STAR-LINEAR EQUATIONAL THEORIES OF GROUPOIDS

P. ĐAPIĆ, J. JEŽEK, P. MARKOVIĆ, R. MCKENZIE AND D. STANOVSKÝ

Abstract. We prove that there are precisely six equational theories E of groupoids with the property that every term is E-equivalent to a unique linear term.

1. Introduction

By a term we always mean a term in the signature of groupoids (algebras with one binary, multiplicatively denoted operation). A term is said to be *linear* if every variable has at most one occurrence in it.

The equational theory of order algebras, introduced and investigated in [4] and [2], turned out to have the following interesting property: every term in at most three variables is equivalent to precisely one linear term.

By a $*$-*linear equational theory* we mean an equational theory E such that every term t is E-equivalent to a unique linear term, denoted usually t^*. In the present paper, we prove that there are precisely six $*$-linear equational theories of groupoids (Theorem 13.1, see constructions in Sections 8,10,12), find finite equational bases for four of them (Theorems 9.1 and 11.2), prove that the other two are inherently non-finitely based (Theorem 14.7), describe all subvarieties of the six corresponding varieties (Theorem 15.5) and find small generating groupoids for each of them (Theorems 16.1, 16.2). As a corollary, we obtain all (two) equational theories of semigroups such that every word is equivalent to a unique linear word (Theorem 17.3).

Every $*$-linear theory defines a locally finite variety. In fact, the universe of the free algebra on n generators in that variety is bijective with the set of all linear terms over $x_1, \dots, x_n$. On two generators, this means that the algebra has four elements, on three generators 21 elements, on four generators 184.

If E is a $*$-linear equational theory and, for instance, a 21-element groupoid $\mathbf{G}$ on three generators belongs to the corresponding variety $\mathrm{Mod}(E)$, then $\mathbf{G}$ must be (isomorphic to) the free groupoid of rank three in this variety.

Observe that two comparable $*$-linear theories must be identical.

Let $S(t)$ denote the set of variables occurring in a term t and let $|t|$ denote the *length* of t, i.e. the total number of occurrences of variables in t. Clearly, $|S(t)| \leq |t|$ with equality precisely when t is linear.

1991 *Mathematics Subject Classification.* 03C05, 03C13, 08B05, 20N02.

Key words and phrases. equational theory, groupoid, linear term, variety.

While working on this paper the first and the third author were supported by the Ministry of Science, Technologies and Development of Republic of Serbia, grant no. 1227; the second and the fifth author were supported by MŠMT, research project MSM 0021620839, and by the Grant Agency of the Czech Republic, grant #201/02/0594; the fourth author was supported by the NSF grant #DMS-0245622.

In every $*$-linear equational theory $S(t^*) \subseteq S(t)$. Indeed, suppose that there is a variable $x \in S(t^*) - S(t)$. Take a variable y not occurring in t^* and denote by r the term obtained from t^* by substituting y for x. Then $t \approx r$ is a consequence of $t \approx t^*$, and thus $t^* \approx r$ in E. But t^*, r are two different linear terms, a contradiction.

Consequently, $xx \approx x$ in every $*$-linear equational theory.

More generally, by an *n-linear equational theory* (for a positive integer n) we mean an equational theory E such that every term in at most n variables is E-equivalent to precisely one linear term (which must be again in at most n variables). If, moreover, E is generated by its at most n-variable equations, then we say that E is *sharply n-linear*. Of course, such an equational theory is uniquely determined by its n-generated free groupoid.

We say that an equational theory E *extends* a groupoid $\mathbf{G}$, if $\mathbf{G}$ is its free groupoid.

We need also the following concepts: A *dual* of the term t (written t^∂) is defined to be equal to t if t is a variable, and if $t = t_1t_2$, then $t^\partial = t_2^\partial t_1^\partial$. The dual of an equational theory E would then mean the class of all identities $t_1^\partial \approx t_2^\partial$, where $t_1 \approx t_2$ in E. An equation $s \approx t$ is called *regular*, if $S(s) = S(t)$. An equation $s \approx t$ is called *left non-permutational*, if the order of first occurrences of the variables in s, counting from the left, is the same as the order in t. It is called *right non-permutational*, if the dual equation is left non-permutational. An equational theory E is called *regular* (*left, right non-permutational*, resp.), if all equations in E are regular (left, right non-permutational, resp.).

In order to avoid writing too many parentheses in terms, $x_1x_2x_3 \dots x_n$ will stand for $(((x_1x_2)x_3)\dots)x_n$ (the parentheses are grouped to the left), $x \cdot yz$ will stand for $x(yz)$, etc.

For notation and terminology not introduced in the paper we refer to the book [6].

We close the introduction with an admission. We have used a computer program (available at `www.karlin.mff.cuni.cz/~jezek`) as aid in our investigation. It has the following capabilities: While entering the multiplication table of a groupoid, it automatically completes all the consequences of an entry which are implied by a set of equations previously typed in. If an entry is contradictory with the equations, the program informs the user of it, as well as where it happens. Also, when a full multiplication table is entered, the program checks if a given set of equations is satisfied in the groupoid, and finds an evaluation of the variables for which some equation fails.

We have used this program extensively, but later found independent proofs for most of the results. The only place where the reader could be challenged to verify the validity of these results without resorting to the computer check are the Sections 5 and 6. And even in these two Sections, we feel that it is not beyond the ability of a (very) patient reader to manually verify that the given groupoids are precisely the 3-generated free algebras for all sharply 3-linear equational theories which have $\mathbf{G}_6$ as their 2-generated free algebra.

All results obtained with computer aid were checked by an independent computation using the automated theorem prover Otter [5] driven by a Perl script. It took about one minute (on a Pentium PC) to compute all strictly 2-linear theories, about two hours to find their strictly 3-linear extensions and several weeks to prove that only $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\mathbf{Q}_4$ may have a 4-linear extension.

2. Sharply 2-linear equational theories

Clearly, there is precisely one sharply 1-linear equational theory: that of idempotent groupoids. The following lemma is an easy consequence of a result of J. Dudek [1], where all theories of groupoids with two strictly binary terms are classified. We include a proof in order to keep our paper self-contained.

Lemma 2.1. *There are precisely twelve sharply* 2*-linear equational theories. Their* 2*-generated free groupoids are the following seven groupoids, plus their duals.* (*The first two of the seven groupoids are self-dual.*)

$\mathbf{G}_0$	x	y	xy	yx
x	x	xy	yx	y
y	yx	y	x	xy
xy	y	yx	xy	x
yx	xy	x	y	yx

$\mathbf{G}_1$	x	y	xy	yx
x	x	xy	xy	x
y	yx	y	y	yx
xy	x	xy	xy	x
yx	yx	y	y	yx

$\mathbf{G}_2$	x	y	xy	yx
x	x	xy	xy	yx
y	yx	y	xy	yx
xy	x	y	xy	x
yx	x	y	y	yx

$\mathbf{G}_3$	x	y	xy	yx
x	x	xy	xy	yx
y	yx	y	xy	yx
xy	x	y	xy	yx
yx	x	y	xy	yx

$\mathbf{G}_4$	x	y	xy	yx
x	x	xy	x	xy
y	yx	y	yx	y
xy	xy	x	xy	x
yx	y	yx	y	yx

$\mathbf{G}_5$	x	y	xy	yx
x	x	xy	x	xy
y	yx	y	yx	y
xy	xy	xy	xy	xy
yx	yx	yx	yx	yx

$\mathbf{G}_6$	x	y	xy	yx
x	x	xy	xy	xy
y	yx	y	yx	yx
xy	xy	xy	xy	xy
yx	yx	yx	yx	yx

Proof. Denote by $\mathbf{G}$ the two-generated free groupoid in the variety corresponding to a 2-linear equational theory.

Case 1: $xy \cdot x \approx y$. We are going to prove that in this case $\mathbf{G}$ is $\mathbf{G}_0$. We have $x \cdot yx \approx (yx \cdot y) \cdot yx \approx y$. Since $(x \cdot xy)x \approx xy$, $(x \cdot xy)^*$ cannot be any of the terms x, y or xy, so that $x \cdot xy \approx yx$. Since $y(xy \cdot y) \approx xy$, we get similarly $xy \cdot y \approx yx$. Now $xy \cdot yx \approx (y \cdot yx) \cdot yx \approx yx \cdot y \approx x$, so $\mathbf{G}$ is $\mathbf{G}_0$.

Case 2: $xy \cdot x \approx x$. We are going to show that $\mathbf{G}$ is either $\mathbf{G}_1$ or $\mathbf{G}_2$ or $\mathbf{G}_3$. We have $(xy \cdot x) \cdot xy \approx xy$, i.e., $x \cdot xy \approx xy$.

Subcase 2a: $x \cdot yx \approx x$. Then $xy \cdot y \approx xy \cdot (y \cdot xy) \approx xy$. If $xy \cdot yx \approx y$ then $x \approx (yx \cdot x)(x \cdot yx) \approx yx \cdot x \approx yx$, a contradiction. If $xy \cdot yx \approx xy$ then $x \approx xy \cdot x \approx (x \cdot xy)(xy \cdot x) \approx x \cdot xy \approx xy$, a contradiction. If $xy \cdot yx \approx yx$ then $x \approx x \cdot yx \approx (x \cdot yx)(yx \cdot x) \approx yx \cdot x \approx yx$, a contradiction. Hence $xy \cdot yx \approx x$ and we get the groupoid $\mathbf{G}_1$.

Subcase 2b: $x \cdot yx \approx y$. This subcase is not possible by the dual of Case 1.

Subcase 2c: $x \cdot yx \approx xy$. Then $x \approx xx \approx x(xy \cdot x) \approx x \cdot xy \approx xy$, a contradiction.

Subcase 2d: $x \cdot yx \approx yx$. Then $xy \cdot y \approx (y \cdot xy)y \approx y$. If $xy \cdot yx \approx y$ then $x \approx yx \cdot x \approx (x \cdot yx)(yx \cdot x) \approx yx$, a contradiction. If $xy \cdot yx \approx xy$ then $x \approx yx \cdot x \approx (x \cdot yx)(yx \cdot x) \approx x \cdot yx \approx yx$, a contradiction. So, we have either $xy \cdot yx \approx x$ or $xy \cdot yx \approx yx$, i.e., we get either $\mathbf{G}_2$ or $\mathbf{G}_3$.

Case 3: $xy \cdot x \approx yx$. We are going to show that $\mathbf{G}$ is the dual of either $\mathbf{G}_4$ or $\mathbf{G}_5$ or $\mathbf{G}_6$. We have $yx \cdot xy \approx (xy \cdot x) \cdot xy \approx x \cdot xy$. There are four possibilities for $x \cdot xy$:

Subcase 3a: $x \cdot xy \approx x$. Then $x \approx xx \approx (x \cdot xy)x \approx xy \cdot x \approx yx$, a contradiction. This subcase is not possible.

Subcase 3b: $x \cdot xy \approx y$. Then $xy \cdot (xy \cdot y) \approx y$ implies that $(xy \cdot y)^*$ cannot be any of the terms x, xy, yx. Hence $xy \cdot y \approx y$. By the duals of the cases 1 and 2, $(x \cdot yx)^*$ is neither x nor y. If $x \cdot yx \approx xy$ then $yx \approx xy \cdot x \approx (x \cdot yx)x \approx yx \cdot x \approx x$, a contradiction. Hence $x \cdot yx \approx yx$ and we get the dual of $\mathbf{G}_4$.

Subcase 3c: $x \cdot xy \approx xy$. By the duals of the cases 1 and 2, $(x \cdot yx)^*$ is neither x nor y. If $x \cdot yx \approx xy$ then $xy \approx xy \cdot xy \approx xy \cdot (x \cdot xy) \approx xy \cdot x \approx yx$, a contradiction. Hence $x \cdot yx \approx yx$. If $xy \cdot y \approx x$ then $xy \approx xy \cdot xy \approx (x \cdot xy) \cdot xy \approx x$, a contradiction. If $xy \cdot y \approx yx$ then $xy \approx xy \cdot xy \approx (x \cdot xy) \cdot xy \approx xy \cdot x \approx yx$, a contradiction. Hence either $xy \cdot y \approx y$ or $xy \cdot y \approx xy$, i.e., we get the dual of either $\mathbf{G}_5$ or $\mathbf{G}_6$.

Subcase 3d: $x \cdot xy \approx yx$. We have $x \cdot yx \approx x(x \cdot xy) \approx xy \cdot x \approx yx$ and $yx \cdot x \approx x(x \cdot yx) \approx x \cdot yx \approx yx$. Now $xy \approx xy \cdot yx \approx xy \cdot (x \cdot xy) \approx x \cdot xy \approx yx$, a contradiction. This subcase is not possible.

Case 4: $xy \cdot x \approx xy$. We are going to show that $\mathbf{G}$ is either $\mathbf{G}_4$ or $\mathbf{G}_5$ or $\mathbf{G}_6$ or the dual of either $\mathbf{G}_2$ or $\mathbf{G}_3$.

Subcase 4a: $x \cdot yx \approx x$. By the dual of case 2, we get either the dual of $\mathbf{G}_2$ or the dual of $\mathbf{G}_3$.

