MATH 3322

Supplemental notes on
Unifying principles in algebra: Universal Algebra

- A *semigroup* is a structure $\langle\, S;\, * \,\rangle$ such that ...
- A *group* is a structure $\langle\, G;\, *, ^{-1}, \mathbf{e} \,\rangle$ such that ...
- A *ring* is a structure $\langle\, R;\, +, -\ , 0, \times, 1 \,\rangle$ such that ...
- A *boolean algebra* is a structure $\langle\, B;\, \wedge, \vee, ', 0, 1 \,\rangle$ such that ...
- A *group* is a structure $\langle\, H;\, * \,\rangle$ such that ...
- A *real vector space* is a structure $\langle\, V;\, +, -\ , 0, \cdot \,\rangle$ such that ...

These objects are all recognizably things that we study in Algebra. For each we have a concept of homomorphism/isomorphism and a concept of substructure. And for each, in introductory courses in algebra, we repeat for each kind of algebraic structure, proofs about the properties of these basic concepts and others.

Universal Algebra is that area of algebra that studies these common ideas in an abstract setting.

Universal Algebra, G. Grätzer, (two editions 1979 and 1968) QA251.G...

Notation: ω is the set of natural numbers $\{\, 0,\ 1,\ 2,\ 3,\ \ldots \,\}$, usually just taken as a set or at most with the usual order.

Definition 0.1 *An* abstract algebra $\mathcal{A}$ *consists of*

i. *A non-empty set* A, *called the* underlying set *or* universe *of* $\mathcal{A}$.

ii. *a set* $(f_i)_{i\in I}$ *of* operations *on* A, *together with a map* $\nu : I \to \omega\backslash\{\, 0 \,\}$, *such that* $f_i : A^{\nu(i)} \to A$;

iii. *a set* $(c_k)_{k\in K}$ *of* individual constants, *that is,* $c_k \in A$.

Each of I and K can in principle be any set whatsoever, including the empty set, but are usually taken to be ω or some natural number n.

We write

$$\mathcal{A} = \left\langle\, A;\, (f_i)_{i\in I}\,,\, (c_k)_{k\in k} \,\right\rangle$$

The numbers $\nu(i)$ are called the *arities* of the corresponding operations; a 1-ary operation is called *unary*, a 2-ary operation is called *binary*, and a 3-ary operation is called *ternary*. Operations of arity greater that 3 are almost unheard of in standard mathematical practice.

From the standpoint of set theory, constants are really just operations that do not depend on an argument: 0-ary or *nullary* operations. But for many technical reasons, it is often convenient to separate them off in a family of their own.

The basic algebraic "shape" of an abstract algebra is determined completely by the numbers $\nu(i)_{i\in I}$ and by K. These two things together are called the (algebraic) (similarity) *type* of the algebra.

The types of the first four examples above would normally be written as, respectively, $\langle\, 2\,\rangle$, $\langle\, 2, 1, 0\,\rangle$, $\langle\, 2, 1, 0, 2, 0\,\rangle$, and $\langle\, 2, 2, 1, 0, 0\,\rangle$.

Definition 0.2 *An* algebraic language $\mathcal{L}$ *consists of*

i. a set of variables v_0, v_1, v_2, ... *normally taken to be indexed by* ω *but in principle, indexed by any set whatsoever, finite or infinite;*

ii. a set $(\mathbf{f}_i)_{i\in I}$ *of* function symbols*, together with a map* $\nu : I \to \omega \setminus \{\, 0\,\}$*, with the intention that* $\mathbf{f}_i$ *be read as a* $\nu(i)$*-ary function symbol;*

iii. a set $(\mathbf{c}_k)_{k\in k}$ *of* constant *symbols.*

An abstract algebra $\mathcal{A}$ is intended to *interpret* the symbols of an algebraic language $\mathcal{L}$ of the same similarity type. We call it a *structure* for $\mathcal{L}$.

Often we have to talk about more than one structure for $\mathcal{L}$ at a time, and so we interpret the symbols of $\mathcal{L}$ in different ways, depending on the context. To make the context clear, we write $f_i = \mathbf{f}_i{}^{\mathcal{A}}$, and $c_k = \mathbf{c}_k{}^{\mathcal{A}}$, read as "$f_i$ interpreted in $\mathcal{A}$" (etc).

