6. Context–free Languages

Let $\sigma : \Sigma^* \to G$ be a choice of generators for the group G. $W = \sigma^{-1}(1)$ is called the word problem of G with respect this choice of generators. We have shown that G is finite if and only if W is a regular language. Before considering other cases we investigate how W depends on the choice of generators.

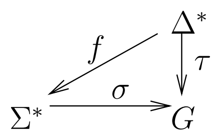
A class of languages, $\mathcal{C}$ is closed under homomorphism if whenever $f : \Sigma^* \to \Delta^*$ is a homomorphism of free monoids and $L \subset \Sigma^*$ is a language in $\mathcal{C}$, then $f(L) \in \mathcal{C}$. Closure under inverse homomorphism is defined similarly. A class $\mathcal{C}$ which contains at least one nonempty language and is closed under homomorphism, inverse homomorphism, intersection with rational languages, and the rational operations of union, product, and generation of submonoid is called a full abstract family of languages.

Exercise 6.1. *The class of rational languages is a full AFL.*

It is a standard result in language theory that context–free languages are a full AFL. Thus by the following theorem it makes sense to speak of a group with a context–free word problem.

Theorem 6.2. *Let $\mathcal{C}$ be a class of languages which is closed under inverse homomorphism. For any finitely generated group G whether or not the word problem of G is in $\mathcal{C}$ does not depend on the choice of generators.*

Proof. Suppose that $\sigma : \Sigma^* \to G$ is a choice of generators for which $W = \sigma^{-1}(1) \in \mathcal{C}$, and take $\tau : \Delta^* \to G$ to be another choice of generators. Choose a homomorphism $f : \Delta^* \to \Sigma^*$ making the diagram below commutative and observe that $\tau^{-1}(1) = f^{-1}(\sigma^{-1}(1)) \in \mathcal{C}$.



□

The argument used in the proof of Theorem 6.2 shows that finitely generated subgroups of groups in $\mathcal{C}$ have their word problems in $\mathcal{C}$ too.

Theorem 6.3. *Let $\mathcal{C}$ be a class of languages which is closed under inverse homomorphism. If the finitely generated group G has it word problem in $\mathcal{C}$, so do all finitely generated subgroups of G.*

Two more properties of context–free languages deserve mention.

Theorem 6.4 (Pumping Lemma). *For every context-free language L there is an integer k such that any word in L of length greater than k can be written as $uvwxy$ so that*

1. $v \neq \lambda$ *or* $x \neq \lambda$;
2. $|vwx| \leq k$;
3. $uv^iwx^iy \in L$, *for all* $i \geq 0$.

Exercise 6.5. *Show that the language $\{a^nb^nc^n \mid n \geq 0\}$ is not context-free.*

Theorem 6.6 (Chomsky Normal Form)**.** *Every context–free language is generated by a grammar in which all productions have the form $A \to a$ or $A \to BC$. A, B, C are nonterminals, and a is a terminal or λ, the empty string.*