Subcase 4b: $x \cdot yx \approx y$. This is impossible by the dual of Case 1.

Subcase 4c: $x \cdot yx \approx xy$. By the dual of Case 3 we get either $\mathbf{G}_4$ or $\mathbf{G}_5$ or $\mathbf{G}_6$.

Subcase 4d: $x \cdot yx \approx yx$. We have $xy \cdot (x \cdot xy) \approx x \cdot xy$, so that $(x \cdot xy)^* \neq x$. We have $(x \cdot xy)x \approx x \cdot xy$, so that $(x \cdot xy)^* \neq y$. We have $(xy \cdot yx) \cdot xy \approx xy \cdot yx$, so that $(xy \cdot yx)^*$ cannot be any of the terms x, y, yx, and we get $xy \cdot yx \approx xy$. But quite similarly $xy \cdot yx \approx yx$, a contradiction. This subcase is not possible. □

3. Extending $\mathbf{G}_0$, $\mathbf{G}_1$, $\mathbf{G}_2$, $\mathbf{G}_3$, and $\mathbf{G}_4$

Lemma 3.1. *We cannot have $\mathbf{G}_0$ as the free two-generated groupoid for a 3-linear equational theory.*

Proof. Let E be a 3-linear equational theory extending $\mathbf{G}_0$ and $\ell = (xy \cdot zx)^*$. By the substitutions $y \mapsto x$, $z \mapsto x$ and $z \mapsto y$ we get $\ell(x,x,y) \approx \ell(x,y,x) \approx \ell(y,x,x) \approx y$ in E. Clearly, in such a case, $S(\ell) = \{x,y,z\}$ and each of the 12 possibilities is easily seen unsuitable. □

Lemma 3.2. *We cannot have $\mathbf{G}_1$ as the free two-generated groupoid for a 3-linear equational theory.*

Proof. Suppose that there is a 3-linear equational theory E with the free two-generated groupoid $\mathbf{G}_1$. If E contains an equation with different leftmost variables at both sides, then we can substitute for all the remaining variables one of these two variables, and obtain an equation with the same property in just two variables, which would yield a contradiction. So, every equation of E must have the same leftmost variables and, quite similarly, also the same rightmost variables at both sides. Thus a term both starting and ending with a variable x must be equivalent to a linear term both starting and ending with x, and therefore equivalent to x. So, $x \cdot yz \approx (xz \cdot x)(y(z \cdot xz)) \approx xz$ in E, a contradiction. □

Lemma 3.3. *We cannot have* $\mathbf{G}_2$ *or* $\mathbf{G}_3$ *as the free two-generated groupoid for a* 3*-linear equational theory.*

Proof. Similarly to the previous case, any terms equivalent in a $*$-linear theory extending $\mathbf{G}_2$ or $\mathbf{G}_3$ must have the same rightmost variable. We will establish that, where $t = x \cdot xyz$, t^* cannot be any of the 7 linear terms in x, y, z ending with z. The substitution $y \mapsto x$ shows that $t^* \neq z$ for both of these groupoids. The substitution $z \mapsto y$ eliminates xyz and yxz and the substitution $z \mapsto x$ eliminates yz, $x(yz)$ and $y(xz)$. Finally, in $\mathbf{G}_2$, the possibility $t^* = xz$ is eliminated by $z \mapsto yx$, while in $\mathbf{G}_3$, the same possibility is eliminated by $x \mapsto yz$. $\square$

Lemma 3.4. *We cannot have* $\mathbf{G}_4$ *as the free two-generated groupoid for a* 3*-linear equational theory.*

Proof. Suppose that $\mathbf{G}_4$ serves for a 3-linear equational theory E. Since (as it is easy to check) $\mathbf{G}_4$ satisfies $xy \cdot yz \approx x$ and there is no linear term ℓ except x for which $\mathbf{G}_4$ would satisfy $\ell \approx x$, the equation $xy \cdot yz \approx x$ belongs to E. Then also $(xy \cdot yz) \cdot yz \approx x \cdot yz$ belongs to E. Now $x \approx xy \cdot y$ is a two-variable equation satisfied in $\mathbf{G}_4$, so it must belong to E. Consequently, $xy \approx (xy \cdot yz) \cdot yz$ belongs to E and we get that $xy \approx x \cdot yz$ belongs to E, a contradiction. $\square$

4. Extending $\mathbf{G}_5$

In this section we suppose that $\mathbf{G}_5$ is the two-generated free groupoid for a 4-linear equational theory E. Thus we have in E the equations

$$x \approx xx,$$
$$x \approx x \cdot xy,$$
$$xy \approx xy \cdot x \approx xy \cdot y \approx x \cdot yx \approx xy \cdot yx.$$

and, again, if $u \approx v$ in E, then u, v have the same leftmost variables. We will use the above facts without explicit quotations in this section.

Lemma 4.1. $xy \cdot xz \approx xy$ *in* E.

Proof. Put $u = (xy \cdot xz)^*$, so that u is a linear term starting with x. If either $u = x$ or $u = xz$, we get a contradiction by substitution $z \mapsto x$. If u is either $xz \cdot y$ or $x \cdot zy$, we get a contradiction by substitution $y \mapsto x$. If $u = x \cdot yz$, then $yx \approx (yx \cdot y)(yx \cdot z) \approx yx \cdot yz \approx y \cdot xz$, a contradiction. If $u = xy \cdot z$, then $xy \approx xy \cdot x(xz) \approx xy \cdot xz \approx xy \cdot z$, a contradiction. The only remaining possibility is $u = xy$. $\square$

Lemma 4.2. $xyzx \approx xyz$ *in* E.

Proof. Put $u = (xyzx)^*$. If u is either x or xz or $x \cdot zy$, we get a contradiction by $z \mapsto x$. If u is either xy or $x \cdot yz$, we get a contradiction by $y \mapsto x$. Suppose $u = xzy$. Then $xzy \approx xyzx \approx xyzxx \approx xzyx \approx xyz$, a contradiction. $\square$

Lemma 4.3. $x(yz)z \approx xyz$ *in* E.

Proof. Put $u = (x(yz)z)^*$. If u is either x or xy or $x \cdot yz$, we get a contradiction by $y \mapsto x$. If u is xz, or $x \cdot zy$, we get a contradiction by putting $z \mapsto x$.

By the way of contradiction, assume $u = xzy$ (1), and by substituting z with yz we get

$$x(yz)y \approx_{(1)} x(y(yz))(yz) \approx xy \cdot yz. \, (2)$$

Let $w = (x(yz)y)^* = (xy \cdot yz)^*$. w is not equal to xz, xyz, xzy, or $x \cdot zy$, because of the substitution $x \mapsto y$. Also, $w \neq x$ because of the substitution $y \mapsto z$.

Case 1: $w = x \cdot yz$. Then $xzy \approx_{(1)} x(yz)z \approx wz \approx (xy \cdot yz)z \approx_{(1)} xyzy$, i.e. $xzy \approx xyzy\,(3)$.

Now we consider $t = (xy \cdot zy)^*$. Clearly t is none of x, xy, or $x \cdot yz$ because of the substitution $x \mapsto y$, and it is not equal to xz or $x \cdot zy$ because of the substitution $x \mapsto z$.

Subcase 1a: $t = xzy\,(4)$. Then $xzy \approx xzyy \approx_{(4)} (xy \cdot zy)y \approx_{(1)} xyyz \approx xyz$.

Subcase 1b: $t = xyz\,(5)$. Then $xzy \approx_{(3)} xyzy \approx_{(5)} xy(zy)y \approx_{(2)} xyy(zy) \approx xy \cdot zy \approx_{(5)} xyz$. This proves that $w = x \cdot yz$ and $u = xzy$ are contradictory.

Case 2: $w = xy\,(6)$. Let $v = (x(yz)(zy))^*$. Then $v \neq x$, because of the substitution $y \mapsto z$, $v \neq xy$ and $v \neq x(yz)$ because of the substitution $x \mapsto y$, and $v \neq xz$ and $v \neq x(zy)$ because of the substitution $x \mapsto z$.

Subcase 2a: $v = xzy$. Then $xzy \approx xzyy \approx vy \approx x(yz)(zy)y \approx_{(1)} x(yz)yz \approx_{(6)} xyz$.

Subcase 2b: $v = xyz$. Then $xyz \approx xyzz \approx vz \approx x(yz)(zy)z \approx_{(6)} x(yz)z \approx_{(1)} xzy$. This final contradiction proves that $u = xzy$ is not possible in a $*$-linear variety extending $\mathbf{G}_5$, so the only remaining possibility is $u = xyz$. □

Lemma 4.4. *We cannot have* $\mathbf{G}_5$ *as the free two-generated groupoid for a* 4*-linear equational theory.*

Proof. Let ℓ be the unique linear term equivalent in E to $x(yz)(wz)$. The proof proceeds by showing that whatever the term ℓ is, either the equation $x(yz)(wz) \approx \ell$ fails in $\mathbf{G}_5$, or else together with the two-variable equations from the beginning of this section, this equation yields a nontrivial linear equation, i.e., an equation $s \approx t$ with $s \neq t$ and both s, t linear.

We have $x \in S(\ell)$, since it is the leftmost variable.

We have $y \in S(\ell)$, else substituting $y \mapsto x$ in $x(yz)(wz)$, and also substituting $y \mapsto z$, yields $xz \cdot wz \approx x \cdot wz$ and then substituting $w \mapsto x$ gives $xz \approx x$ — a non-trivial linear equation.

We have $z \in S(\ell)$, else substituting $z \mapsto yz$ in $x(yz)(wz)$ yields $x(yz)(wz) \approx xy(w \cdot yz)$. Then substituting $w \mapsto y$ in this equation yields $x \cdot yz \approx xy$, a non-trivial linear equation.

We have $w \in S(\ell)$, else substituting $w \mapsto z$, and also substituting $w \mapsto x(yz)$, yields $(x \cdot yz)z \approx x \cdot yz$, which becomes $xz \approx x$ after substituting $y \mapsto x$.

Thus $S(\ell) = \{x, y, z, w\}$. Write $\ell = ab$.

Case $a = x$: Here b cannot be one of the terms $y \cdot zw, \dots, w \cdot zy$, i.e., cannot be right-associated. For if it were, then by identifying some two of y, z, w we would obtain one of the equations $xy \cdot wy \approx xy$, $x \cdot yz \approx xy$, $(x \cdot yz)z \approx xz$. The first leads to $xw \approx x$ upon replacing y by x; the second is linear, the third leads to $xy \approx x$ upon replacing z by x. Thus b is one of the left-associated terms $yzw, \dots, wzy$. If b begins with y then the substitution $w \mapsto z, y \mapsto x$ gives $xz \approx x$. If b begins with z, then $w \mapsto y$ gives $x \cdot yz \approx x \cdot zy$, which is linear. If b begins with w, then replacing z by y gives $xy \cdot wy \approx x \cdot wy$, leading to $xy \approx x$ when w is replaced by x.

Case $b = y$: Taking $y \mapsto x$ yields $ax \approx x \cdot wz$. Since a must be one of the terms $xwz, xzw, x \cdot zw, x \cdot wz$, then $ax \approx a$. (See Lemma 4.2 and the equation $xyx \approx xy$ above.) Thus we have $a \approx x \cdot wz$, so a is identically $x \cdot wz$ and the equation

$x(yz)(wz) \approx ab$ is $x(yz)(wz) \approx x(wz)y$. Taking $w \mapsto z$ gives $x(yz)z \approx xzy$. However, this contradicts Lemma 4.3.

Case $b = z$: There are four subcases. $a = x \cdot yw$ is destroyed by taking $w \mapsto y$. $a = x \cdot wy$ is destroyed by taking $w \mapsto z$. $a = xyw$ and $a = xwy$ are destroyed by $y \mapsto x$.

Case $b = w$: There are four subcases. In every one, taking $y \mapsto x$ yields either $xw \approx x \cdot wz$ or $xzw \approx x \cdot wz$.

The only remaining cases are where both a and b have two variables. Thus for some $e \in \{y, z, w\}$, a is xe and b is a product of the two members of $S(\ell) \setminus \{x, e\}$. If $a = xz$, then taking $w \mapsto y$ yields $x \cdot yz \approx xzy$. If $a = xw, b = yz$ then taking $y \mapsto x$ yields $x \cdot wz \approx xw$ (by Lemma 4.1). If $a = xw, b = zy$, then taking $z \mapsto wz, x \mapsto wz$ yields $wzyw \approx wzw \cdot wzy$ which produces $wzy \approx wz$ (using Lemmas 4.1 and 4.2). If $a = xy$, $b = zw$ then $y \mapsto x$ yields $x \cdot zw \approx x \cdot wz$.

Only one possible value of ℓ remains. It is $xy \cdot wz$. Thus, we have $x(yz)(wz) \approx xy \cdot wz$ in E. But then taking $w \mapsto x$ gives $x \cdot yz \approx xy$ by Lemma 4.1. □

Remark 4.5. One can prove that there are precisely nine sharply 3-linear equational theories extending $\mathbf{G}_5$. According to the preceding lemma, none of them can be extended to a 4-linear theory.