Definition 0.3 *An* assignment α in A *is a map* $\alpha : \omega \to A$*, assigning to the variable* v_n *the value* $\alpha(n)$ *in* A*. We can also write* $\alpha(v_n)$ *or* α_n *or even* a_n *for* $\alpha(n)$*. [If the set of variables is indexed by some set* V *other than* ω*, we of course take* α *to be a map defined on* V*.]*

An $\mathcal{L}$-valuation *is a pair* $\langle\, \mathcal{A}, \alpha\,\rangle$ *where* $\mathcal{A}$ *is an* $\mathcal{L}$ *structure and* α *is an assignment in* A*.*

If v *is a variable and* $a \in A$*, then* $\alpha[a/v]$ *(read "*α *a for* v*") is the assignment in* A *defined by*

$$\alpha[a/v] = \begin{cases} a & \text{if } w = v \\ \alpha(w) & \text{if } w \neq v \end{cases}$$

Definition 0.4 *Let* $\langle\, \mathcal{A},\, \alpha \,\rangle$ *be an* $\mathcal{L}$*-valuation.*

For terms $\mathbf{t}$ *of* $\mathcal{L}$*, we define* the interpretation of $\mathbf{t}$ in $\mathcal{A}$ *at* α (*or* the value *of* $\mathbf{t}$ in $\mathcal{A}$ *at* α), $\mathbf{t}^{\mathcal{A}}[\alpha]$*, by the following:*

(*T1*) *if* $\mathbf{t} = v$*, a variable, then* $\mathbf{t}^{\mathcal{A}}[\alpha] = \alpha(v)$*;*

(*T2*) *if* $\mathbf{t} = \mathbf{c}$*, a constant symbol, then* $\mathbf{t}^{\mathcal{A}}[\alpha] = \mathbf{c}^{\mathcal{A}}$*;*

(*T3*) *if* $\mathbf{t} = \mathbf{f}\mathbf{t}_1 \dots t_n$*, where* $\mathbf{f}$ *is an* n*-ary function symbol and* $\mathbf{t}_1 \dots t_n$ *are* $\mathcal{L}$*-terms, then* $\mathbf{t}^{\mathcal{A}}[\alpha] = \mathbf{f}^{\mathcal{A}}(\mathbf{t}_1^{\mathcal{A}}[\alpha], \dots, \mathbf{t}_n^{\mathcal{A}}[\alpha])$*.*

Definition 0.5 *An* identity (*for* $\mathcal{L}$) *is an expression of the form* $\mathbf{s} \doteq \mathbf{t}$*, where* $\mathbf{s}$ *and* $\mathbf{t}$ *are* $\mathcal{L}$*-terms.*

We define "$\mathcal{A} \models \mathbf{s} \doteq \mathbf{t}[\alpha]$*",* $\mathcal{A}$ satisfies $\mathbf{s} \doteq \mathbf{t}$ at α*, iff* $\mathbf{s}^{\mathcal{A}}[\alpha] = \mathbf{t}^{\mathcal{A}}[\alpha]$*;*

We then define $\mathcal{A} \models \mathbf{s} \doteq \mathbf{t}$ *if* $\mathcal{A} \models \mathbf{s} \doteq \mathbf{t}[\alpha]$ *for all assignments* $\alpha \in \mathcal{A}$*, and say that "the identity* $\mathbf{s} \doteq \mathbf{t}$ holds *in* $\mathcal{A}$*".*

Note that the "..." in the first 4 examples in the introduction can be completed by small sets of identities. What about the last two examples?

