mathcal{A}$ over $\mathfrak{F}$, define $\mathscr{E} := \{L(M_i) \mid \mathfrak{F}$ proves $\phi(i) \wedge$ 'M_i is total'$\}^f$. Then $\mathscr{E}$ is closed in $\mathfrak{R}$, as is $\mathbf{P}$. $\mathscr{D}$ is hyperconnected in $\mathfrak{R}$. From this alone we conclude that since $\mathscr{E}$ fails to contain all of $\mathbf{P}$, it also fails to contain all of $\mathscr{D} \setminus \mathbf{P}$. In cases where $\mathcal{A}$ is disjoint from $\mathbf{P}$, such as $\mathcal{A} := \mathbf{REC} \setminus \mathbf{P}$, the connectedness of $\mathscr{D}$ suffices to show $\mathscr{D} \setminus \mathbf{P} \nsubseteq$ "provable-$\mathcal{A}$."

under $\leqslant_m^{\text{lin}}$. In each case we can even render it unprovable that E is infinite, as we expound upon in Section 6.4.

Returning to Theorem 6.1, we remark that E itself technically does depend on both $\mathfrak{F}$ and ψ. Let M_e be a fixed TM accepting E. One may first define a new formal system $\mathfrak{F}'$ by adding to $\mathfrak{F}$ the axiom '$\psi(e)$.' $\mathfrak{F}'$ is still recursively axiomatized and sound since we have added a single, true axiom. However, $\mathfrak{F}' \vdash \psi(e)$ trivially. Second, one may keep $\mathfrak{F}$ fixed and replace ψ by defining $\eta(i) := [\psi(i) \vee i = e]$ for all i. Then $\eta(\cdot)$ still represents $\mathbf{NP} \backslash \mathbf{P}$, but $\mathfrak{F}$ does prove the sentence '$L(M_e) \notin \mathbf{P}$' formalized via η in place of ψ.

Nevertheless, the formal system one implicitly works within while trying to prove properties Π for specific languages do not have any such *ad hoc* axioms. Moreover, one generally refers to a definition ψ of Π set down in the first place. For example, the following generic definition of $\mathbf{P}$ is used by many researchers: for each i let P_i be the TM obtained by attaching an $n^i + i$ *clock* to M_i; then $[P_i]_{i=1}^{\infty}$ is a recursive presentation of $\mathbf{P}$. The formula $\psi(i) := (\forall j)(\exists x)$'$x \in L(M_i) \triangle L(P_j)$' then expresses "$L(M_i) \notin \mathbf{P}$" in a uniform manner in $\mathfrak{F}$, and represents $\mathbf{RE} \backslash \mathbf{P}$. (That this may not be the best definition to use in practice is hinted by results of D. Kozen [Koz80].)

To obtain the stronger conclusions of [Reg83b] mentioned above, verify that the class of recursive languages having a provably recursive witnss function with respect to $[P_j]_{j=1}^{\infty}$, namely $\{L(M_i) \mid$ for some transducer T: $\mathfrak{F} \vdash$ ['T is total' $\wedge$ $(\forall j)$'$T(x) \in L(M_i) \triangle L(P_j)$']$\}$, has the r.r. property. Then use the (hyper)connectedness of $\mathbf{NP}$ exactly as before.

The same technique gives analogous results for classes such as $\mathbf{P}$ vs $\mathbf{NC}$, $\mathbf{L}$ vs $\mathbf{NL}$, $\mathbf{PH}$ vs $\mathbf{PSPACE}$, and $\mathbf{DSPACE}[n]$ vs $\mathbf{NSPACE}[n]$. The general idea is that for any effective reducibility relation $\leqslant_r$ extending $\leqslant_m^{rl}$, non-membership in the zero-degree of $\leqslant_r$ is not a provable property of languages.

6.2. *A Barely Provable Property of Languages*

This example is intended to show some of the subtlety involved in the results of Section 4. Let Π be the property 'If L is infinite then L is in **PSPACE** but not in **LOGSPACE**,' so that the class $\mathscr{C}_\Pi$ equals $(\mathbf{PSPACE} \backslash \mathbf{LOGPACE}) \cup \mathbf{FIN}$.

THEOREM 6.4. *For Π as above, $\mathscr{C}_\Pi$ is representable by a provable formula which hits total TMs. However, for any formula $\psi(\cdot)$ representing $\mathscr{C}_\Pi$ and recursive function u: $\mathbb{N} \to \mathbb{N}^+$, there exists a language $E \in \mathscr{C}_\Pi$ such that for any TM M_e accepting E, if $\mathfrak{F} \vdash \psi(e)$ then M_e does not run in time $u(n)$.*

Proof. By Theorem 3.12(e), the full index set of $\mathbf{PSPACE} \backslash \mathbf{LOGSPACE}$ is in Σ_3^0. Since $I_{\mathbf{FIN}} \in \Sigma_2^0$ and Σ_3^0 is closed under union, $I_{\mathscr{C}_\Pi} \in \Sigma_3^0$. Since $\mathscr{C}_\Pi \supseteq \mathbf{FIN}$, the first statement follows by Lemma 4.2(b).

However, **PSPACE** is hyperconnected in $\mathfrak{R}$ (by Theorem 5.6), and equals the nontrivial union of $\mathscr{C}_\Pi$ and **LOGSPACE**. Since **LOGSPACE** is r.p., $\mathscr{C}_\Pi$ cannot be r.p., and since $\mathscr{C}_\Pi$ is scfv, the conclusion follows from Theorem 4.4. ∎

See Fig. 6.2. Though membership of languages E in $\mathscr{C}_\Pi$ is provable, this shows that there is no recursive upper bound on the "badness" of TMs M_e under which one is able to prove '$L(M_e) \in \mathscr{C}_\Pi$.' For any practical purpose, Π is not a provable property of languages.

6.3. *Unprovably Immune Languages*

Conventionallly, a language A is **P**-*immune* if A is infinite and no infinite subset of A is in **P**. The following result does not hold if one also calls finite sets **P**-immune, as observed in Section 3.3.3.