Theorem 4.6. *Every $*$-linear theory is regular.*

Proof. According to Lemmas 3.1, 3.2, 3.3, 3.4 and 4.4, no $*$-linear theory extends any of the groupoids $\mathbf{G}_0$, $\mathbf{G}_1$, $\mathbf{G}_2$, $\mathbf{G}_3$, $\mathbf{G}_4$, $\mathbf{G}_5$ or their duals. Therefore, if any exist, they must extend $\mathbf{G}_6$ or its dual. Since both of these are groupoids which satisfy only regular 2-variable equations, it is easy to show that if the 2-variable identities of an idempotent variety are all regular, then this variety satisfies only regular identities. □

5. Extending $\mathbf{G}_6$

Let us define seven 21-element groupoids $\mathbf{Q}_1, \dots, \mathbf{Q}_7$, all with the same underlying set $\{a, b, c, \dots, u\}$, all of them 3-generated ($a = x$, $b = y$, $c = z$, $d = xy$, $e = xz$, $f = yz$, $g = yx$, $h = zx$, $i = zy$, $j = xyz$, $k = yxz$, $l = xzy$, $m = zxy$, $n = yzx$, $o = zyx$, $p = x \cdot yz$, $q = x \cdot zy$, $r = y \cdot xz$, $s = y \cdot zx$, $t = z \cdot xy$, $u = z \cdot yx$), by the multiplication tables below; the multiplication table of $\mathbf{Q}_6$ is obtained from that of $\mathbf{Q}_5$ by setting $ar = dr = p$, $at = et = q$, $bp = gp = r$, $bu = fu = s$, $cq = hq = t$, $cs = is = u$.

$\mathbf{Q}_1$	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
a	a	d	e	d	e	p	d	e	q	p	p	q	q	p	q	p	q	p	p	q	q
b	g	b	f	g	r	f	g	s	f	r	r	r	s	s	s	r	r	r	s	s	s
c	h	i	c	t	h	i	u	h	i	t	u	t	t	u	u	t	t	u	u	t	u
d	d	d	j	d	j	j	d	j	j	j	j	j	j	j	j	j	j	j	j	j	j
e	e	l	e	l	e	l	l	e	l	l	l	l	l	l	l	l	l	l	l	l	l
f	n	f	f	n	n	f	n	n	f	n	n	n	n	n	n	n	n	n	n	n	n
g	g	g	k	g	k	k	g	k	k	k	k	k	k	k	k	k	k	k	k	k	k
h	h	m	h	m	h	m	m	h	m	m	m	m	m	m	m	m	m	m	m	m	m
i	o	i	i	o	o	i	o	o	i	o	o	o	o	o	o	o	o	o	o	o	o
j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j
k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k
l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l
m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p
q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q
r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s
t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u

$\mathbf{Q}_2$	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
a	a	d	e	d	e	p	d	e	q	j	p	l	q	p	q	p	q	p	p	q	q
b	g	b	f	g	r	f	g	s	f	r	k	r	s	n	s	r	r	r	s	s	s
c	h	i	c	t	h	i	u	h	i	t	u	t	m	u	o	t	t	u	u	t	u
d	d	d	j	d	j	j	d	j	j	j	j	j	j	j	j	j	j	j	j	j	j
e	e	l	e	l	e	l	l	e	l	l	l	l	l	l	l	l	l	l	l	l	l
f	n	f	f	n	n	f	n	n	f	n	n	n	n	n	n	n	n	n	n	n	n
g	g	g	k	g	k	k	g	k	k	k	k	k	k	k	k	k	k	k	k	k	k
h	h	m	h	m	h	m	m	h	m	m	m	m	m	m	m	m	m	m	m	m	m
i	o	i	i	o	o	i	o	o	i	o	o	o	o	o	o	o	o	o	o	o	o
j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j
k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k
l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l
m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p
q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q
r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s
t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u

$\mathbf{Q}_3$	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
a	a	d	e	d	e	p	d	e	q	j	j	l	l	p	q	p	q	p	p	q	q
b	g	b	f	g	r	f	g	s	f	k	k	r	s	n	n	r	r	r	s	s	s
c	h	i	c	t	h	i	u	h	i	t	u	m	m	o	o	t	t	u	u	t	u
d	d	d	j	d	j	j	d	j	j	j	j	j	j	j	j	j	j	j	j	j	j
e	e	l	e	l	e	l	l	e	l	l	l	l	l	l	l	l	l	l	l	l	l
f	n	f	f	n	n	f	n	n	f	n	n	n	n	n	n	n	n	n	n	n	n
g	g	g	k	g	k	k	g	k	k	k	k	k	k	k	k	k	k	k	k	k	k
h	h	m	h	m	h	m	m	h	m	m	m	m	m	m	m	m	m	m	m	m	m
i	o	i	i	o	o	i	o	o	i	o	o	o	o	o	o	o	o	o	o	o	o
j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j
k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k
l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l
m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p
q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q
r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s
t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u

$\mathbf{Q}_4$	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
a	a	d	e	d	e	p	d	e	q	j	j	l	l	p	q	p	q	j	p	l	q
b	g	b	f	g	r	f	g	s	f	k	k	r	s	n	n	k	r	r	s	s	n
c	h	i	c	t	h	i	u	h	i	t	u	m	m	o	o	t	m	u	o	t	u
d	d	d	j	d	j	j	d	j	j	j	j	j	j	j	j	j	j	j	j	j	j
e	e	l	e	l	e	l	l	e	l	l	l	l	l	l	l	l	l	l	l	l	l
f	n	f	f	n	n	f	n	n	f	n	n	n	n	n	n	n	n	n	n	n	n
g	g	g	k	g	k	k	g	k	k	k	k	k	k	k	k	k	k	k	k	k	k
h	h	m	h	m	h	m	m	h	m	m	m	m	m	m	m	m	m	m	m	m	m
i	o	i	i	o	o	i	o	o	i	o	o	o	o	o	o	o	o	o	o	o	o
j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j
k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k
l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l
m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p
q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q
r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s
t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u

$\mathbf{Q}_5$	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
a	a	d	e	d	e	p	d	e	q	j	p	l	q	p	q	p	q	j	p	l	q
b	g	b	f	g	r	f	g	s	f	r	k	r	s	n	s	k	r	r	s	s	n
c	h	i	c	t	h	i	u	h	i	t	u	t	m	u	o	t	m	u	o	t	u
d	d	d	j	d	j	p	d	j	j	j	p	j	j	p	j	p	j	j	p	j	j
e	e	l	e	l	e	l	l	e	q	l	l	l	q	l	q	l	q	l	l	l	q
f	n	f	f	n	n	f	n	s	f	n	n	n	s	n	s	n	n	n	s	s	n
g	g	g	k	g	r	k	g	k	k	r	k	r	k	k	k	k	r	r	k	k	k
h	h	m	h	t	h	m	m	h	m	t	m	t	m	m	m	t	m	m	m	t	m
i	o	i	i	o	o	i	u	o	i	o	u	o	o	u	o	o	o	u	o	o	u
j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j
k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k
l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l
m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p
q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q	q
r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s	s
t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u

$\mathbf{Q}_7$	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
a	a	d	e	d	e	p	d	e	q	j	p	l	q	p	q	p	q	p	p	q	q
b	g	b	f	g	r	f	g	s	f	r	k	r	s	n	s	r	r	r	s	s	s
c	h	i	c	t	h	i	u	h	i	t	u	t	m	u	o	t	t	u	u	t	u
d	d	d	j	d	j	p	d	j	j	j	p	j	j	p	j	p	j	p	p	j	j
e	e	l	e	l	e	l	l	e	q	l	l	l	q	l	q	l	q	l	l	q	q
f	n	f	f	n	n	f	n	s	f	n	n	n	s	n	s	n	n	n	s	s	s
g	g	g	k	g	r	k	g	k	k	r	k	r	k	k	k	r	r	r	k	k	k
h	h	m	h	t	h	m	m	h	m	t	m	t	m	m	m	t	t	m	m	t	m
i	o	i	i	o	o	i	u	o	i	o	u	o	o	u	o	o	o	u	u	o	u
j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j	j
k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k	k
l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l
m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m	m
n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
p	p	p	j	p	j	p	p	j	j	j	p	j	j	p	j	p	j	p	p	j	j
q	q	l	q	l	q	l	l	q	q	l	l	l	q	l	q	l	q	l	l	q	q
r	r	r	k	r	r	k	r	k	k	r	k	r	k	k	k	r	r	r	k	k	k
s	n	s	s	n	n	s	n	s	s	n	n	n	s	n	s	n	n	n	s	s	s
t	t	m	t	t	t	m	m	t	m	t	m	t	m	m	m	t	t	m	m	t	m
u	o	u	u	o	o	u	u	o	u	o	u	o	o	u	o	o	o	u	u	o	u

The equational theory corresponding to $\mathbf{Q}_7$ is the dual of the equational theory based on the 3-variable equations of order algebras, described in [4].

Lemma 5.1. *There are precisely seven sharply* 3*-linear equational theories with the* 2*-generated free groupoid isomorphic to* $\mathbf{G}_6$*; their corresponding* 3*-generated groupoids are the groupoids* $\mathbf{Q}_1, \dots, \mathbf{Q}_7$*.*

Proof. Let E be a 3-linear equational theory with the 2-generated free groupoid isomorphic to $\mathbf{G}_6$. For every term t in the variables x, y, z we have $S(t^*) = S(t)$ and the leftmost variables in t and t^* are the same. Hence, if x is the leftmost variable, there are just four candidates for the term t^*, namely, the terms xyz, xzy, $x \cdot yz$ and $x \cdot zy$. Unfortunately, all these four terms are identical on $\mathbf{G}_6$. We are going to distinguish four cases according to the four possible normal forms for the term $xy \cdot yz$.

Case 1: $(xy \cdot yz)^* = x \cdot zy$. By the substitution $y \mapsto yz$ we obtain (using equations of $\mathbf{G}_6$) $x \cdot yz \approx x \cdot zy$, a contradiction. This case is not possible.

Case 2: $(xy \cdot yz)^* = xzy$. Then $xyz \approx xy(xyz) \approx (x \cdot xy)(xy \cdot z) \approx xz \cdot xy$ and hence $xyz \approx (xy \cdot y)z \approx (xy \cdot z)(xy \cdot y) \approx (xz \cdot xy)(xy \cdot y) \approx (xz \cdot y)(xy) \approx (xy \cdot yz) \cdot xy \approx xy \cdot yz \approx xz \cdot y$, a contradiction. This case is not possible.

Case 3: $(xy \cdot yz)^* = xyz$. By running the program (cf. the introduction) we obtain that, after the completion, all products are defined except for the products of a variable with a term containing all the three variables.

Subcase 3a: $(x(yxz))^* = xzy$. A contradiction can be obtained by the substitution $y \mapsto yz$, $z \mapsto y$.

Subcase 3b: $(x(yxz))^* = x \cdot zy$. A contradiction can be obtained by the substitution $y \mapsto yz$.

Subcase 3c: $(x(yxz))^* = x \cdot yz$. With this equation, all products, except $x(xyz)$ (and those obtained by permuting x, y, z), turn out to be defined. If $x(xyz) \approx x \cdot yz$, we obtain the groupoid $\mathbf{Q}_1$. If $x(xyz) \approx xyz$, we obtain the groupoid $\mathbf{Q}_2$. The remaining two possibilities for $(x(xyz))^*$ turn out to be contradictory.

Subcase 3d: $(x(yxz))^* = xyz$. All products except $x(y \cdot xz)$ and $x(y \cdot zx)$ turn out to be defined. If $x(y \cdot xz) \approx x \cdot yz$, we obtain the groupoid $\mathbf{Q}_3$. If $x(y \cdot xz) \approx xyz$, then $x(y \cdot zx) \approx x \cdot yz$ (the other three possibilities for $(x(y \cdot zx))^*$ yield contradictions) and we obtain the groupoid $\mathbf{Q}_4$. The remaining two possibilities for $(x(y \cdot xz))^*$ turn out to be contradictory.

Case 4: $(xy \cdot yz)^* = x \cdot yz$. In that case we have $xy \cdot zy \approx xyz$ and $xy \cdot xz \approx xyz$, since the remaining three possibilities for $xy \cdot zy$ (and also for $xy \cdot xz$) turn out to be contradictory.

Subcase 4a: $(x(y \cdot xz))^* = xyz$. All products except $x(y \cdot zx)$ turn out to be defined. We have $x(y \cdot zx) = x \cdot yz$, since the remaining three possibilities for $x(y \cdot zx)$ turn out to be contradictory. We get the groupoid $\mathbf{Q}_5$.

Subcase 4b: $(x(y \cdot xz))^* = xzy$. This yields a contradiction.

Subcase 4c: $(x(y \cdot xz))^* = x \cdot zy$. This yields a contradiction.