Definition 0.6 *An* equational theory *is a set* Σ *of identities (for* $\mathcal{L}$*). We write* $\mathcal{A} \models \Sigma$ *to mean "*$\mathcal{A} \models \sigma$ *for each identity* $\sigma \in \Sigma$*". An* equational class *or* variety $\mathcal{V}$ *is a class of* $\mathcal{L}$*-structures defined by a set of identities: for some* Σ*,* $\mathcal{V} = \{\, \mathcal{A} \text{ an } \mathcal{L}\text{-structure} \colon \mathcal{A} \models \Sigma \,\}$*.*

[***Warning:*** *The word "variety" means something else entirely in algebraic geometry.*]

Subalgebras

Definition 0.7 *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be* $\mathcal{L}$*-algebras such that* $A \subseteq B$*. We call* $\mathcal{A}$ *a* subalgebra *of* $\mathcal{B}$*, and write* $\mathcal{A} \subseteq \mathcal{B}$*, if*

i. for all $i \in I$*,* $n = \nu(i)$*, and* $a_1, \dots, a_n \in A$*,*

$$\mathbf{f}_i^{\mathcal{B}}(a_1, \dots, a_n) = \mathbf{f}_i^{\mathcal{A}}(a_1, \dots, a_n), \text{ and}$$

ii. for all $k \in K$*,*

$$\mathbf{c}_k^{\mathcal{B}} = \mathbf{c}_k^{\mathcal{A}}.$$

Lemma 0.8 *Let* $\mathcal{B}$ *be an* $\mathcal{L}$*-algebra and* $A \subseteq B$*.*

Then we can define an $\mathcal{L}$*-structure* $\mathcal{A}$ *on* A *so that* $\mathcal{A} \subseteq \mathcal{B}$ *iff "*A *is closed under the operations", that is, iff*

i. for all $i \in I$*,* $n = \nu(i)$*, and* $a_1, \dots, a_n \in A$*,*

$$\mathbf{f}_i^{\mathcal{A}}(a_1, \dots, a_n) \in A, \text{ and}$$

ii. for all $k \in K$,

$$\mathbf{c}_k^{\mathcal{A}} \in A\,.$$

Proof: Exercise. ∎

Homomorphisms

Definition 0.9 *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be* $\mathcal{L}$*-algebras and* $\varphi : A \to B$ *be a map such that*

i. for all $i \in I$, $n = \nu(i)$, *and* $a_1, \ldots, a_n \in A$,

$$\mathbf{f}_i^{\mathcal{B}}(\varphi(a_1), \ldots, \varphi(a_n)) = \varphi(\mathbf{f}_i^{\mathcal{A}}(a_1, \ldots, a_n))\,, \text{ and}$$

ii. for all $k \in K$,

$$\mathbf{c}_k^{\mathcal{B}} = \varphi(\mathbf{c}_k^{\mathcal{A}})\,.$$

Then we call φ *a* homomorphism *from* $\mathcal{A}$ *to* $\mathcal{B}$ *, and write* $\varphi : \mathcal{A} \to \mathcal{B}$ *.*

$\varphi[A] = \{\, \varphi(a) : a \in A \,\}$ *is called the* image *of* φ *.*

A homomorphism $\mathcal{A} \to \mathcal{A}$ *is called an* endomorphism *of* $\mathcal{A}$ *, and if it is a bijection as well, we call it an* automorphism *of* $\mathcal{A}$ *.*

A one-to-one homomorphism is called an injection *or a* monomorphism *, and an onto homomorphism is called a* surjection*.*

[***Warning:*** *"epimorphism" means something different., and not every epimorphism is a surjection.*]

Lemma 0.10 *Let* $\varphi : \mathcal{A} \to \mathcal{B}$ *be a homomorphism. Then the image of* φ , $\varphi[A] = \{\, \varphi(a) : a \in A \,\}$, *is the underlying set of a subalgebra (denoted by* $\mathrm{im}(\varphi)$*) of* $\mathcal{B}$ *.*

Proof: Exercise. ∎

Congruences

Recall from elementary arithmetic that for integers a , b , and $m > 0$, we define

$$a \equiv b \pmod{m}$$

iff $m|(b - a)$, and we saw that this is an equivalence relation with nice arithmetic properties.