THEOREM 6.5. *The class* **P-IMMUNE** *of recursive* **P**-*immune languages is neither closed in* $\mathfrak{R}$ *nor presentable by r.e. indices.*

Proof. Take $\mathscr{D} := \mathbf{P\text{-}IMMUNE} \cup \mathbf{FIN}$, and $\mathscr{C} := \mathbf{FIN}$. Proposition 5.9(c) shows that $\mathscr{D}$ is connected (in fact, hyperconnected) in $\mathfrak{R}$, and so $\mathscr{D} \setminus \mathscr{C}$ is not closed in $\mathfrak{R}$. That $\mathscr{D} \setminus \mathscr{C}$ is not r.e.-presentable either follows from Theorem 5.7. ▮

COROLLARY 6.6. (a) *For any formula* $\psi(\cdot)$ *representing* **P**-*immunity over* $\mathfrak{F}$, *there exists a recursive* **P**-*immune language* E *such that* $\mathfrak{F}$ *does not prove* $\psi(e)$ *for any TM* M_e *accepting* E.

(b) *Moreover, given any recursive* **P**-*immune language* A *we can get* $E \leqslant_m^{\text{lin}} A$.

Proof. (a) This follows from Theorem 6.5 and Theorem 4.3(a). (b) Take $\mathscr{D} := \{L \mid L \leqslant_m^{\text{lin}} A\} \cap (\mathbf{P\text{-}IMMUNE} \cup \mathbf{FIN})$. Proposition 3.9(d) showed that $\mathscr{D}$ is closed under splices by sets in **RL**. With $\mathscr{C} := \mathbf{FIN}$ we have $\varnothing \subsetneq \mathscr{C} \subsetneq \mathscr{D}$ since A is **P**-immune. Then $\mathscr{D} \setminus \mathscr{C}$ comprises the **P**-immune languages which $\leqslant_m^{\text{lin}}$-reduce to A.

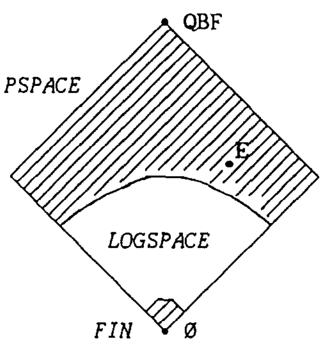

FIG. 6.2. A barely provable property of languages. Although $\mathscr{C} := (\mathbf{PSPACE} \setminus \mathbf{LOGSPACE}) \cup \mathbf{FIN}$ (shaded) is not closed in $\mathfrak{R}$, the full index set of $\mathscr{C}$ is in Σ_3^0, and $\mathscr{C}$ contains all finite sets. This makes $\mathscr{C}$ r.e.-presentable, and so membership in $\mathscr{C}$ is a provable property. However, for any recursive space bound $u(n)$, there are languages $E \in \mathscr{C}$ whose membership in $\mathscr{C}$ can be proved only under representations by TMs using more than $u(n)$ space.

By Theorem 5.7, no r.e.-representable class $\mathscr{E}$ such that $\mathscr{E} \cap \mathbf{FIN} = \emptyset$ contains $\mathscr{D} \setminus \mathscr{C}$. This is true in particular of $\mathscr{E} := \{L(M_i) \mid \mathfrak{F} \vdash \psi(i)\}$. ∎

Figure 6.3 shows this graphically. In particular, such "unprovably **P**-immune" languages exist in **EXPTIME** since this class is closed under $\leqslant_m^{\text{lin}}$ and contains not only **P**-immune but also **P**-biimmune languages [BeHa77].

6.4. *Unprovably Infinite Languages*

The actual rub in Corollary 6.6 is that $\mathfrak{F}$ may not even be able to prove that E is infinite. We formalize this by defining '$L(M_i)$ is infinite' $\leftrightarrow (\forall x)(\exists y, m)$ $[\text{Accept}(y, i, m) \wedge y \geqslant x]$ for all i. Much the same proof shows:

THEOREM 6.7. *Let $\mathscr{D}$ be any class closed downward under $\leqslant_m^{\text{lin}}$, and let $\mathscr{C}$ be any r.p. cfv class which does not contain $\mathscr{D}$. Then $\mathscr{D} \setminus \mathscr{C}$ contains infinite languages E such that for all M_e accepting E, $\mathfrak{F} \nvdash$ '$L(M_e)$ is infinite.'*

Proof. $\mathscr{D} \setminus \mathscr{C}$ cannot equal **FIN**, since $\mathscr{D}$ is connected in $\mathfrak{R}$ and **FIN** is closed. So $\mathscr{D} \setminus \mathscr{C}$ contains an infinite set A. Now take $\mathscr{D}' := \{L \mid L \leqslant_m^{\text{lin}} A\}$ and $\mathscr{C}' := \mathscr{C} \cup \mathbf{FIN}$. By Theorem 5.7, the r.e.-presentable class $\mathscr{E} := \{L(M_i) \mid \mathfrak{F} \vdash$ '$L(M_i)$ is infinite'$\}$ does not contain all of $\mathscr{D}' \setminus \mathscr{C}'$, so $\mathscr{D} \setminus \mathscr{C} \nsubseteq \mathscr{E}$. ∎

Reviewing Theorem 6.1, we see that if $\mathbf{NP} \neq \mathbf{P}$ then $\mathbf{NP} \setminus \mathbf{P}$ contains languages E such that it is consistent with $\mathfrak{F}$ to believe that E is finite, more than that $E \in \mathbf{P}$. By examining the mechanics of Section 5.1 more closely (see [Reg86b]), one can in fact construct E directly as follows: Let $s_{\mathfrak{F}}$ be an increasing recursive function which grows faster than any function $\mathfrak{F}$ can prove recursive, and let D be any language in **RL** which outruns $s_{\mathfrak{F}}$. Then for any $A \in \mathbf{NP}$, the splice of A and $\emptyset$ by D is in **NP**, but is not provably infinite.

Furthermore, any infinite language A equals the disjoint union of two languages E_1 and E_2, both of which are unprovably infinite and $\leqslant_m^{\text{lin}}$-reduce to A. If A is a polynomial cylinder other than Σ^*, such as $A := \text{SAT}$, then also $A \equiv_{\text{iso}}^p (E_1 \oplus E_2)$.

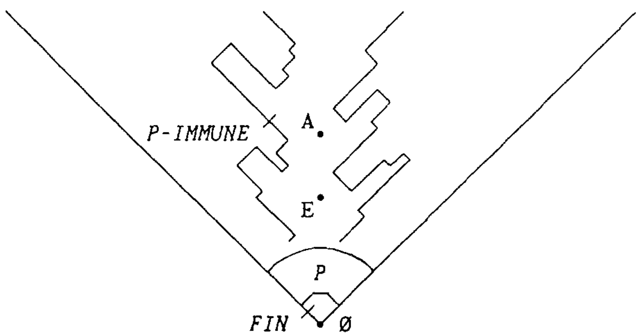

FIG. 6.3. **P-IMMUNE** is not r.e.-presentable. **P-IMMUNE** equals the difference of the connected class **P-IMMUNE** $\cup$ **FIN** and the closed class **FIN**, and so is neither closed in $\mathfrak{R}$ nor r.e.-presentable.