Subcase 4d: $(x(y \cdot xz))^* = x \cdot yz$. Now consider the term $(x \cdot yz)(zy)$. If it is equivalent to either $x \cdot zy$ or xzy, we get a contradiction. If it is equivalent to $x \cdot yz$, we get the groupoid $\mathbf{Q}_6$. Finally, if it is equivalent to xyz, we get the groupoid $\mathbf{Q}_7$. □

Lemma 5.2. *The sharply* 3*-linear equational theories extending* $\mathbf{G}_6$ *are left non-permutational.*

Proof. Easy to check (it is enough to check the lines a, d, j, p in the tables of $\mathbf{Q}_1, \dots, \mathbf{Q}_7$). □

6. Extending $\mathbf{Q}_3$, $\mathbf{Q}_5$, $\mathbf{Q}_6$, and $\mathbf{Q}_7$

Lemma 6.1. *There is no 4-linear equational theory with 3-generated free groupoid isomorphic to either $\mathbf{Q}_3$ or $\mathbf{Q}_5$ or $\mathbf{Q}_6$ or $\mathbf{Q}_7$.*

Proof. If there is such an equational theory, then every term in four variables must be equivalent to a linear term in four variables, and that equation must be satisfied in the 3-generated free groupoid. Now one can check that the term $wxy(x \cdot zw)$ is not equivalent to any linear term for any of the groupoids $\mathbf{Q}_5$, $\mathbf{Q}_6$, $\mathbf{Q}_7$. Also, the term $w(x(ywz))$ is not equivalent to any linear term in the case of $\mathbf{Q}_3$. □

This leaves the groupoids $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\mathbf{Q}_4$ as the only candidates for a 3-generated free groupoid of a $*$-linear equational theory.

7. $*$-linear extensions of $\mathbf{Q}_2$ and $\mathbf{Q}_4$ are unique

Theorem 7.1. *Every $*$-linear theory is left or right non-permutational.*

Proof. For terms s, t, we write $s \sim_\ell t$, iff the equation $s \approx t$ is regular and left non-permutational. The relation $\sim_\ell$ is an equational theory.

We show that a $*$-linear theory E extending $\mathbf{G}_6$ is left non-permutational (the dual case can be proven similarly). Suppose there is an equation $s \approx t$ in E such that $s \not\sim_\ell t$. Thus there is a term t such that $t \not\sim_\ell t^*$. Such a t is not linear, and since $S(t) = S(t^*)$ (by Theorem 4.6), we have that $|t| > |t^*|$ for such a t. Let n be minimal so that there exists t with $|t| = n$ and $t \not\sim_\ell t^*$. Choose a variable x so that x has at least two occurrences in t. Replace all occurrences of variables in t except two chosen occurrences of x by occurrences of distinct new variables, creating a term s. Thus x occurs exactly twice in s and all other variables in $S(s)$ occur exactly once in s. Hence $|s| = n$ and $|S(s)| = n - 1$. If $s \sim_\ell s^*$, then substituting back so that s becomes t, we obtain an equation $t \approx \bar{t}$ in E where $|\bar{t}| = n - 1$. By minimality of n, we have $t \sim_\ell \bar{t} \sim_\ell (\bar{t})^* = t^*$, a contradiction. Consequently, $s \not\sim_\ell s^*$.

Now we can choose variables y, z so that y occurs before z in s^* (counting from the left) and the first occurrence of z in s is to the left of all occurrences of y in s. (We do not know if $x \in \{y, z\}$.) Now in $s \approx s^*$ replace all occurrences of variables other than y, z by x and create an equation $r \approx r'$ in E where $S(r) = S(r') = \{x, y, z\}$, $|r| = n$, $|r'| = n - 1$ and $r \not\sim_\ell r'$. By minimality of n, we have $r' \sim_\ell (r')^*$. Thus $r \not\sim_\ell r' \sim_\ell (r')^* = r^*$, which contradicts Lemma 5.2. □

Theorem 7.2. *Every $*$-linear equational theory is generated by its 4-generated free groupoid.*

Proof. Let E be a $*$-linear equational theory and $\mathbf{F}$ its 4-generated free groupoid. By Theorem 4.6, E is regular and according to Theorem 7.1, we can assume it is left non-permutational (the dual case can be proven similarly). We show that every equation valid in $\mathbf{F}$ belongs to E.

Let $\mathbf{F}$ satisfy $s \approx t$, we can assume that s and t are linear. To get a contradiction, we assume that $s \neq t$.

We claim that the equation $s \approx t$ is regular and left non-permutational. Indeed, suppose that s, say, has a variable x that does not occur in t. Replacing all variables of s, t other than x by a variable $y \neq x$ gives us an equation $p \approx q$, valid in $\mathbf{F}$, where

$p = p(x, y)$ has an occurrence of x while $q = q(y)$ has only the variable y. Since $\mathbf{F}$ is the free algebra on 4 free generators, we have that $p \approx q$ belongs to E. This contradicts our assumption that E is regular. Next, suppose that there are variables $x, y \in S(s) = S(t)$ such that the unique occurrence of x in s is to the left of the occurrence of y, while in t, the unique occurrence of y is to the left of the occurrence of x. Replacing all variables except x, y by a third variable z, we obtain an equation $p \approx q$, valid in $\mathbf{F}$, such that $S(p) = S(q)$ contains $\{x, y\}$ and is contained in $\{x, y, z\}$ and the unique occurrences of x, y in p have x to the left of y, while in q it is y to the left of x. As before, $p \approx q$ must belong to E, and this contradicts our assumption that E is left non-permutational. The claim is proved.

Thus we can write $s = s(x_1, \dots, x_n)$, $t = t(x_1, \dots, x_n)$, where $S(s) = S(t) = \{x_1, \dots, x_n\}$ and the ith occurrence of a variable (from the left) in s (and likewise in t) is of x_i. Finally, we can assume that n is minimal, that is, if $s' \approx t'$ is any equation valid in $\mathbf{F}$ with $|S(s')| < n$ then $s' \approx t'$ belongs to E.

Clearly, $n > 4$ and we have $s = a_s b_s$, $t = a_t b_t$. Suppose first that a_s and a_t do not have the same variables, say $S(a_s) = \{x_1, \dots, x_{i+j}\}$, $S(a_t) = \{x_1, \dots, x_i\}$, $j > 0$. Let x, y, w be distinct variables. Replace all the variables $x_1, \dots, x_i$ by x, replace $x_{i+1}, \dots, x_{i+j}$ by y, and replace the remaining variables by w. We get the equation

$$a_s(x, \dots, x, y, \dots, y) \cdot b_s(w, \dots, w) \approx a_t(x, \dots, x) \cdot b_t(y, \dots, y, w, \dots, w)\,,$$

valid in $\mathbf{F}$. Obviously, we have in $\mathbf{F}$

$$\begin{aligned} a_s(x, \dots, x, y, \dots, y) &\approx xy, \\ b_s(w, \dots, w) &\approx w, \\ a_t(x, \dots, x) &\approx x, \\ b_t(y, \dots, y, w, \dots, w) &\approx yw\,. \end{aligned}$$

Thus the equation $(xy)w \approx x(yw)$ is valid in $\mathbf{F}$. But this three-variable linear equation does not belong to E, so cannot be valid in $\mathbf{F}$. Contradiction.

So we are reduced to the case where, say, $S(a_s) = S(a_t) = \{x_1, \dots, x_i\}$, and $S(b_s) = S(b_t) = \{x_{i+1}, \dots, x_n\}$. There are two subcases. In the first subcase, $a_s \neq a_t$. In this subcase, we replace all variables $x_{i+1}, \dots, x_n$ by a new variable u, obtaining that $a_s u \approx a_t u$ holds in $\mathbf{F}$. In the second subcase, $a_s = a_t$ and $b_s \neq b_t$. In this case, we replace all variables $x_1, \dots, x_i$ by u and obtain that $ub_s \approx ub_t$ holds in $\mathbf{F}$. By minimality of n, we have $i = n - 1$ in the first subcase, and $i = 1$ in the second subcase.

Now in the first subcase, write $a_s = c_s d_s$, $a_t = c_t d_t$. If $S(c_s) \neq S(c_t)$, then the above argument gives that $\mathbf{F}$ satisfies $((xy)w)u \approx (x(yw))u$; again, a contradiction. Now just as above, if $c_s \neq c_t$ then we obtain that $\mathbf{F}$ satisfies $(c_s v)u \approx (c_t v)u$ where v is a new variable. If $c_s = c_t$ then $d_s \neq d_t$ and we get that $(vd_s)u \approx (vd_t)u$ is valid in $\mathbf{F}$. Note that E must contain both the equations $(xu)u \approx xu$ and $(ux)u \approx ux$ (since E extends $\mathbf{G}_6$). Hence, thus substitution $v \mapsto u$ gives that $\mathbf{F}$ satisfies either the linear equation $c_s u \approx c_t u$ with $c_s \neq c_t$, or the linear equation $ud_s \approx ud_t$ with $d_s \neq d_t$. Either way, we have a contradiction to the minimality of n. The argument in the second subcase is analogous, using that $u(ux) \approx ux, u(xu) \approx ux$ belong to E. This concludes our proof. $\square$

Theorem 7.3. *For each of the groupoids $\mathbf{Q}_2$, $\mathbf{Q}_4$, there is at most one $*$-linear theory extending it.*

Proof. Assume that $E \neq E'$ are $*$-linear theories extending $\mathbf{Q} \in \{\mathbf{Q}_2, \mathbf{Q}_4\}$. Then $E \not\subseteq E'$ and by Theorem 7.2, there is a four-variable equation which belongs to E and not to E'. Thus there must be a four-variable term s which is equivalent to a linear term t over E and to a linear term t' over E', where $t \neq t'$. Since E and E' have precisely the same three-variable equations, every three variable equation obtained by a substitution from $t \approx t'$ holds in $\mathbf{Q}$.

By Theorems 4.6 and 7.1, $S(s) = S(t) = S(t')$ is a four-element and we can assume that the variables of t and t' are w, x, y, z and they occur in alphabetical order in both these linear terms.

Suppose that wx is a subterm of t (written $wx \leq t$). Then the three-variable equation $s(x,x,y,z) \approx t(x,x,y,z)$ belongs to E and $s(x,x,y,z)$ is equivalent to a linear term $\ell \in \{xyz, x \cdot yz\}$. Hence $t = \ell(wx, y, z)$. If also $wx \leq t'$, then $s(x,x,y,z)$ is equivalent to the same linear term $\ell(x,y,z)$ in E', and we find that $t' = t$, a contradiction. Thus wx cannot be a subterm of both t and t'. Likewise for xy, yz. But clearly, one of the terms wx, xy, yz is a subterm of t, and one is a subterm of t'.

Case $wx \leq t$, $yz \leq t'$: (This proof also takes care of the symmetric case $yz \leq t$, $wx \leq t'$.) Here, $t \approx t'$ is one of the equations

$$wx \cdot yz \approx w(x(yz)) \quad \text{and} \quad wxyz \approx w(x(yz))$$

(or one obtained by switching left-side and right-side terms in one of these equations). In the first equation, the substitution $z \mapsto y$ yields $wx \cdot y \approx w \cdot xy$, and in the second equation, the substitution $y \mapsto x$ yields $wxxz \approx w(x(xz))$, which is in any theory extending $\mathbf{G}_6$ equivalent to $wxz \approx w \cdot xz$. Both cases thus contradict 3-linearity.

Case $xy \leq t$, $yz \leq t'$: (This proof also takes care of the symmetric case $yz \leq t$, $xy \leq t'$.) Here, $t \approx t'$ is one of the equations

$$w(xy)z \approx w(x(yz)), \quad w(xy)z \approx wx \cdot yz,$$
$$w \cdot xyz \approx wx \cdot yz, \quad w \cdot xyz \approx w(x(yz))\,.$$

In the first equation, the substitution $y \mapsto x$ yields $wxz \approx w \cdot xz$, in the second equation, the substitution $w \mapsto x$ yields $xyz \approx x \cdot yz$, in the third equation, the substitution $z \mapsto y$ yields $w \cdot xy \approx wxy$ (in all cases, use again equations of $\mathbf{G}_6$). All three cases thus contradict 3-linearity. Finally, the substitution $w \mapsto x$ in the last equation yields $x \cdot xyz \approx x \cdot yz$, which is not valid in each of $\mathbf{Q}_2$, $\mathbf{Q}_4$.

Case $xy \leq t$, $wx \leq t'$: (This proof also takes care of the symmetric case $wx \leq t$, $xy \leq t'$.) Since the case $t' = wx \cdot yz$ is already covered under the last case, we are here looking at two possibilities for $t \approx t'$, namely,

$$w(xy)z \approx wxyz \quad \text{and} \quad w \cdot xyz \approx wxyz.$$

In the first equation, the substitution $z \mapsto x$ yields $w(xy)x \approx wxyx$, which is not valid in each of $\mathbf{Q}_2$, $\mathbf{Q}_4$. In the second equation, the substitution $y \mapsto x$ yields $w \cdot xz \approx wxz$. □

8. Extending $\mathbf{Q}_1$

Let X be a countably infinite set of variables. We denote by $\mathbf{T}$ the free groupoid over X, and by $\mathbf{T}'$ its extension by a unit element, denoted by $\emptyset$. Put $S(\emptyset) = \emptyset$, so that $S(t)$ is now defined for all $t \in T'$. The length of $\emptyset$ is 0.