In general, if Θ is an equivalence relation on a set A and a, $b \in A$, we write $a \equiv b(\Theta)$ or $a \equiv_\Theta b$ to say that two elements of A are equivalent. We define the *equivalence class* of $a \in A$ (with a variety of notations) by

$$a/\Theta = [a]\Theta = [a]_\Theta = \{\, b \in A : a \equiv b(\Theta) \,\}\,.$$

Recall that the equivalence classes of Θ *partition* A , that is, $A = \bigcup_{a \in A} a/\Theta$ and $a \not\equiv b(\Theta)$ implies $a/\Theta \cap b/\Theta = \emptyset$. A/Θ denotes the set of all equivalence classes of Θ :

$$A/\Theta = \{\, a/\Theta : a \in A \,\}\,.$$

Definition 0.11 *Let Θ be an equivalence relation on A, the underlying set of an $\mathcal{L}$-structure $\mathcal{A}$. We say that Θ is a* congruence relation *on $\mathcal{A}$ if*

i. for all $i \in I$, $n = \nu(i)$, and $a_1, \ldots, a_n \in A$, $b_1, \ldots, b_n \in A$, $a_j \equiv b_j(\Theta)$ (for $1 \leq j \leq n$) implies that

$$\mathbf{f}_i^{\mathcal{A}}(a_1, \ldots, a_n) \equiv \mathbf{f}_i^{\mathcal{A}}(b_1, \ldots, b_n)(\Theta)\,.$$

Note that since the value of a constant does not depend on any parameters, we don't need a "second case" for this definition.

Definition 0.12 *We define an $\mathcal{L}$-structure $\mathcal{A}/\Theta$ on A/Θ by setting*

i. for each $i \in I$, $n = \nu(i)$, $\mathbf{f} = \mathbf{f}_i$, and $a_1, \ldots, a_n \in A$:

$$\mathbf{f}^{\mathcal{A}/\Theta}(a_1/\Theta, \ldots, a_n/\Theta) = \mathbf{f}^{\mathcal{A}}(a_1, \ldots, a_n)/\Theta\,;$$

ii. For each $k \in K$,

$$\mathbf{c}_k^{\mathcal{A}/\Theta} = \left(\mathbf{c}_k^{\mathcal{A}}\right)^{\Theta}\,.$$

Lemma 0.13 *Each $\mathbf{f}_i^{\mathcal{A}/\Theta}$ is* well-defined*, that is, the value assigned to $\mathbf{f}_i^{\mathcal{A}/\Theta}(a_1/\Theta, \ldots, a_n/\Theta)$ does not depend on the particular choices of $a_j' \in [a_j]\Theta$.*

Proof: Exercise. ∎

Lemma 0.14 *The map $\natural : A \to A/\Theta : a \mapsto a/\Theta$ is a surjective homomorphism $\mathcal{A} \to \mathcal{A}/\Theta$.*

Proof: Exercise. ∎

Lemma 0.15 *If $\varphi : \mathcal{A} \to \mathcal{B}$ is a homomorphism, then the relation Θ on A defined by $a \equiv b(\Theta)$ iff $\varphi(a) = \varphi(b)$ is a congruence relation on $\mathcal{A}$, called the* kernel *of φ,* $\ker(\varphi)$.

Proof: Exercise. ∎

Theorem 0.16 *Let $\varphi : \mathcal{A} \to \mathcal{B}$ be a homomorphism.*
Then we have the following commutative diagram:

$$\begin{array}{ccc} \mathcal{A} & \xrightarrow{\quad\varphi\quad} & \mathcal{B} \\ {\scriptstyle\natural}\downarrow & & \uparrow{\scriptstyle\iota} \\ \mathcal{A}/\ker(\varphi) & \xrightarrow[\overline{\varphi}]{} & \mathrm{im}(\varphi) \end{array}$$

where $\overline{\varphi} : \mathcal{A}/\ker(\varphi) \to \mathrm{im}(\varphi) : a/\Theta :\mapsto \varphi(a)$ and $\iota : \mathrm{im}(\varphi) \to \mathcal{B}$ is just the inclusion map.
$\natural$ is a surjection, $\overline{\varphi}$ is an isomorphism, and ι is a monomorphism.