One might hope that languages E in $\mathbf{NP}\backslash\mathbf{P}$ which are not provably infinite would be "nearly" in $\mathbf{P}$. One notion of "nearly in $\mathbf{P}$" is that E be $\mathbf{P}$-*close*, i.e., that $E = A \triangle B$ for some $A \in \mathbf{P}$ and polynomially sparse language B. However, the corresponding class **P-CLOSE** has the r.r. property, and contains SAT iff $\mathbf{NP} = \mathbf{P}$ (cf. [Sö82] and [Sö85], pp. 48–49). Purely on account of this we have the situation shown in Fig. 6.4:

COROLLARY 6.8. *Assuming* $\mathbf{NP} \neq \mathbf{P}$, *there exist unprovably infinite languages in* $\mathbf{NP}\backslash\mathbf{P}$ *which are not* $\mathbf{P}$-*close.*

Proof. Take $\mathscr{C} :=$ **P-CLOSE**, $\mathscr{D} := \mathbf{NP}$, and $A :=$ SAT in the last proof. ∎

This refutes the idea, and several others meet similar fates. Intuitively speaking, one can diagonalize away from "positive" properties almost at will.

We remark also that unprovably infinite languages exist in profusion between any two languages A and B, one of which is either finite or unprovably infinite. That is, let $\leqslant_r$ be any effective reducibility relation extending $\leqslant_m^{rl}$ such that '$\oplus$' acts as a least-upper-bound operation. Suppose $A <_r B$, and let $\mathscr{C}$ be the class of unprovably infinite languages C such that $A <_r C <_r B$. Given any countable partially ordered set $(\mathscr{X}, \preccurlyeq_{\mathscr{X}})$, one can then find a 1–1 mapping $f: \mathscr{X} \to \mathscr{C}$ which *embeds* $\mathscr{X}$ into $\mathscr{C}$, meaning that $x \preccurlyeq_{\mathscr{X}} y \Leftrightarrow f(x) \leqslant_r f(y)$ for all $x, y \in \mathscr{X}$. In particular, $\mathscr{C} \neq \varnothing$. This can be shown by combining Theorems 3.9(a, b) and 6.7 with the main theorems of [Meh76] or [A-S84]; for further details see [Reg86b].

We inquire whether one can sharpen the condition in Theorem 6.7 that $\mathscr{D}$ be closed under $\leqslant_m^{\text{lin}}$ (alternatively, closed under $\bigcap$ with sets in **RL**) even further and still obtain that infiniteness is not a probable property of languages in $\mathscr{D}$. [Sdt85] observes that finiteness is a decidable property of regular and context-free languages, and in parallel that neither **REG** nor **CFL** is "recursive gap closed," on account of the *Pumping Lemmas* for these classes. We ask for the maximum growth

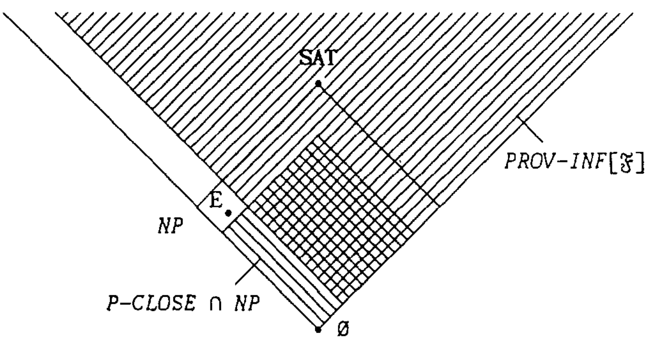

FIG. 6.4. There exist unprovably infinite languages which are not **P**-close. *Assumption*: $\mathbf{NP} \neq \mathbf{P}$ (so SAT $\notin$ **P-CLOSE**). *Conclusion*: $(\exists E \in \mathbf{NP})$: E is neither **P**-close nor provably infinite. Since **NP** is hyperconnected in $\mathfrak{R}$, **NP** cannot be contained in the union of the classes of **P**-close and provably infinite languages (unless $\mathbf{NP} = \mathbf{P}$).

rate that recursive functions can have and still be out-run by some CFL, with a view to obtaining independence results about CFLs for subsystems of PA.

Our last two applications are somewhat more complex than the above, and lead up to several open problems.

6.5. *Unprovability of* $\mathbf{NP}^E \neq \mathbf{P}^E$

One indicated way to formalize '$\mathbf{NP}^A = \mathbf{P}^A$' over a given formal system $\mathfrak{F}$ is to take an enumeration $[P_k]_{k=1}^{\infty}$ of polynomial-time bounded OTMs, and use the OTM Z constructed in [BGS75] such that $L(Z^A)$ is $\mathbf{NP}^A$-complete for any $A \subseteq \Sigma^*$. Then the formula $\eta(\cdot)$ defined for all i by

$$\eta(i) := (\forall j)[`L(P_j^{L(M_i)}) \neq L(Z^{L(M_i)})\text{'}] \tag{6.1}$$

can be taken to express '$\mathbf{P}^{L(M_i)} \neq \mathbf{NP}^{L(M_i)}$.'

All we actually need to suppose is that $\eta(\cdot)$ is *some* formula over $\mathfrak{F}$ representing $\mathbf{REC} \backslash \mathbf{EQ}$. We refer to $\eta(\cdot)$ above for the sake of definiteness.

THEOREM 6.9 [Reg86a]. *Let $\mathfrak{F}$ be any sound r.a. formal system which encodes the formula $\eta(\cdot)$ given in* (6.1). *Then there exists a language $E \in (\mathbf{PSPACE} \cup \mathbf{EXPTIME})$ such that $\mathbf{NP}^E \neq \mathbf{P}^E$, but $\mathfrak{F}$ fails to prove $\eta(e)$ for any provably total M_e accepting E.*

The theorem holds with the conclusion $E \in \mathbf{PSPACE}$ if $\mathbf{NP} \neq \mathbf{P}$, and with $E \in \mathbf{EXPTIME}$ if $\mathbf{NP} = \mathbf{P}$. It is not known whether **PSPACE** is contained in **EXPTIME** or vice versa, In [Har84], such a language E is asserted to exist in $\mathbf{DSPACE}[2^{2^n}]$, which properly contains the union of these classes, so in either case the conclusion is more acute.