For every subset Y of X we denote by δ_Y the endomorphism of $\mathbf{T}'$ such that $\delta_Y(x)=\emptyset$ for $x\in Y$ and $\delta_Y(x)=x$ for $x\in X-Y$. Clearly, for two subsets Y_1,Y_2 of X we have $\delta_{Y_1}\delta_{Y_2}=\delta_{Y_2}\delta_{Y_1}=\delta_{Y_1\cup Y_2}$. For a subset M of T' put $\delta_M=\delta_Y$, where $Y=\bigcup\{S(t):t\in M\}$; for $t\in T'$ put $\delta_t=\delta_{\{t\}}$.

Denote by L the set of linear terms over X and put $L'=L\cup\{\emptyset\}$. Define a binary operation $\circ$ on L' by $u\circ v=u\cdot\delta_u(v)$. Let $\mathbf{L}=(L,\circ)$ and $\mathbf{L}'=(L',\circ)$.

Lemma 8.1. *Let Y be a subset of X. The restriction of δ_Y to L' is an endomorphism of $\mathbf{L}'$.*

Proof. Let $u,v\in L$. Clearly, δ_Y maps L' into L'. We have

$$\delta_Y(u\circ v)=\delta_Y(u\cdot\delta_u(v))=\delta_Y(u)\cdot\delta_Y\delta_u(v)$$

and

$$\delta_Y(u)\circ\delta_Y(v)=\delta_Y(u)\cdot\delta_{\delta_Y(u)}\delta_Y(v);$$

these terms are equal, since $Y\cup S(u)=S(\delta_Y(u))\cup Y$. □

Denote by ℓ_1 the unique homomorphism of $\mathbf{T}'$ into $\mathbf{L}'$ with $\ell_1(x)=x$ for all $x\in X$.

Lemma 8.2. *Let f be a homomorphism of $\mathbf{T}'$ into $\mathbf{L}'$. Then $f\ell_1(t)=f(t)$ for any $t\in T'$.*

Proof. By induction on the length of t. If $t\in X\cup\{\emptyset\}$, then it follows from $\ell_1(t)=t$. Let $t=uv$ where $u,v\in T$. By the induction assumption, $f\ell_1(u)=f(u)$ and $f\ell_1(v)=f(v)$. We have

$$\begin{aligned}f\ell_1(t)=&f(\ell_1(u)\circ\ell_1(v))=f(\ell_1(u)\cdot\delta_u\ell_1(v))\\=&f\ell_1(u)\circ f\delta_u\ell_1(v)=f(u)\cdot\delta_{f(u)}f\delta_u\ell_1(v)\end{aligned}$$

and

$$f(t)=f(u)\circ f(v)=f(u)\cdot\delta_{f(u)}f(v)=f(u)\cdot\delta_{f(u)}f\ell_1(v),$$

so it is sufficient to show that $\delta_{f(u)}f\delta_u=\delta_{f(u)}f$. But, applying 8.1, these are two homomorphisms of $\mathbf{T}'$ into $\mathbf{L}$ that coincide on $X\cup\{\emptyset\}$. □

Let us denote by $\mathcal{L}_1$ the variety generated by $\mathbf{L}$ and by $\sim_1$ the corresponding equational theory.

Theorem 8.3. *$\sim_1$ is a $*$-linear equational theory extending $\mathbf{Q}_1$. It has a normal form function ℓ_1, i.e., $u\sim_1 v$ if and only if $\ell_1(u)=\ell_1(v)$. The groupoid $\mathbf{L}$ is a free $\mathcal{L}_1$-groupoid over X and the groupoid $\mathbf{L}'$ also belongs to $\mathcal{L}_1$.*

Proof. It follows from 8.2. □

9. A base of equations of the variety $\mathcal{L}_1$

Theorem 9.1. *The variety $\mathcal{L}_1$ has a base consisting of the following three equations:*

(1) $xx\approx x$,
(2) $x\cdot yx\approx xy$ *and*
(3) $x(xyz)\approx x\cdot yz$,

Proof. Denote by E the equational theory based on the equations (1)–(3). Observe that E is contained in $\sim_1$. Let us first list some consequences of (1)–(3):

(4) $xy(yx) \approx xyy$. Indeed, $xy(yx) \approx_{(2)} xy(y(xy)) \approx_{(2)} xyy$.

(5) $x(yxz) \approx x \cdot yz$. Indeed, $x(yxz) \approx_{(3)} x((x \cdot yx)z) \approx_{(2)} x(xyz) \approx_{(3)} x \cdot yz$.

(6) $x \cdot xy \approx xy$. Indeed, $x \cdot xy \approx_{(1)} x(xy \cdot xy) \approx_{(3)} x(y \cdot xy) \approx_{(2)} x \cdot yx \approx_{(2)} xy$.

(7) $x(y \cdot xz) \approx x \cdot yz$. Indeed, $x(y \cdot xz) \approx_{(3)} x(y(yxz)) \approx_{(5)} x(yx \cdot (yxz)) \approx_{(6)} x(yxz) \approx_{(5)} x \cdot yz$.

(8) $xyx \approx xy$. Indeed, $xyx \approx_{(2)} xy(x(xy)) \approx_{(6)} xy(xy) \approx_{(1)} xy$.

(9) $xy(xz) \approx xyz$. Indeed, $xy(xz) \approx_{(5)} xy(xyxz) \approx_{(8)} xy(xyz) \approx_{(6)} xyz$.

(10) $xy \cdot yz \approx xyz$. Indeed, $xy(yz) \approx_{(9)} xy(x \cdot yz) \approx_{(3)} xy(x \cdot xyz) \approx_{(9)} xy(xyz) \approx_{(6)} xyz$.

(11) $xyy \approx y$. Indeed, $xyy \approx_{(4)} xy(yx) \approx_{(10)} xyx \approx_{(8)} xy$.

(12) $x(yz)z \approx x \cdot yz$. Indeed, $x(yz)z \approx_{(10)} x(yz)(yzz) \approx_{(11)} x(yz)(yz) \approx_{(11)} x(yz)$.

(13) $xy(zy) \approx xyz$. Indeed, $xy(zy) \approx_{(10)} xy(y \cdot zy) \approx_{(2)} xy \cdot yz \approx_{(10)} xyz$.

(14) $xyzy \approx xyz$. Indeed, $xyzy \approx_{(13)} xy(zy)y \approx_{(12)} xy \cdot zy \approx_{(13)} xyz$.

We are going to prove by induction on the length of t that $t \approx \ell_1(t)$ belongs to E. If t is a variable (or any linear term), this is clear. Let $t = t_1t_2$. By induction we can assume that t_1, t_2 are both linear. If they have no variable in common, then t is linear and we are done. Take a variable $x \in S(t_1) \cap S(t_2)$.

Let $t_1 \neq x$ and $t_2 \neq x$. Then t_1x is shorter than t, so that $t_1x \approx \ell_1(t_1x) = t_1$ by induction. Similarly, xt_2 is shorter than t and hence $xt_2 \approx \ell_1(xt_2) = x\delta_x(t_2)$. We get $t \approx t_1x \cdot t_2 \approx_{(10)} t_1x \cdot xt_2 \approx t_1x \cdot x\delta_x(t_2) \approx_{(10)} t_1x \cdot \delta_x(t_2) \approx t_1 \cdot \delta_x(t_2)$. The term $t_1 \cdot \delta_x(t_2)$ is shorter than t, so that $t_1 \cdot \delta_x(t_2) \approx \ell_1(t_1 \cdot \delta_x(t_2)) = \ell_1(t)$ and we get $t \approx \ell_1(t)$.

Let $t_1 = x$. If $t_2 = x$, use (1). Otherwise, we can write $t_2 = t_{21}t_{22}$. If $t_{21} = x$ then $t = x(x \cdot t_{22}) \approx_{(6)} xt_{22} \approx \ell_1(xt_{22}) = \ell_1(t)$. If $x \in S(t_{21})$ and $t_{21} \neq x$ then $t = x \cdot t_{21}t_{22} \approx_{(3)} x(xt_{21} \cdot t_{22}) \approx x(x\delta_x(t_{21}) \cdot t_{22}) \approx_{(3)} x(\delta_x(t_{21})t_{22}) \approx \ell_1(x(\delta_x(t_{21})t_{22})) = \ell_1(t)$. If $t_{22} = x$ then $t = x(t_{21}x) \approx_{(2)} xt_{21} \approx \ell_1(xt_{21}) = \ell_1(t)$. If $x \in S(t_{22})$ and $t_{22} \neq x$ then $t = x \cdot t_{21}t_{22} \approx_{(7)} x(t_{21} \cdot xt_{22}) \approx x(t_{21} \cdot x\delta_x(t_{22})) \approx_{(7)} x(t_{21}\delta_x(t_{22})) \approx \ell_1(x(t_{21}\delta_x(t_{22}))) = \ell_1(t)$.

Let $t_1 \neq x$ and $t_2 = x$. Write $t_1 = t_{11}t_{12}$. If $x \in S(t_{11})$ then $t = t_{11}t_{12} \cdot x \approx (t_{11}x \cdot t_{12})x \approx_{(14)} t_{11}x \cdot t_{12} \approx t_{11}t_{12} = t_1 = \ell_1(t)$. If $x \in S(t_{12})$ then $t = t_{11}t_{12} \cdot x \approx (t_{11} \cdot t_{12}x)x \approx_{(12)} t_{11} \cdot t_{12}x \approx t_{11}t_{12} = t_1 = \ell_1(t)$. $\square$

Corollary 9.2. *There is exactly one $*$-linear theory extending the groupoid $\mathbf{Q}_1$.*

Proof. Since the equational theory $\sim_1$ has a base consisting of equations in three variables, any $*$-linear theory extending $\mathbf{Q}_1$ must contain $\sim_1$. Hence, it must coincide with $\sim_1$. $\square$

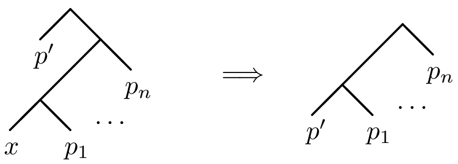
10. Extending $\mathbf{Q}_2$

Let t be a non-linear term and consider a variable x occurring more than once in t. For $i \geq 2$, we denote $p_{x,i}$ the subterm of t of the form

$$p_{x,i} = p'(xp_1p_2 \dots p_n),$$

where the occurrence of x above is the i-th one in t, n is a non-negative number (if $n = 0$ then $p = p'x$) and $p', p_1, \dots, p_n$ are terms.

Let $\sim_2$ be the equivalence on the free groupoid $\mathbf{T}$ generated by all pairs $(t, t^{x,i})$, where t is a term, x is a variable occurring at least i-times in t, $i \geq 2$, and $t^{x,i}$ is the term obtained from t by replacing $p_{x,i} = p'(xp_1p_2 \dots p_n)$ with $p'p_1p_2 \dots p_n$.



Theorem 10.1. $\sim_2$ *is a $*$-linear equational theory extending* $\mathbf{Q}_2$.

Proof. We first prove that there is a unique linear term $\ell_2(t)$ with $t \sim_2 \ell_2(t)$. To do this it is sufficient to prove that $(t^{x,i})^{y,j} = (t^{y,j})^{x,i}$ and that $(t^{x,i})^{x,j-1} = (t^{x,j})^{x,i}$ for $2 \leq i < j$. Let $p = p_{x,i} = p'(xp_1p_2 \dots p_n)$ and $q = p_{y,j} = q'(yq_1q_2 \dots q_m)$. If neither of p and q is a subterm of the other, then it is clear. So, let q be a subterm of p. Again, we have no problems if q is a subterm of p' or of one of the p_is. The remaining case is if the jth occurrence of the variable y is the leftmost variable of a subterm p_i. Then we have that $q = xp_1p_2 \dots p_{i-1}$ and $p_i = yq_1q_2 \dots q_m$. Now, by the definition, the term p gets replaced by $p'p_1p_2 \dots p_{i-1}q_1q_2 \dots q_mp_{i+1} \dots p_n$ in both of the terms $(t^{x,i})^{y,j}$ and $(t^{y,j})^{x,i}$, so those two terms are equal. The other case, $(t^{x,i})^{x,j-1} = (t^{x,j})^{x,i}$ for $2 \leq i < j$, is dealt with analogously. Therefore we have proved that we can transpose the order in which we cancel two different occurrences of variables, so we get that, no matter what order we cancel the occurrences in, we obtain the same linear term.

Now we see that the set of linear terms is a transversal of the equivalence $\sim_2$, so two terms t_1 and t_2 are equivalent modulo $\sim_2$ iff $\ell_2(t_1) = \ell_2(t_2)$. It is easy to see that $\sim_2$ is a congruence of the term algebra.