Proof: Exercise. ∎

Natural order relations on subalgebras and congruences

Definition 0.17 *Fix an algebra* $\mathcal{A}$ *.*

i. $\mathrm{Sub}(\mathcal{A})$ *is the set of all subalgebras of* $\mathcal{A}$ *, ordered by* $\mathcal{B} \leq \mathcal{B}'$ *iff* $\mathcal{B}$ *is a subalgebra of* $\mathcal{B}'$ *.*

ii. $\mathrm{Con}(\mathcal{A})$ *is the set of all congruence relations on* $\mathcal{A}$ *ordered by* $\Theta \leq \Theta'$ *iff for all* $a, b \in A$ *,* $a \equiv b(\Theta)$ *implies* $a \equiv b(\Theta')$ *.*

Both $\mathrm{Sub}(\mathcal{A})$ and $\mathrm{Con}(\mathcal{A})$ form *complete, algebraic , lattices* .

Lattices

Definition 0.18

i. A lattice-i *is an algebra* $\mathbb{L} = \langle L; \wedge, \vee \rangle$ *satisfying the identities*

$$\begin{array}{ccc} x \wedge x = x & x \vee x = x & \text{(idempotent)} \\ x \wedge y = y \wedge x & x \vee y = y \vee x & \text{(commutative)} \\ x \wedge (y \wedge z) = (x \wedge y) \wedge z & x \vee (y \vee z) = (x \vee y) \vee z & \text{(associative)} \\ x \wedge (x \vee y) = x & x \vee (x \wedge y) = x & \text{(absorptive)} \end{array}$$

The operations "$\wedge$*" and "*$\vee$*" are called* meet *and* join *respectively.*

A bounded lattice-i *is an algebra* $\mathbb{L} = \langle L; \wedge, \vee, 0, 1 \rangle$ *satisfying* (*in addition to the lattice identities*)

$$x \wedge 1 = x \qquad x \vee 0 = x \qquad \text{(identity)}$$

ii. A lattice-ii *is a partially ordered set* $\mathbb{P} = \langle L; \leq \rangle$ *such that every pair of elements of* L *has a greatest lower bound and every pair of elements of* L *has a least upper bound.*

A bounded lattice-ii *is a lattice-ii with a least element* 0 *and a greatest element* 1 *.*

Proposition 0.19

i. Let $\mathbb{L}$ *be a lattice-i. Then the conditions "*$a \leq b$ *iff* $a \wedge b = a$*" and "*$a \leq b$ *iff* $a \vee b = b$*" define the same relation, and* $\langle L; \leq \rangle$ *is a lattice-ii, with the greatest lower bound of* $\{a, b\}$ *being given by* $a \wedge b$ *and the least upper bound of* $\{a, b\}$ *being given by* $a \vee b$ *.*

If $\mathbb{L}$ *is a bounded lattice-i, then 0 and 1 are the least element and the greatest element respectively of* $\leq$ *.*

ii. If $\mathbb{P}$ is a lattice-ii, and we define operations on L by "$a \wedge b$ is the greatest lower bound of $\{a, b\}$" and "$a \vee b$ is the least upper bound of $\{a, b\}$", then $\langle L; \wedge, \vee \rangle$ is a lattice-i.

If $\mathbb{P}$ is a bounded lattice-ii, then the least element 0 and the greatest element 1 define the identity elements of a bounded lattice-i $\langle L; \wedge, \vee, 0, 1 \rangle$.

Proof: Exercise. ∎

Remark: Once having established the equivalence between the two concepts, we *never* distinguish between the two, and just say "lattice".

Definition 0.20 *A lattice is* complete *if* every *subset X has a greatest lower bound $\bigwedge X$ and a least upper bound $\bigvee X$.*

Clearly if $\mathbb{L}$ is a complete lattice, it is bounded: $0 = \bigvee \emptyset$ and $1 = \bigwedge \emptyset$.

Lemma 0.21 *If $\mathbb{P}$ is a poset in which every subset X has a greatest lower bound [equivalently, every subset has a least upper bound], then $\mathbb{P}$ is a complete lattice.*

Proof: $\bigvee X = \bigwedge \{ y \in L : (\forall x \in X)\, (x \leq y) \}$ ∎

Proposition 0.22 *Let $\mathcal{A}$ be an algebra. Then* $\mathrm{Sub}(\mathcal{A})$ *and* $\mathrm{Con}(\mathcal{A})$ *are complete lattices.*

Proof: If $X \subseteq \mathrm{Sub}\,\mathcal{A}$, then let B be the intersection of the underlying sets of the subalgebras in X. It is easy to check that the conditions of Lemma 8 hold for B, and then that the subalgebra defined on B is in fact the meet of X. We write $\mathcal{B} = \bigcap X = \bigwedge X$.