To prove Theorem 6.9, we devote separate lemmas to the cases $\mathbf{NP} = \mathbf{P}$ and $\mathbf{NP} \neq \mathbf{P}$. Compare Figs. 6.5(a) and 6.5(b). In the first we show that E can be made linear-time reducible to any given recursive language B such that $\mathbf{NP}^B \neq \mathbf{P}^B$. The demonstration in [BGS75] that such languages B exist in **EXPTIME** and the closure of **EXPTIME** under $\leqslant_m^{\text{lin}}$ account for half of the conclusion about E in Theorem 6.9.

LEMMA 6.10. *Suppose $\mathbf{NP} = \mathbf{P}$, and let B be any recursive language such that $\mathbf{NP}^B \neq \mathbf{P}^B$. Then there exists a language E such that $E \leqslant_m^{\text{lin}} B$ and $\mathbf{NP}^E \neq \mathbf{P}^E$, but $\mathfrak{F}$ fails to prove $\eta(e)$ for any M_e accepting E.*

Proof. Take $\mathscr{D} := \{L \mid L \leqslant_m^{\text{lin}} B\}$ and $\mathscr{C} := \mathbf{EQ} \cap \mathscr{D}$. Then $\mathscr{C}$ is closed in $\mathfrak{R}$. By the hypotheses $\mathbf{P} = \mathbf{NP}$ and $B \notin \mathbf{EQ}$, $\mathbf{FIN} \subseteq \mathscr{C} \subsetneq \mathscr{D}$. By Theorem 5.7, no r.e.-presentable class contains $\mathscr{D} \backslash \mathscr{C}$ without containing **FIN**. Since "provable $\sim$**EQ**" is disjoint from **EQ**, and hence from **FIN**, the rest follows via Theorem 4.3(a). ∎

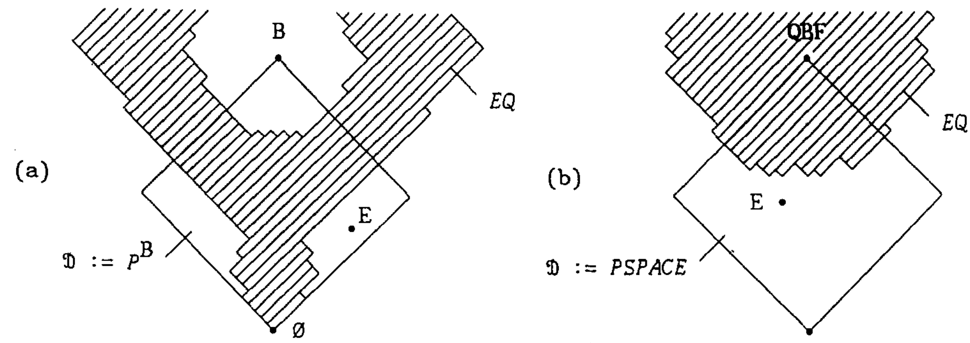

FIG. 6.5. (a) *Assumption*: $\mathbf{NP} = \mathbf{P}$. (B is chosen so that $\mathbf{NP}^B \neq \mathbf{P}^B$; B can be in **EXPTIME**.) *Conclusion*: $(\exists E \in \mathbf{P}^B)$: $\mathbf{NP}^E \neq \mathbf{P}^E$, but $\mathfrak{F}$ fails to prove '$\mathbf{NP}^E = \mathbf{P}^E$.' (b) *Assumption*: $\mathbf{NP} \neq \mathbf{P}$ (Fact: $\mathbf{NP}^{\mathrm{QBF}} = \mathbf{P}^{\mathrm{QBF}}$.) *Conclusion*: $(\exists E \in \mathbf{PSPACE})$: $\mathbf{NP}^E \neq \mathbf{P}^E$, but $\mathfrak{F}$ fails to prove '$\mathbf{NP}^E \neq \mathbf{P}^E$.' In both cases, the class $\mathscr{D}$ represented by the diamond is connected in $\mathfrak{R}$. Let $\mathscr{C} := \{L \mid \mathfrak{F} \vdash \text{'}\mathbf{NP}^L \neq \mathbf{P}^L\text{'}\}$, as formalized in the text. Then $\mathscr{C}$ is closed in $\mathfrak{R}$, and $\mathscr{C}$ is disjoint from **EQ** by the soundness of $\mathfrak{F}$. Since **EQ** itself is closed in $\mathfrak{R}$, $\mathscr{D} \setminus \mathbf{EQ}$ (i.e., the unshaded part of the diamond) does not equal $\mathscr{D} \cap \mathscr{C}$.

In the second lemma one can in fact obtain $E \leqslant_m^{\mathrm{lin}} A$ for any given $A \in \mathbf{REC}$ such that $\mathbf{NP}^A = \mathbf{P}^A$. However, no such languages A are known to exist in **PSPACE** other than those which, like QBF, are **PSPACE**-complete under $\leqslant_T^p$.

LEMMA 6.11. *Suppose* $\mathbf{NP} \neq \mathbf{P}$. *Then there exists* $E \in \mathbf{PSPACE}$ *such that* $\mathbf{NP}^E \neq \mathbf{P}^E$, *but* $\eta(e)$ *is not provable for any provably total* M_e *accepting* E.

Proof. Take $\mathscr{C} := \mathbf{EQ} \cap \mathbf{PSPACE}$. Then $\mathscr{C}$ contains QBF but excludes all finite sets by the hypothesis $\mathbf{NP} \neq \mathbf{P}$, so it is a nontrivial property of languages in **PSPACE**. $\mathscr{C}$ is r.p., hence closed in $\mathfrak{R}$, and **PSPACE** is connected in $\mathfrak{R}$. So $\mathbf{PSPACE} \setminus \mathscr{C}$ is not closed in $\mathfrak{R}$, hence not r.p. The conclusion then follows from Theorem 4.3(c). ∎

Proof of Theorem 6.9. Combine Lemmas 6.10 and 6.11. ∎

In Lemma 6.11, $\mathbf{PSPACE} \setminus \mathscr{C}$ contains **FIN** and is a difference of r.p. classes, so it follows from Theorem 3.12 that $\mathbf{PSPACE} \setminus \mathscr{C}$ is presentable by r.e. indices. Then $\mathbf{PSPACE} \setminus \mathscr{C}$ is representable by some provable formula $\psi(\cdot)$, in fact by one which hits total TMs. Whether $\psi(i)$ can be made to express '$\mathbf{NP}^{L(M_i)} \neq \mathbf{P}^{L(M_i)}$' in the "sensible" manner of $\eta(i)$ in Eq. (6.1) above is something we do not know and strongly doubt. We also inquire whether the division into two cases is necessary to reach the stated conclusions about the language E.