Finally, we need to show that $\sim_2$ is fully invariant. Let $t(x, y_1, \dots, y_k)$ and p be terms. It is sufficient to show that, if t' is the term obtained from $\ell_2(t)$ by substituting a variable x with p, then $\ell_2(t(p, y_1, \dots, y_k)) = \ell_2(t')$. Let y be the leftmost variable of p. We consider an occurrence of the subterm p in $t(p, y_1, \dots, y_k)$ obtained from the substitution of an occurrence of x in t which is not the leftmost one. Then each occurrence of any variable z of p within this subterm is not the leftmost occurrence of z in $t(p, y_1, \dots, y_k)$ (as at least one copy of the whole p lies left of it), so it can be cancelled. We cancel first all the occurrences of variables of p in this subterm, except for the leftmost occurrence of y. The parentheses were affected only within the subterm, so we can replace the whole occurrence of the subterm p with the variable y. Working this way, we reduce $t(p, y_1, \dots, y_k)$ to a term t'' obtained from t by replacing the leftmost occurrence of x with p, while all the other occurrences of x get replaced by y. Now, all of these occurrences of y which replace x in t'' are not the leftmost ones, since y is a variable that occurs in p. Therefore, all of them get cancelled in the precisely same way as the corresponding occurrences of x get cancelled in t when we reduce t to $\ell_2(t)$. Finally, we have obtained t' from t''.

We have proved that $\sim_2$ is a $*$-linear equational theory. Clearly, $\mathbf{Q}_2$ is its 3-generated free groupoid. $\square$

We denote the corresponding variety $\mathcal{L}_2$.

11. A base of equations of the variety $\mathcal{L}_2$

Lemma 11.1. *The variety $\mathcal{L}_2$ has a base consisting of the at most 3-variable equations that are given by the multiplication table of $\mathbf{Q}_2$, together with the equations*

$$xy(xzu) \approx xyzu \quad \text{and} \quad xy(yzu) \approx xyzu.$$

Proof. Denote this set of equations by S.

Claim 1. $S \vdash (yx)(xy_1y_2\dots y_n) \approx yxy_1y_2\dots y_n$. By using the identity $x(xyz) \approx xyz$ $(n-2)$ times, we transform the left hand side to

$$(yx)(x(x\dots(x(xy_1y_2)y_3)\dots)y_{n-1})y_n).$$

Then we use the identity $(xy)(yzu) \approx xyzu$ $(n-1)$ times to transform this expression to the right hand side. (Note that for $n \le 1$ this proof does not work, but these are just the identities $xyy \approx xy$ and $yx(xz) \approx yxz$.)

Claim 2. $S \vdash x(yy_1y_2\dots y_n) \approx x(yxy_1y_2\dots y_n)$. Again, using the identity $x(xyz) \approx xyz$ $(n-1)$ times, we transform the left hand side to

$$x(y((y((y\dots(y((yy_1)y_2))\dots)y_{n-1}))y_n)).$$

Then, because of the identity $x(yz) \approx x(yxz)$, this expression becomes

$$x((yx)((y((y\dots(y((yy_1)y_2))\dots)y_{n-1}))y_n)).$$

Finally, using the identity $(xy)(xzu) \approx xyzu$ $(n-1)$ times, we transform this expression to the right hand side. (Again, note that for $n=0$ this proof won't work, but that this is just the identity $x(yx) \approx xy$.)

Claim 3. $S \vdash y_1(y_2\dots(y_{n-1}(y_nx))\dots) \approx (y_1(y_2\dots(y_{n-1}(y_nx))\dots))x$. We use the identity $x(yz) \approx (x(yz))z$ $(n-1)$ times to transform the left hand side to

$$y_1(y_2\dots y_{n-2}(y_{n-1}(y_nx)x)\dots)x)x,$$

and then the same identity $(n-2)$ times to obtain the right hand side from the above expression.

Claim 4. *Let t be a term and let an occurrence of the variable x lie immediately to the left of an occurrence of the variable y in t. Let t' be the term obtained from t by replacing this occurrence of y by yx. Then $S \vdash t \approx t'$.* In general, this means that t has a subterm of the form

$$(p_1(p_2\dots(p_{n-1}(p_nx))\dots))((\dots(yq_1)q_2\dots)q_m),$$

where $n, m \ge 0$, and x and y are the occurrences in question. In particular, $n = m = 0$ means that we have xy as a subterm in this place. We obtain from this subterm

$$((p_1(p_2\dots(p_{n-1}(p_nx))\dots))x)((\dots(yq_1)q_2\dots)q_m),$$

by Claim 3, then

$$((p_1(p_2\dots(p_{n-1}(p_nx))\dots))x)(x((\dots(yq_1)q_2\dots)q_m)),$$

by the identity $(xy)z \approx (xy)(yz)$, and

$$((p_1(p_2\dots(p_{n-1}(p_nx))\dots))x)(x((\dots((yx)q_1)q_2\dots)q_m)),$$

by Claim 2. We finish by again using $(xy)z \approx (xy)(yz)$ and Claim 3 to cancel the two occurrences of x in the middle and get

$$(p_1(p_2\dots(p_{n-1}(p_nx))\dots))((\dots((yx)q_1)q_2\dots)q_m),$$

which proves our Claim.

Claim 5. *Let t be a term and let an occurrence of the variable x lie to the left of an occurrence of the variable y in t. Let t' be the term obtained from t by replacing this occurrence of y by yx. Then $S \vdash t \approx t'$.* We do this by an induction on k, the number of occurrences of variables which lie between the occurrences of x and of y in question. For $k = 0$, this is precisely the Claim 4. Otherwise, let $k > 0$ and assume the Claim is proved for $k-1$. Let an occurrence of the variable z lie in t immediately to the left of the occurrence of y we are considering. Let t'' be the term obtained from t by replacing this occurrence of z by zx and t''' the term obtained from t by replacing both of the considered occurrences of y and z by yx and zx respectively. Then $t \approx t''$ by the induction hypothesis, $t'' \approx t'''$ by Claim 4, and $t''' \approx t'$ by the induction hypothesis.

We now finish the proof that S is a base of equations for $\mathcal{L}_2$. Let t be a term in which x occurs more than once and $p_{x,i} = p'(xp_1p_2 \dots p_n)$, for some $i \geq 2$, be the subterm of t from our definition of $\sim_2$. Let y be the rightmost variable in p' and let p'' be the term obtained from p' by replacing this rightmost occurrence of y with yx. Then, since $i \geq 2$, there must exist an occurrence of x in t to the left of the considered occurrence of y in p', or at worst $y = x$. In both cases, t can be transformed to the term where p' is replaced by p'', in the first case by Claim 5, and in the second by idempotence. Furthermore, by Claim 3, Claim 1 and Claim 3,

$$S \vdash p''(xp_1p_2 \dots p_n) \approx (p''x)(xp_1p_2 \dots p_n) \approx p''xp_1p_2 \dots p_n \approx p''p_1p_2 \dots p_n$$

and, finally, by Claim 5, or the idempotence, we can replace p'' by p'. □

Theorem 11.2. *The variety $\mathcal{L}_2$ has a base consisting of the following four equations:*

(1) $xx \approx x$,
(2) $x(yx) \approx xy$,
(3) $x(yxz) \approx x(yz)$ *and*
(4) $xy(yzu) \approx xyzu$.

Proof. By a careful analysis of the proof of Lemma 11.1, we see that the identities actually used are the above four, together with these five: (5) $x(yz)z \approx x(yz)$, (6) $xyy \approx xy$, (7) $xy(yz) \approx xyz$, (8) $x(xyz) \approx xyz$ and (9) $xy(xzu) \approx xyzu$. So, we need to prove them from (1)–(4).

For (7), $xy(yz) \approx_{(3)} xy(y(xy)z \approx_{(4)} xy(xy)z \approx_{(1)} xyz$.

For (8), $x \cdot xyz \approx_{(1)} xx \cdot xyz \approx_{(4)} xxyz \approx_{(1)} xyz$.

For (6), $xyy \approx_{(2)} xy(y(xy)) \approx_{(2)} xy(yx) \approx_{(7)} xyx \approx_{(8)} x(xyx) \approx_{(2)} x(xy) \approx_{(1)} x(xxy) \approx_{(8)} xxy \approx_{(1)} xy$.

For (5), $x(yz)z \approx_{(7)} x(yz)(yzz) \approx_{(6)} x(yz)(yz) \approx_{(6)} x(yz)$.

Now, we prove that (10) $xyx \approx xy$. Indeed, $xyx \approx_{(2)} x(yx)x \approx_{(5)} x(yx) \approx_{(2)} xy$.

Finally, for (9), $xy(xzu) \approx_{(10)} xyx(xzu) \approx_{(4)} xyxzu \approx_{(10)} xyzu$. □

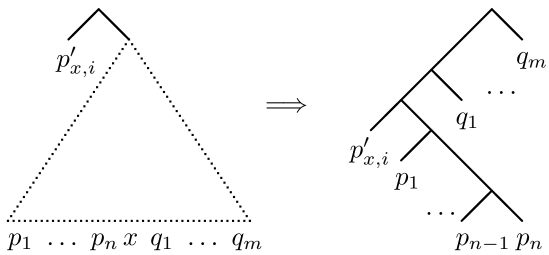
12. Extending $\mathbf{Q}_4$

We start with a technical definition. For a term t, we define inductively the *left* and the *right sequence* corresponding to an occurrence of a variable in t. If t is itself a variable, both sequences are empty. Let $t = t_1t_2$ and assume the occurrence is in t_1. Then the left sequence for t is exactly that for t_1, while the right sequence is $q_1, \dots, q_n, t_2$, where $q_1, \dots q_n$ is the right sequence for the occurrence in t_1. Analogously, assume the occurrence is in t_2. Then the left sequence for t is

$t_1, p_1, \dots, p_n$, where $p_1, \dots p_n$ is the left sequence for the occurrence in t_2, while the right sequence for t is exactly that for t_2.

Let t be a non-linear term and consider a variable x occurring more than once in t. For $i \geq 2$, we denote $p_{x,i} = p'_{x,i}p''_{x,i}$ the subterm of t such that $p'_{x,i}$ contains the $(i-1)$-th occurrence of the variable x in t and $p''_{x,i}$ contains the i-th occurrence of x in t.

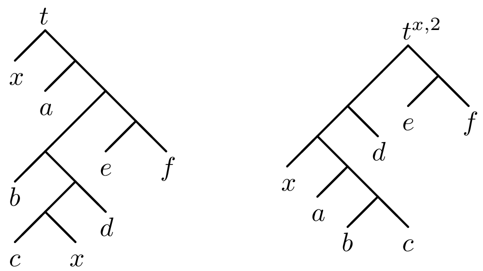
Let $\sim_3$ be the equivalence on the free groupoid $\mathbf{T}$ generated by all pairs $(t, t^{x,i})$, where t is a term, x is a variable occurring at least i-times in t, $i \geq 2$, and $t^{x,i}$ is the term obtained from t by replacing $p_{x,i}$ with $(p'_{x,i}(p_1(p_2(\dots(p_{n-1}p_n)))))q_1q_2\dots q_m$, where $p_1, \dots, p_n$ is the left sequence of the first occurrence of x in $p''_{x,i}$ and $q_1, \dots, q_m$ is the right sequence of this occurrence in $p''_{x,i}$.



In the present section, we adopt a less formal notation. $\{q_1q_2q_3\dots q_\omega\}$ will stand for the bracketing $(((q_1q_2)q_3)\dots)q_\omega$, while $[q_1q_2\dots q_\omega]$ will denote the bracketing $q_1(q_2(\dots(q_{\omega-1}q_\omega)))$. In this notation, the term $p''_{x,i}$ can be written as

$$\{[p_1 \dots \{[p_{\alpha+1} \dots p_\beta\{xp_{\beta+1} \dots p_\gamma\}]p_{\gamma+1} \dots p_\delta\} \dots p_\omega\}.$$

(It means that $p_1, \dots, p_\beta$ is the left sequence for the occurrence of x in $p''_{x,i}$ and $p_{\beta+1}, \dots, p_\omega$ is the right sequence.) So $t^{x,i}$ is obtained from t by replacing the subterm $p_{x,i}$ with $\{[p'_{x,i}p_1\dots p_\beta]p_{\beta+1}\dots p_\omega\}$. An example illustrating this definition is pictured below.



First, we prove that for every term t there is a unique linear term $\ell_3(t)$ equivalent to t modulo $\sim_3$ (clearly, there exists some).

Lemma 12.1. $(p^{x,i})^{x,j-1} = (p^{x,j})^{x,i}$ *for* $2 \leq i < j$.

Proof. Let

$$p_{x,i} = p'_{x,i}\{[p_1 \dots \{[p_{\alpha+1} \dots p_\beta\{xp_{\beta+1} \dots p_\gamma\}]p_{\gamma+1} \dots p_\delta\} \dots p_\omega\} \text{ and}$$
$$p_{x,j} = p'_{x,j}\{[q_1 \dots \{[q_{\alpha'+1} \dots q_{\beta'}\{xq_{\beta'+1} \dots q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\delta'}\} \dots q_{\omega'}\}.$$

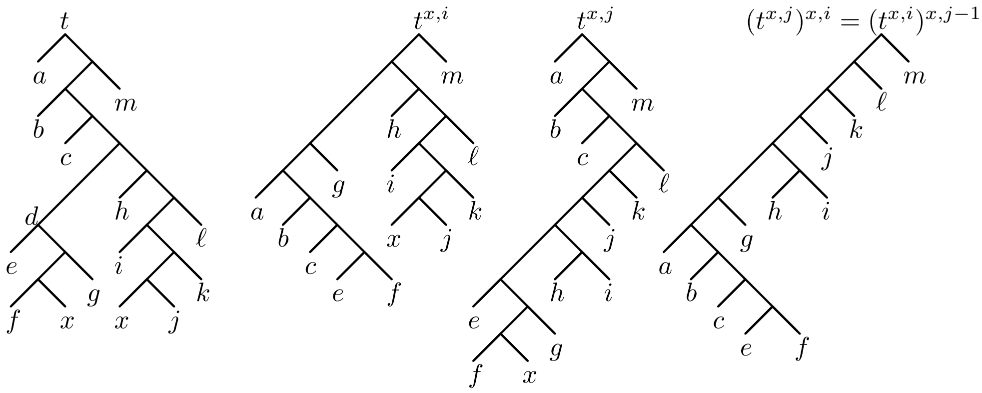
In the case when neither of these two terms is a subterm of the other one, the lemma is easy to prove.