If $Y \subseteq \mathrm{Con}(\mathcal{A})$, define a relation Θ on A by $a \equiv b(\Theta)$ iff $a \equiv B(\Upsilon)$ for all $\Upsilon \in Y$. It is a straightforward exercise that Θ is a congruence relation on $\mathcal{A}$, and that $\Theta = \bigwedge Y$. ∎

Definition 0.23 *Let $\mathcal{A}$ is an algebra and $X \subseteq A$.*

$\bigwedge \{ \mathcal{B} \in \mathrm{Sub}(\mathcal{A}) : X \subseteq B \}$ *is called* the subalgebra generated by X, *denoted by* $\langle\langle X \rangle\rangle$.

$\bigwedge \{ \Theta \in \mathrm{Con}(\mathcal{A}) : x \equiv y(\Theta)$ *for all* $x, y \in X \}$ *is called* the congruence generated by X, *denoted by* $\Theta(X)$.

In particular, if $a \neq b \in A$, $\Theta(a, b)$ is called the principle congruence *generated by $\{a, b\}$.*

Remark: We are not going to make significant use of the concept of "algebraic" lattice, but I include the relevant definitions here for completeness.

Definition 0.24 *Let* **L** *be a lattice. $a \in L$ is* compact *if whenever $X \subseteq L$ and $a \leq \bigvee X$, then there is finite $X' \subseteq X$ such that $a \leq \bigvee X'$.*

A complete lattice L is algebraic *if every element is the join of algebraic elements.*

It is relatively straightforward to see that the compact elements of $\mathrm{Sub}(\mathcal{A})$ and of $\mathrm{Con}(\mathcal{A})$ are the finitely generated ones; clearly any subalgebra is the join of all the finitely generated (in fact, 1-generated) subalgebras contained in it; and any congruence is the join of all the finitely generated (in fact, principle) congruences contained in it.

Theorem 0.25 *Sub($\mathcal{A}$) and* $\mathrm{Con}(\mathcal{A})$ *are complete algebraic lattices.*

General remarks:

I've said "obvious" or "easy" quite a few times here, but it would still take a lot of time to fill in all the details.All this and much more is covered in Grätzer's book mentioned above; to the end of §10 (in the first edition).

I cannot resist adding the following, although the proofs would require a lot more work than we have time for. The first and second Isomorphism Theorems for groups are actually theorems of Universal Algebra. We note that if Θ is a congruence relation on a group G, then the equivalence class of the identity is a normal subgroup; and every congruence on G is "congruence modulo some normal subgroup".

Theorem 0.26 "First Isomorphism Theorem"

Let $\varphi : \mathcal{A} \to \mathcal{B}$ be a surjective homomorphism, and $\Theta \geq \ker(\varphi)$ a congruence on $\mathcal{A}$. Define a relation $\varphi(\Theta)$ on B by "$b \equiv b'(\varphi(\Theta))$ iff for some $a, a' \in A$, $a \equiv a'(\Theta)$ and $\varphi(a) = b$, $\varphi(a') = b'$"

Then $\varphi(\Theta)$ is a congruence relation on $\mathcal{B}$, and

$$\mathcal{A}/\Theta \cong \mathcal{B}/\varphi(\Theta)\,.$$

by the map $a/\Theta \mapsto \varphi(a)/\varphi(\Theta)$.

Theorem 0.27 "Second Isomorphism Theorem"

Let Θ and Ψ be congruence relations on an algebra $\mathcal{A}$ with $\Psi \geq \Theta$. Define a relation Ψ/Θ on A/Θ by "$a/\Theta \equiv b/\Theta(\Psi/\Theta)$ iff $a \equiv b(\Psi)$.

Then Ψ/Θ is a congruence relation on $\mathcal{A}/\Theta$ and

$$(\mathcal{A}/\Theta)/(\Psi/\Theta) \cong \mathcal{A}/\Psi$$

by the map $(a/\Theta)//(\Psi/\Theta) \mapsto a/\Psi$.