6.6. *Unprovability of Non-$\mathbf{P}$-Isomorphism*

J. Hartmanis [Har84] proved a result similar to Theorem 6.12 below, but under the assumption that the sound, recursively axiomatised formal system $\mathfrak{F}$ concerned is strong enough to prove '$\mathbf{NP} \neq \mathbf{P}$.' Hence he posed the question of whether the conclusion holds for $\mathfrak{F} := \mathrm{PA}$, or for any particular $\mathfrak{F}$.

We show nevertheless that the unprovability phenomenon arises for any sound r.a. formal system $\mathfrak{F}$. The main trick involves the use of the join operation $\oplus$ to make up for the present lack of knowledge as to whether **NPC** is connected in the topology $\mathfrak{R}$. (*Remark*. The original use of this trick in my first paper [Re83a] made the error or assuming that $A \equiv^p_{\mathrm{iso}} \mathrm{SAT} \Leftrightarrow (A \oplus \varnothing) \equiv^p_{\mathrm{iso}} \mathrm{SAT}$ for all (**NP**-complete) languages A. This is in fact a conjecture of S. Mahaney [Mah81], and the proof below corrects the error.)

THEOREM 6.12 [Reg86a, b]. *Let $\mathfrak{F}$ be any sound r.a. formal system, and let $\phi(\cdot)$ represent* $\mathbf{NPC} \setminus \mathbf{NPI}$ *over $\mathfrak{F}$. Suppose* $\mathbf{NPC} \setminus \mathbf{NPI}$ *is nonempty. Then there exists a language E which is* **NP**-*complete but not p-isomorphic to SAT, and such that $\phi(e)$ is not provable in $\mathfrak{F}$ for any provably total M_e accepting E.*

Before giving the proof we state two lemmas. We do not need to make any assumptions in the second because the empty class is not r.p.

LEMMA 6.13 [Mah81]. *For any language A, if both $(A \oplus \varnothing) \equiv^p_{\mathrm{iso}} \mathrm{SAT}$ and $(A \oplus \Sigma^*) \equiv^p_{\mathrm{iso}} \mathrm{SAT}$, then $A \equiv^p_{\mathrm{iso}} \mathrm{SAT}$.*

LEMMA 6.14. *The class* $\mathbf{NPC} \setminus \mathbf{NPI}$ *is not recursively presentable.*

Proof. Suppose to the contrary that $\mathbf{NPC} \setminus \mathbf{NPI}$ is recursively presentable. Define $\mathscr{C}_1 := \{L \mid (L \oplus A) \in \mathbf{NPC} \setminus \mathbf{NPI}\}$; that $\mathscr{C}_1$ is r.p. is shown in [Reg83a]. Since **NPC** and **NPI** are cfv, $\mathscr{C}_1$ is also cfv. One may check that for any $A \leqslant^p_m \mathrm{SAT}$, $\mathrm{SAT} \oplus A \equiv^p_{\mathrm{iso}} \mathrm{SAT}$, so it follows that taking $A_1 := \mathrm{SAT}$ gives $A_1 \notin \mathscr{C}_1$.

Define $\mathscr{C}_2 := \{L \mid (A \oplus L) \in \mathbf{NPI}\}$. Then $\mathscr{C}_2$ is r.p. and cfv, with no supposition needed about **NPI** or **NPC**. Since $\mathbf{NPC} \setminus \mathbf{NPI}$ is assumed to be r.p., $\mathbf{NPC} \setminus \mathbf{NPI}$ is nonempty. Let any $A \in \mathbf{NPC} \setminus \mathbf{NPI}$ be given. By Lemma 6.13, either $A \oplus \varnothing \not\equiv^p_{\mathrm{iso}} \mathrm{SAT}$ or $A \oplus \Sigma^* \not\equiv^p_{\mathrm{iso}} \mathrm{SAT}$. If the former, take $A_2 := \varnothing$; if the latter, take $A_2 := \Sigma^*$ instead. In either case $A_2 \notin \mathscr{C}_2$.

Then $\mathscr{C}_1, \mathscr{C}_2$ and A_1, A_2 satisfy the hypotheses of Theorem 5.1(a). Hence there exists a language $E \notin (\mathscr{C}_1 \cup \mathscr{C}_2)$ such that $E \leqslant^p_m (A_1 \oplus A_2)$, so $E \in \mathbf{NP}$. Then $D := (E \oplus A)$ is **NP**-complete since $E \in \mathbf{NP}$ and A is **NP**-complete. However, $E \notin \mathscr{C}_1$ and $E \notin \mathscr{C}_2$ imply that D is in neither **NPI** nor $\mathbf{NPC} \setminus \mathbf{NPI}$. This is a contradiction. ∎

Proof of Theorem 6.12. Since $\mathbf{NPC} \setminus \mathbf{NPI}$ is not r.p., and $\mathbf{NPC} \setminus \mathbf{NPI}$ is not empty by hypothesis, the conclusion follows from Theorem 4.3(c). ∎

Figure 6.6 is an intuitive rendering of **NPC**, **NPI**, and the classes $\mathscr{D}_1 := \{L \mid L \oplus \Sigma^* \equiv^p_{\mathrm{iso}} \mathrm{SAT}\}$ and $\mathscr{D}_2 := \{L \mid L \oplus \varnothing \equiv^p_{\mathrm{iso}} \mathrm{SAT}\}$. It assumes that all containments not known to hold actually fail. We informally think of $\mathscr{D}_1 \setminus \mathbf{NPI}$ as the class of **NP**-complete languages which fail to be p-isomorphic to SAT through being too "sparse," and of $\mathscr{D}_2 \setminus \mathbf{NPI}$ similarly as those **NP**-complete languages which