If the term $p_{x,i}$ is a subterm of $p_{x,j}$, then $p_{x,i}$ must be a subterm of $p'_{x,j}$ because the terms $q_1, \dots, q_{\beta'}$ do not contain an occurrence of the variable x. The lemma is again easy to prove.

It remains to consider the case when $p_{x,j}$ is a subterm of $p_{x,i}$. This case contains two subcases.

First subcase: $p_{x,j}$ is a subterm of one of the terms $p_{\beta+1}, \dots, p_\omega$ (because j-th occurrence of x is located to the right from the i-th occurrence of x). This case is easy, too.

Second subcase: $p_{x,j}$ is not a subterm of any of the terms $p_{\beta+1}, \dots, p_\omega$. (It may be helpful to consider the following example, where i-th and j-th occurrences of x are indicated, the $(i-1)$-th occurrence is contained in a and the $(j-1)$-th occurrence is contained in d.)



Then the j-th occurrence of x in p is in a term p_δ, where $\beta+1 \le \delta \le \omega$. Then $p_\delta = p''_{x,j}$ and $p'_{x,j}$ is the largest subterm of $p''_{x,i}$ which does not contain the occurrence of the term p_δ we took for $p''_{x,j}$, and does contain the i-th occurrence of x in p, i.e.

$$p'_{x,j} = \{[p_{\alpha+1}\dots p_\beta\{xp_{\beta+1}\dots p_\gamma\}]p_{\gamma+1}\dots p_{\delta-1}\}.$$

Then

$$p_{x,i} = p'_{x,i}\{[p_1\dots p_\alpha(p'_{x,j}\{^1 q_1\dots[q_{\alpha'+1}\dots q_{\beta'}\{xq_{\beta'+1}\dots q_{\gamma'}\}]\dots q_{\omega'}\}^1)p_{\delta+1}\dots p_\omega\}.$$

The term $p^{x,j}$ is obtained from p by replacing $p_{x,i}$ with

$$p' = p'_{x,i}\{[p_1\dots p_\alpha\{^1[p'_{x,j}q_1\dots q_{\alpha'+1}\dots q_{\beta'}]q_{\beta'+1}\dots q_{\gamma'}\dots q_{\omega'}\}^1 p_{\delta+1}\dots p_\omega\}.$$

Since $p'_{x,j} = \{[p_{\alpha+1}\dots p_\beta\{xp_{\beta+1}\dots p_\gamma\}]p_{\gamma+1}\dots p_{\delta-1}\}$, it follows that

$$p' = p'_{x,i}\{[p_1\dots p_\alpha\{^1[^2\{^3[p_{\alpha+1}\dots p_\beta\{xp_{\beta+1}\dots p_\gamma\}]p_{\gamma+1}\dots p_{\delta-1}\}^3$$
$$q_1\dots q_{\beta'}]^2 q_{\beta'+1}\dots q_{\omega'}\}^1 p_{\delta+1}\dots p_\omega\}.$$

The term $(p^{x,j})^{x,i}$ is obtained from $p^{x,j}$ by replacing p' with the term

$$p'' = \{[p'_{x,i}p_1\dots p_\beta]p_{\beta+1}\dots p_{\delta-1}[q_1\dots q_{\beta'}]q_{\beta'+1}\dots q_{\omega'}p_{\delta+1}\dots p_\omega\}.$$

On the other hand, the term $p^{x,i}$ is obtained from the term p by replacing $p_{x,i}$ with

$$p''' = \{\{[p'_{x,i}p_1\dots p_\beta]p_{\beta+1}\dots p_\gamma\dots p_{\delta-1}\}p_\delta\dots p_\omega\},$$

so $(p^{x,i})'_{x,j-1} = \{[p'_{x,i}p_1\dots p_\beta]p_{\beta+1}\dots p_\gamma\dots p_{\delta-1}\}$, and as p_δ is given above, it follows that

$$p''' = \{(p^{x,i})'_{x,j-1}\{^1[q_1\dots[q_{\alpha'+1}\dots q_{\beta'}\{xq_{\beta'+1}\dots q_{\gamma'}\}]q_{\gamma'+1}\dots q_{\omega'}\}^1 p_{\delta+1}\dots p_\omega\}.$$

The term $(p^{x,i})^{x,j-1}$ is obtained from $p^{x,i}$ by replacing p''' with

$$p'''' = \{[(p^{x,i})'_{x,j-1}q_1\dots q_{\alpha'+1}\dots q_{\beta'}]q_{\beta'+1}\dots q_{\gamma'}\dots q_{\omega'}p_{\delta+1}\dots p_\omega\},$$

which can be written as

$$p'''' = \{(p^{x,i})'_{x,j-1}[q_1\dots q_{\beta'}]q_{\beta'+1}\dots q_{\omega'}q_{\delta'+1}\dots p_\omega\},$$

ie. when we replace $(p^{x,i})'_{x,j-1}$, we get

$$p'''' = \{[p'_{x,i}p_1\dots p_\beta]p_{\beta+1}\dots p_\gamma\dots p_{\delta-1}[q_1\dots q_{\beta'}]q_{\beta'+1}\dots q_{\omega'}p_{\delta+1}\dots p_\omega\}.$$

From above it follows that $p'' = p''''$, ie. that $(p^{x,i})^{x,j-1} = (p^{x,j})^{x,i}$. □

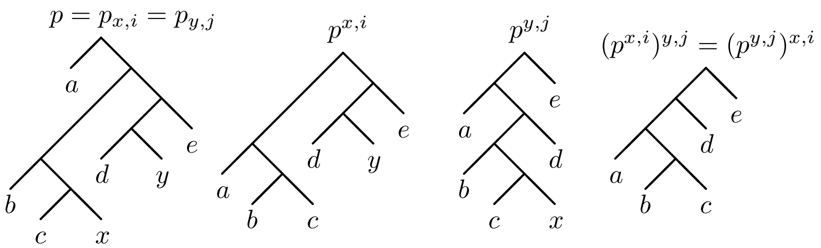
Lemma 12.2. $(p^{x,i})^{y,j} = (p^{y,j})^{x,i}$.

Proof. Without loss of generality, the i-th occurrence of x is located before (to the right of) j-th occurrence of y in the term p. Let

$$\begin{aligned} p_{x,i} &= p'_{x,i}\{[p_1 \dots \{[p_{\alpha+1} \dots p_\beta\{xp_{\beta+1} \dots p_\gamma\}]p_{\gamma+1} \dots p_\delta\} \dots p_\omega\} \\ p_{y,j} &= p'_{y,j}\{[q_1 \dots \{[q_{\alpha'+1} \dots q_{\beta'}\{yq_{\beta'+1} \dots q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\delta'}\} \dots q_{\omega'}\}. \end{aligned}$$

If the subterms $p_{x,i}$ and $p_{y,j}$ are not subterms of each other, or if $p_{x,i}$ is a subterm of $p'_{y,j}$ or of $q_{\lambda'}$ for some $1 \le \lambda' \le \omega'$, or $p_{y,j}$ is a subterm of $p'_{x,i}$ or of p_λ for some $1 \le \lambda \le \omega$, then the lemma is clearly true. Otherwise, consider the following cases.

First case: Let $p_{x,i} = p_{y,j}$. (On the following picture, the previous occurrence of x and y is contained in a.)



Then $p'_{x,i} = p'_{y,j}$, $p''_{x,i} = p''_{y,j}$, and the j-th occurrence of y in p is in the subterm p_λ for some $\beta + 1 \le \lambda \le \omega$ (since i-th x occurs before j-th y). Then p_λ is a subterm of $p''_{y,j}$ and $p_\lambda = \{[q_{\kappa'} \dots \{[q_{\alpha'+1} \dots q_{\beta'}\{yq_{\beta'+1} \dots q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\mu'}\}$, and also $q_{\kappa'-1}$ is a subterm of $p''_{y,j}$ such that $q_{\kappa'-1}$ multiplies p_λ from the left and $q_{\kappa'-1} = \{[p_\kappa \dots \{[p_{\alpha+1} \dots p_\beta\{xp_{\beta+1} \dots p_\gamma\}]p_{\gamma+1} \dots p_{\lambda-1}\}$. It follows that

$$p_{x,i} = p_{y,j} = p'_{x,i}\{[p_1 \dots [^1 p_\rho \dots p_{\kappa-1}(q_{\kappa'-1}p_\lambda)]p_{\lambda+1} \dots p_\eta\}]^1 \dots p_\omega\},$$

where $q_{\kappa'-1}$ contains the i-th occurrence of x and p_λ the j-th occurrence of y in p.

Then the term $p^{x,i}$ is obtained from p by replacing $p_{x,i}$ with the term

$$p' = \{[p'_{x,i}p_1 \dots p_{\kappa-1}p_\kappa \dots p_\beta]p_{\beta+1} \dots p_{\lambda-1}p_\lambda p_{\lambda+1} \dots p_\omega\}, \text{ i.e.}$$

$$\begin{gathered} p' = \{[p'_{x,i}p_1 \dots p_{\kappa-1}p_\kappa \dots p_\beta]p_{\beta+1} \dots p_{\lambda-1} \\ \{^1[q_{\kappa'} \dots \{[q_{\alpha'+1} \dots q_{\beta'}\{yq_{\beta'+1} \dots q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\mu'}\}^1 p_{\lambda+1} \dots p_\omega\}. \end{gathered}$$

The term $(p^{x,i})^{y,j}$ is obtained by replacing p' in $p^{x,i}$ with the term p'' which is equal to

$$\begin{aligned} &\{[\{[p'_{x,i}p_1 \dots p_\beta]p_{\beta+1} \dots p_{\lambda-1}\}q_{\kappa'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\gamma'} \dots q_{\mu'}p_{\lambda+1} \dots p_\omega\} \\ &= \{[p'_{x,i}p_1 \dots p_\beta]p_{\beta+1} \dots p_{\lambda-1}[q_{\kappa'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\gamma'} \dots q_{\mu'}p_{\lambda+1} \dots p_\omega\}. \end{aligned}$$

The term $p^{y,j}$ is obtained from p by replacing the subterm $p_{y,j}$ with the term p''' which is equal to

$$\{[p'_{x,i}p_1 \dots p_{\kappa-1}q_{\kappa'-1}q_{\kappa'} \dots q_{\alpha'+1} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\gamma'} \dots q_{\mu'}p_{\lambda+1} \dots p_\eta \dots p_\omega\}.$$

By replacing $q_{\kappa'-1}$, from above we get

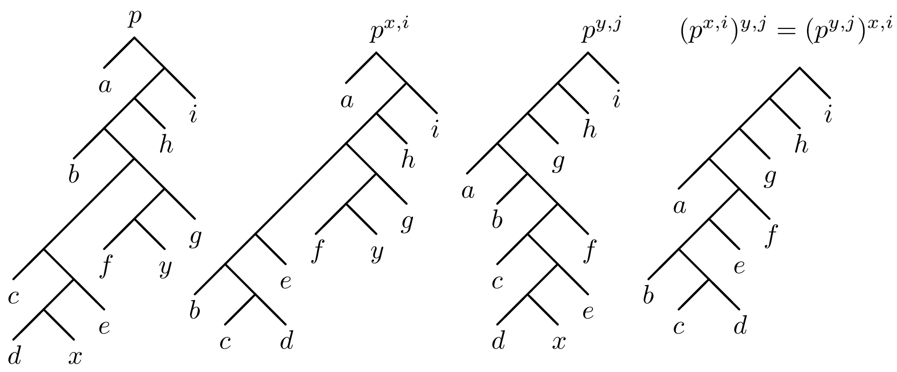
$$\begin{gathered} p''' = \{[p'_{x,i}p_1 \dots p_{\kappa-1}\{^1[p_\kappa \dots \{[p_{\alpha+1} \dots p_\beta\{xp_{\beta+1} \dots p_\gamma\}]p_{\gamma+1} \dots p_{\lambda-1}\}^1 \\ q_{\kappa'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\mu'}p_{\lambda+1} \dots p_\omega\}. \end{gathered}$$

The term $(p^{y,j})^{x,i}$ is obtained from $p^{y,j}$ by replacing p''' with p'''', which equals

$$\{[p'_{x,i}p_1 \dots p_\kappa \dots p_{\alpha+1} \dots p_\beta]p_{\beta+1} \dots p_\gamma \dots p_{\lambda-1}[q_{\kappa'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\mu'}p_{\lambda+1} \dots p_\omega\}.$$

This means that $p'' = p''''$, and then $(p^{x,i})^{y,j} = (p^{y,j})^{x,i}$.