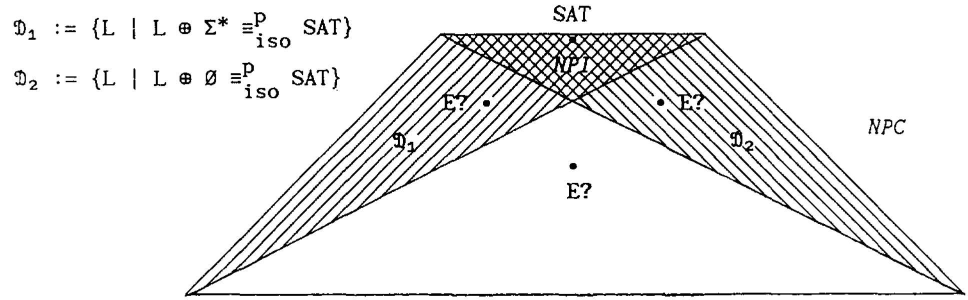

FIG. 6.6. Informal rendition of Theorem 6.12. *Assumption*: $\mathbf{NPC} \neq \mathbf{NPI}$. (The diagram itself also assumes $\mathscr{D}_1 \not\subseteq \mathscr{D}_2$ & $\mathscr{D}_2 \not\subseteq \mathscr{D}_1$.) *Conclusion*: $(\exists E \in \mathbf{NPC})$: $E \not\equiv^p_{\text{iso}}$ SAT, but $\mathfrak{F}$ fails to prove '$E \not\equiv^p_{\text{iso}}$ SAT.' Think of $\mathscr{D}_1$ and $\mathscr{D}_2$ as the classes of **NP**-complete sets which, if they are not already p-isomorphic to SAT, fail only because they are respectively too "light" or too "heavy." *Facts*: $\mathbf{NPI} = \mathscr{D}_1 \cap \mathscr{D}_2$, and **NPI**, **NPC**, $\mathscr{D}_1$, and $\mathscr{D}_2$ are all r.p. (thus closed in $\mathfrak{R}$). *Open question*: Can one show, assuming $\mathbf{NPC} \neq \mathbf{NPI}$, that $\mathbf{NPC} \backslash (\mathscr{D}_1 \cup \mathscr{D}_2)$, i.e., the white region in the diagram, is nonempty? This follows if $\mathbf{NPC} \neq \mathscr{D}_1$, $\mathbf{NPC} \neq \mathscr{D}_2$, and **NPC** is *hyper*connected in $\mathfrak{R}$. The mere connectedness of **NPC** still allows $\mathbf{NPC} \neq \mathscr{D}_1$, $\mathbf{NPC} \neq \mathscr{D}_2$ yet $\mathbf{NPC} = \mathscr{D}_1 \cup \mathscr{D}_2$.

are too "dense." $\mathscr{D}_1 \cap \mathscr{D}_2$ equals **NPI** by Lemma 6.13. Theorem 6.12 still leaves open the possibility that $E \in \mathscr{D}_1$ or $E \in \mathscr{D}_2$. Whether $\mathbf{NPI} \backslash (\mathscr{D}_1 \cup \mathscr{D}_2)$ is empty on the assumption $\mathbf{NPC} \neq \mathbf{NPI}$ is also left open. We conjecture 'no' to the latter, and further conjecture a negative answer to the following

Open Question. Is $\mathbf{NPC} \backslash \mathbf{NPI}$ presentable by r.e. indices?

One open problem bearing on this is whether the implication '$\mathscr{C}$ is r.e.-presentable $\Rightarrow \{L \mid A \oplus L \in \mathscr{C}\}$ is r.e.-presentable' holds for any (scfv) class $\mathscr{C}$ and recursive language A. A 'yes' answer would improve Lemma 6.14 and answer the preceding question negatively.

7. The Topology $\mathfrak{R}$ and Complexity Theory

What does $\mathfrak{R}$ mean for the structure of complexity classes? We have shown that the concepts underlying $\mathfrak{R}$, together with the equivalent characterizations offered in Sections 3 and 4, are natural, significant, and worthy of study. What matters here, though, are applications for the topological properties in themselves. One small contribution is our noting that $\mathfrak{R}$ is not metrizable, which may partly explain why it is so difficult to measure the concept of complexity.

The approach we envision is to find a good characterization of the threshold of complexity for families of languages, past which all nontrivial properties (that are invariant under $\equiv^f$) become undecidable in the strong sense we have given in Sections 4–6. Our work has placed this threshold at or below simultaneous real time and log space. Put another way, properties of languages become generally undecidable whenever an amount of time or (on-line) space that grows with the input size becomes available. The point is that $\mathfrak{R}$ provides a means of expressing this threshold—via conditions for classes to be connected in $\mathfrak{R}$—which does not refer to machines or notions of complexity.

7.1. *The Connectedness Problem*

To reiterate our findings (Theorems 5.6–5.8): if a class $\mathscr{D} \subseteq$ **REC** is closed under $\bigcap$ with languages in **RL**, then $\mathscr{D}$ is connected in the topology $\mathfrak{R}$. If $\mathscr{D}$ is closed under splices by sets in **RL**, then $\mathscr{D}$ is hyperconnected in $\mathfrak{R}$. The same hold true if $\mathscr{D}$ is respectively closed downward under $\leqslant_m^{\text{lin}}$ or under $\leqslant_{li}^{rl}$ and $\oplus$, and *a fortiori* if $\mathscr{D}$ is closed under the more familiar polynomial-time or log-space analogs of these reducibilities. We show a progression of the degrees to which certain classes are or might be connected in $\mathfrak{R}$.

EXAMPLES. (a) **PSPACE**, **CSL**, **P**, **NP**, **co-NP**, and **NP** $\cap$ **co-NP** are all hyperconnected in $\mathfrak{R}$.

(b) **NP** $\cup$ **co-NP** is connected, and hyperconnected if and only if **NP** = **co-NP**.

(c) **P** $\cup$ **NPC** is disconnected if and only if **P** $\neq$ **NP**.

(d) **CFL** is disconnected.

(e) **REG** is *totally disconnected*, meaning that the only connected subsets of $(\mathbf{REG}/\equiv^f, \mathfrak{R} \restriction \mathbf{REG})$ are singletons. The reason is that for every regular language L, the classes L^f and $\mathbf{REG} \setminus L^f$ are both closed in $\mathfrak{R}$.

The last clause follows because there is a way to present regular languages, namely by finite automata, so that it is possible to decide whether two given automata accept the same language up to finite variations. Similarly it is possible to decide whether a given CFL L is finite when L is given by a pushdown automaton or a context-free grammar, so **CFL** $\setminus$ **FIN** as well as **FIN** is closed in $\mathfrak{R}$. We do not know whether **CFL** is totally disconnected, or in particular whether **CFL** $\setminus$ **co-FIN** is closed in $\mathfrak{R}$; note that it is undecidable whether a given CFL is co-finite. **CFL** $\setminus$ **co-FIN** is not a recursive translation of **CFL** $\setminus$ **FIN** because **CFL** is not closed under complements.

These examples indicate that the following question is at least significant, in that it embraces many difficult open problems in complexity theory.

GENERAL PROBLEM. Which classes $\mathscr{C} \subseteq$ **REC** are connected in the topology $\mathfrak{R}$? Which classes are hyperconnected?

Even though an "extrinsic" answer will likely be very hard to come by, it may well be possible to find *instrinsic* criteria on the structure of a class $\mathscr{C}$ which hold iff $\mathscr{C}$ is connected or hyperconnected in $\mathfrak{R}$. Such a characterization may illuminate the obstacles for proving properties of recursive languages.

We have been unable to determine whether classes $\mathscr{C}$ which are closed *upward* under effective reducibilities $\leqslant_r$ extending $\leqslant_{li}^{rl}$, and equivalence classes under $\equiv_r$ such as **NPC** and **NPI**, are generally connected or hyperconnected in $\mathfrak{R}$. As remarked before, neither **NPC** nor **NPI** is closed under splices by sets in **RL**, nor even by $\{\emptyset \oplus \Sigma^*\}$. However, a closer look at Section 5 reveals that we need only have **NPC** or **NPI** closed under "enough" splices D to out-run all recursive functions. We make a speculative guess that this is true for any $\equiv_{\text{iso}}^{p}$-degree, and for **NPC** iff **NPC** = **NPI**.

7.2. *Conclusion*

This paper has developed a clean, pictorial approach to recognizing structural features of complexity classes. In particular we have improved much work in the literature on independence results, and provided very general conditions under which one can tell whether a given property of languages is or is not provable. It would not be hard to extend the applications of Section 6 into a long list of independence results, all of them having very much the same character.

This returns us to the query raised in Section 1: What do these results mean? They do not mean that $\mathbf{P} \overset{?}{=} \mathbf{NP}$ and related questions are more likely to be undecidable in formal systems such as Peano arithmetic or set theory. They do mean that unless one makes distinctions finer than equivalence up to real-time log-space reductions, one cannot even come close to deciding any nontrivial properties of general languages in the class under study.

In regard to whether given languages L belong to **P**, they may mean that whether one can prove '$L \notin \mathbf{P}$,' using a given formal system $\mathfrak{F}$, depends wholly on the distribution of hard and easy instances to L. The examples E constructed in Section 6 all have highly irregular distributions, and we ask whether all unprovably intractable languages must have this form when $\mathfrak{F}$ is a natural system such as PA or ZFC. The refinement of the delayed diagonalization technique sketched in Section 5.1, with its explicit link to the maximum growth rate for functions whose total recursiveness is provable in $\mathfrak{F}$, may supply some of the power needed to answer this question.

We have also found that the complexity classes which have been considered attractive objects of study in the literature all share the same basic features: (i) they have relatively simple definitions in arithmetic, (ii) it is possible to enumerate their members effectively, and (iii) all of their members provably belong to the class. Up to some differences brought out in Sections 3.3, 4, and 6.2, these all correspond to the feature of (iv) being closed in the topology $\mathfrak{R}$ generated by the "recursively refutable" classes. Though we have mainly used $\mathfrak{R}$ as a conceptual tool in this paper, we look forward to research in complexity which exploits results that have already been obtained using topology in other branches of mathematics.

One possible avenue for this is suggested by a family of topologies due to A. Visser [Vis80]. He develops them for the study of *numerations*, which are abstract indexings of sets by natural numbers, as applied to the *λ-calculus*. The spaces have many properties in common with (**REC**, $\mathfrak{R}$): notably they are hyperconnected and have naturally associated relational structures which are universal for countable partial orders. Visser remarks that he "hasn't seen the topology in the litarature." We inquire whether the methods of [Vis80] can be used to obtain results about the structure of complexity classes.

REFERENCES

[A-S84] K. AMBOS-SPIES, Sublattices of the polynomial-time degrees, lecture draft, "STACS '84, Paris, France, April 1984;" *Inform. and Control* **65**, No. 1 (1985), 63–84.

[Bak79] T. BAKER, On "provable" analogs of **P** and **NP**, *Math. Systems Theory* **12** (1979), 213–218.

[BGS75] T. BAKER, J. GILL, AND R. SOLOVAY, Relativisations of the **P** = **NP**? question, *SIAM J. Comput.* **4**, No. 4 (1975), 431–442.

[BeGi81] C. BENNETT AND J. GILL, Relative to a random oracle A, $\mathbf{P}(A) \neq \mathbf{NP}(A)$ with probability 1, *SIAM J. Comput.* **10** (1981), 96–113.

[BeHa77] L. BERMAN AND J. HARTMANIS, On isomorphisms and density of **NP** and other complete sets, *SIAM J. Comput.* **6** (1977), 305–321.

[BJ74] G. BOOLOS AND R. JEFFREY, "Computability and Logic," Cambridge Univ. Press, Cambridge, 1974.

[BCH70] A. BORODIN, R. CONSTABLE, AND J. HOPCROFT, Dense and non-dense families of complexity classes, *in* "Proceedings, 2nd ACM Sympos. Theory of Comput., 1970," pp. 7–19.

[CM81] P. CHEW AND M. MACHTEY, A note on structure and looking-back..., *J. Comput. System Sci.* **22** (1981), 53–59.

[Cu80] N. CUTLAND, "Computability," Cambridge Univ. Press, Cambridge, 1980.

[DL80] R. DEMILLO AND R. LIPTON, The consistency of "**P** = **NP**" and related problems with fragments of number theory, *in* "Proceedings, 12th ACM Sympos. Theory of Comput., 1980," pp. 45–57.

[Dowd82] M. DOWD, "Forcing and the **P**-Hierarchy," LCSR-TR-35, Laboratory for Computer Science Research, Rutgers University, 1982.

[Dug66] J. DUGUNDJI, "Topology," Allyn & Bacon, Boston, 1966.

[Gor79] D. GORDON, Complexity classes of probably recursive functions, *J. Comput. System Sci.* **18**, No. 3 (1979), 294–303.

[Gra80] P. GRANT, Some more independence results in complexity theory, *Theoret. Comput. Sci.* **12** (1980), 119–126.

[Háj79] P. HÁJEK, Arithmetical hierarchy and complexity of computation, *Theoret. Comput. Sci.* **8** (1979), 227–237.

[Har78] J. HARTMANIS, "Feasible Computations and Provable Complexity Properties," monograph, Soc. Indus. Appl. Math. Philadelphia, 1978.

[Har84] J. HARTMANIS, Independence results about context-free languages and lower bounds, draft, Cornell University, 1984; *Inform. Process. Lett.* **20**, No. 5 (1985), 241–248.

[HaHo76] J. HARTMANIS AND J. HOPCROFT, Independence results in computer science, *SIGACT News* **8**, No. 4 (1976), 13–24.

[HaSt65] J. HARTMANIS AND R. STEARNS, On the computational complexity of algorithms, *Trans. Amer. Math. Soc.* **117**, No. 5 (1985), 285–306.

[Hu79] J. HOPCROFT AND J. ULLMAN, "Introduction to Automata Theory, Languages, and Computation," Addison–Wesley, Reading, MA, 1979.

[Jos83] D. JOSEPH, Computational complexity in models of ET, *J. Comput. System Sci.* **25** (1983), 205–228.

[JoYg81] D. JOSEPH AND P. YOUNG, Independence results in computer science? *J. Comput. System Sci.* **23** (1981), 205–222.

[Kow84] W. KOWALCZYK, Some connections between presentability of complexity classes and the power of formal systems of reasoning, *in* "Proceedings, Math. Found. Comput. Sci., '84, Prague, August 1984," Lecture Notes in Computer Sci. Vol. 201, Springer-Verlag, Berlin/New York, pp. 364–368, 1984.

[Koz80] D. KOZEN, Indexings of subrecursive classes, *Theoret. Comput. Sci.* **11** (1980), 277–301.

[Lad75] R. LADNER, On the structure of polynomial-time reducibility, *J. Assoc. Comput. Mach.* **22**, No. 1 (1975), 155–171.

[LaRo72] L. LANDWEBER AND E. ROBERTSON, Recursive properties of abstract complexity classes, *J. Assoc. Comput. Mach.* **19**, No. 2 (1972), 296–308.

[Lav77] I. LAVROV, Computable numberings, *in* "Logic, Foundations of Mathematics, and Computability" (R. Butts and J. Hintikka, Eds.), Reidel, New York, 1977.

[LLR81] L. LANDWEBER, R. LIPTON, AND E. ROBERTSON, On the structure of sets in **NP** and other complexity classes, *Theoret. Comput. Sci.* **15** (1981), 181–200.

[Lei82] D. LEIVANT, Unprovability of theorems of complexity theory in weak number theories, *Theoret. Comput. Sci.* **18** (1982), 259–268.

[Lip78] R. LIPTON, Model theoretic aspects of complexity theory, *in* "Proceedings, 19th IEEE Found. of Comput. Sci., 1978," pp. 193–200.

[Lutz87] J. LUTZ, Resource-bounded Baire catagory and small circuits in exponential space, *in* "Proceedings, 2nd Structure in Complexity Theory Conference, Ithaca, NY, June 1987."

[MY78] M. MACHTEY AND P. YOUNG, "An Introduction to the General Theory of Algorithms," North-Holland, New York, 1978.

[Mah81] S. MAHANEY, On the number of isomorphism classes of **NP**-complete sets, *in* "Proceedings, 22nd IEEE Found. of Comput. Sci., 1981," pp. 271–278.

[Meh73] K. MEHLHORN, On the size of sets of computable functions, *in* "Proceedings, 14th IEEE Sympos. on Switching and Automata Theory, 1973," pp. 96–109.

[Meh76] K. MEHLHORN, Polynomial and abstract subrecursive classes, *in* "Proceedings, 17th IEEE Found. of Comput. Sci., 1976," pp. 96–109.

[PaHa77] J. PARIS AND L. HARRINGTON, A mathematical incompleteness in Peano's arithmetic, *in* "Handbook of Mathematical Logic" (J. Barwise, Ed.), Academic Press, New York, 1978.

[Phi79] N. PHILLIPS, The **P** = **NP** question and recent independence results, *Quaestiones Inform.* **1** (1979), 26–29.

[Reg83a] K. REGAN, On diagonalization methods and the structure of language classes, *in* "Proceedings, Found. Computation Theory '83, Borgholm, Sweden, Aug. 1983," Lecture Notes in Computer Sci. Vol. 158, pp. 368–380, Springer-Verlag, Berlin/New York, 1983.

[Reg83b] K. REGAN, Arithmetical degrees of index sets for complexity classes, *in* "Proceedings, Logic and Machines '83, Muenster, W. Germany, May 1983," Lecture Notes in Computer Sci. Vol. 171, pp. 118–130, Springer-Verlag, Berlin/New York, 1984.

[Reg.86a] K. REGAN, The topology of provability in complexity theory, *in* "Proceedings, Structure in Complexity Theory, Berkeley, CA, June 1986," Lecture Notes in Computer Sci. Vol. 223, pp. 291–310, Springer-Verlag, Berlin/New York, 1986.

[Reg86b] K. REGAN, "On the Separation of Complexity Classes," D. Phil. dissertation, Oxford University, 1986.

[Rog67] H. ROGERS, "Theory of Recursive Functions and Effective Computability," McGraw–Hill, New York, 1967.

[Ros67] A. ROSENBERG, Real time definable languages, *J. Assoc. Comput. Mach.* **14**, No. 4 (1967), 645–662.

[Sdt85] D. SCHMIDT, The recursion-theoretic structure of complexity classes, *Theoret. Comput. Sci.* **38**, Nos. 2-3 (1985), 143–156.

[Sö81] U. SCHÖNING, "Untersuchungen zur Struktur von **NP**... ," Ph.D. dissertation, University of Stuttgart, 1981.

[Sö82] U. Schöning, A uniform approach to obtain diagonal sets in complexity classes, *Theoret. Comput. Sci.* **18** (1982), 95–103.

[Sö85] U. Schöning, "Complexity and Structure," Lecture Notes in Computer Sci. Vol. 211, Springer-Verlag, Berlin, 1985.

[Vis80] A. Visser, Numerations, λ-calculus, and arithmetic, *in* "To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism," Academic Press, London/New York, 1980.