Second case: Let $p_{x,i}$ be a proper subterm of $p_{y,j}$. (On the following picture, the previous occurrence of x is contained in b and the previous occurrence of y is in a.)



Then the term $q_{\xi'}$ contains the i-th occurrence of x for some $1 < \xi' \leq \beta'$ ($\xi' = 1$ would mean that either $p_{x,i} = p_{y,j}$ or that $p_{x,i}$ is a subterm of q_1). The term $q_{\xi'}$ is a subterm of $p''_{x,i}$ and equal to $\{[p_\phi \dots \{[p_{\alpha+1} \dots p_\beta\{xp_{\beta+1} \dots p_\gamma\}]p_{\gamma+1} \dots p_\nu\}$. The subterm $p'_{x,i}$ is equal to some $q_{\rho'}$, $1 \leq \rho' < \xi'$. Therefore,

$$p''_{x,i} = \{^1[q_{\rho'+1} \dots \{[q_{\xi'} \dots \{[q_{\alpha'+1} \dots q_{\beta'}\{yq_{\beta'+1} \dots q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\sigma'}\} \dots q_{\tau'}\}^1,$$

and $p_{\nu+1}$ will be equal to

$$p_{\nu+1} = \{[q_{\xi'+1} \dots \{[q_{\alpha'+1} \dots q_{\beta'}\{yq_{\beta'+1} \dots q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\eta'}\}.$$

Then $p_{y,j}$ is equal to

$$p'_{y,j}\{[q_1 \dots (q_{\rho'}\{^1[q_{\rho'+1} \dots q_{\xi'-1}\{(q_{\xi'}p_{\nu+1})q_{\eta'+1} \dots q_\theta\} \dots q_{\tau'}\}^1) \dots q_{\omega'}\}.$$

Here $p_k = q_{\rho'+i}$, for all $1 \leq k \leq \phi - 1 = \xi' - \rho' - 2$ and $p_k = q_{\eta'+k-\nu-1}$ for all $\nu + 1 < k \leq \omega$.

The term $p^{x,i}$ is obtained by replacing the subterm $p_{y,j}$ with the term p' which is equal to

$$p'_{y,j}\{[q_1 \dots \{[q_{\varepsilon'} \dots \{^2[q_{\rho'}p_1 \dots p_\beta]p_{\beta+1} \dots p_\nu p_{\nu+1}q_{\eta'+1} \dots q_{\tau'}\}^2] \dots q_{\delta'}\} \dots q_{\omega'}\}.$$

The term $(p^{x,i})^{y,j}$ is obtained from $p^{x,i}$ by replacing the subterm p' with

$$p'' = \{[^4p'_{y,j}q_1 \dots \{^3[q_{\rho'}p_1 \dots p_\beta]p_{\beta+1} \dots p_\nu\}^3 q_{\xi'+1} \dots q_{\beta'}]^4 q_{\beta'+1} \dots q_{\omega'}\}.$$

On the other hand, the term $p^{y,j}$ is obtained from p by replacing $p_{y,j}$ with

$$p''' = \{[p'_{y,j}q_1 \dots q_{\rho'} \dots q_{\xi'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\delta'} \dots q_{\omega'}\}.$$

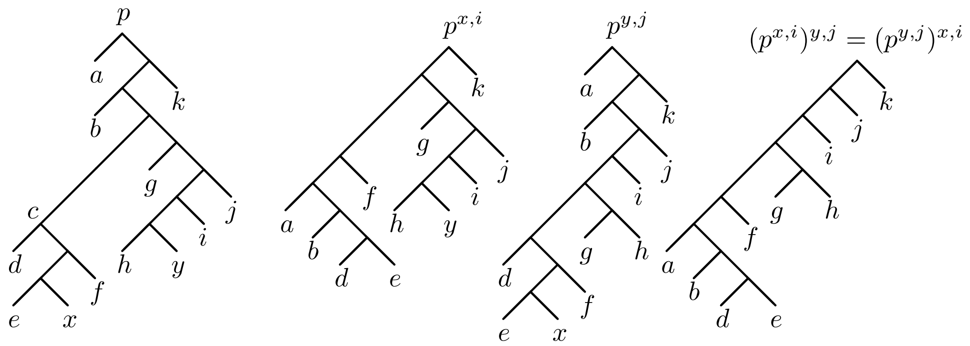
For the term $p^{y,j}$ it holds that $(p^{y,j})_{x,i} = [q_{\rho'} \dots q_{\xi'}[q_{\xi'+1} \dots q_{\beta'}]]$. Therefore, $(p^{y,j})^{x,i}$ is obtained from $p^{y,j}$ by replacing the subterm p''' with

$$\begin{aligned} p'''' &= \{[p'_{y,j}q_1 \dots \{^1[q_{\rho'}p_1 \dots p_\beta]p_{\beta+1} \dots p_\nu[q_{\xi'+1} \dots q_{\beta'}]\}^1]q_{\beta'+1} \dots q_{\omega'}\} \\ &= \{[p'_{y,j}q_1 \dots \{^1[q_{\rho'}p_1 \dots p_\beta]p_{\beta+1} \dots p_\nu\}^1 q_{\xi'+1} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\omega'}\} = p'', \end{aligned}$$

as desired.

Third case: Let $p_{y,j}$ be a proper subterm of $p_{x,i}$. Let p_ψ contain the jth occurrence of y in p, $\beta + 1 \leq \psi \leq \omega$. Let t be the maximal subterm of $p_{x,i}$ which does not contain p_ψ, but does contain the i-th occurrence of x in p. In other words, $t = \{[p_\lambda \dots \{[p_{\alpha+1} \dots p_\beta\{xp_{\beta+1} \dots p_\gamma\}]p_{\gamma+1} \dots p_{\psi-1}\}$. Since p_ψ is a subterm of $p''_{y,j}$, it follows that $p_\psi = \{[q_{\sigma'} \dots \{[q_{\alpha'+1} \dots q_{\beta'}\{yq_{\beta'+1} \dots\ q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\tau'}\}$. Consider two subcases:

First subcase: $p'_{y,j} = t$. (On the following picture, the previous occurrence of x is contained in a and the previous occurrence of y is in c.)



Then

$$p_{x,i} = p'_{x,i}\{[p_1 \dots (p_{\lambda-1}(p'_{y,j}p_\psi)) \dots p_\omega\}.$$

The term $p^{y,j}$ is obtained by replacing the subterm $p_{x,i}$ in p with the term p', which equals

$$p'_{x,i}\{[p_1 \dots (p_{\lambda-1}\{[p'_{y,j}q_{\sigma'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\gamma'} \dots q_{\tau'}\}) \dots p_\omega\}.$$

The term $(p^{y,j})^{x,i}$ we get from $(p^{y,j})$ by replacing p' with

$$p'' = \{[p'_{x,i}p_1 \dots p_{\lambda-1}p_\lambda \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}[q_{\sigma'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\tau'}p_{\psi+1} \dots p_\omega\}.$$

On the other hand, the term $p^{x,i}$ is obtained from p by replacing $p_{x,i}$ with the term p''', which equals

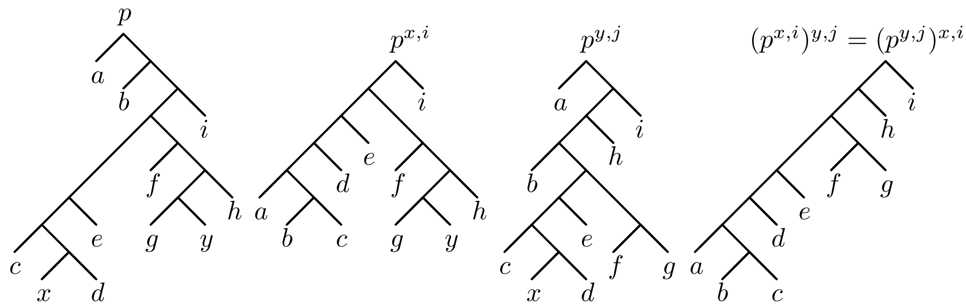
$$\{[p'_{x,i}p_1 \dots p_{\lambda-1}p_\lambda \dots p_{\alpha+1} \dots p_\beta]p_{\beta+1} \dots p_\gamma \dots p_{\psi-1}p_\psi p_{\psi+1} \dots p_\omega\}.$$

Now, $(p^{x,i})'_{y,j} = \{[p'_{x,i}p_1 \dots p_{\lambda-1} \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}\}$; it follows that $(p^{x,i})^{y,j}$ is obtained from $p^{x,i}$ by replacing p''' with

$$\begin{aligned} p'''' &= \{[\{[p'_{x,i}p_1 \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}\}q_{\sigma'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\tau'}p_{\psi+1} \dots p_\omega\} \\ &= \{[p'_{x,i}p_1 \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}[q_{\sigma'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\tau'}p_{\psi+1} \dots p_\omega\} = p'', \end{aligned}$$

which is what we needed.

Second subcase: $p'_{y,j} = p_\rho$, for some $1 \le \rho < \lambda$. (On the following picture, the previous occurrence of x is contained in a and the previous occurrence of y is in b.)



Then

$$p_{x,i} = p'_{x,i}(p_1 \dots p'_{y,j}(p_{\rho+1} \dots p_{\lambda-1}(tp_\psi)p_{\psi+1} \dots p_\sigma) \dots p_\omega).$$

The term $p^{y,j}$ is obtained by replacing $p_{x,i}$ in p with the term p' which equals

$$p'_{x,i}(p_1 \dots p_{\rho-1}\{[p'_{y,j}p_{\rho+1} \dots p_{\lambda-1}tq_{\sigma'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\tau'}p_{\psi+1} \dots p_\sigma\}p_{\sigma+1} \dots p_\omega).$$

The term $(p^{y,j})^{x,i}$ is obtained from $p^{y,j}$ by replacing the previous subterm with p'' which equals

$$\{[p'_{x,i}p_1 \dots p_{\rho-1}p'_{y,j}p_{\rho+1} \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}[q_{\sigma'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\tau'}p_{\psi+1} \dots p_\omega\}.$$

On the other hand, $p^{x,i}$ is obtained from p by replacing $p_{x,i}$ with the subterm

$$p''' = \{[p'_{x,i}p_1 \dots p_{\rho-1}p'_{y,j}p_{\rho+1} \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}p_\psi \dots p_\omega\}.$$

Since now $(p^{x,i})'_{y,j} = \{[p'_{x,i}p_1 \dots p_{\rho-1}p'_{y,j}p_{\rho+1} \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}\}$, it follows that $(p^{x,i})^{y,j}$ is obtained from $p^{x,i}$ by replacing the subterm p''' with p'''' which equals

$$\{[\{[p'_{x,i}p_1 \dots p_{\rho-1}p'_{y,j}p_{\rho+1} \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}\}q_{\sigma'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\tau'}p_{\psi+1} \dots p_\omega\}.$$

Then this subterm equals

$$\{[p'_{x,i}p_1 \dots p_{\rho-1}p'_{y,j}p_{\rho+1} \dots p_\beta]p_{\beta+1} \dots p_{\psi-1}[q_{\sigma'} \dots q_{\beta'}]q_{\beta'+1} \dots q_{\tau'}p_{\psi+1} \dots p_\omega\},$$

which is what we desired to prove. □

Theorem 12.3. *Any term p is equivalent to a unique linear groupoid term $\ell_3(p)$ modulo $\sim_3$.*

Proof. It follows directly from Lemmas 12.1 and 12.2. □

Next, we show that $\sim_3$ is a fully invariant congruence of the free groupoid $\mathbf{T}$.

Lemma 12.4. *$\sim_3$ is a congruence of $\mathbf{T}$.*

Proof. This follows obviously from the definition of $\sim_3$. □

Lemma 12.5. *Let the term p_x contain an occurrence of x. Then*

$$p_x\{[p_1 \dots \{[p_\alpha \dots \{xp_\beta \dots p_\gamma\}] \dots p_\omega\} \sim_3 \{[p_xp_1 \dots p_\alpha \dots p_{\beta-1}]p_\beta \dots p_\gamma \dots p_\omega\}.$$

Proof. We use the induction on the number of terms p_ψ, $1 \le \psi < \beta$, containing at least one occurrence of x.

Assume that only one term p_ψ contains an occurrence of x. Let $\ell_3(p_\psi) = \{[q_1 \dots \{[q_{\alpha'} \dots \{xq_{\beta'} \dots q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\omega'}\}$ and $p_\psi\{[p_{\psi+1} \dots \{[p_\alpha \dots \{xp_\beta \dots p_\gamma\}] \dots p_\chi\}$ be a subterm of the left side expression. Then

$$p_x\{[p_1 \dots p_{\psi-1}\{^1 p_\psi\{[p_{\psi+1} \dots \{[p_\alpha \dots \{xp_\beta \dots p_\gamma\}] \dots p_\chi\} \dots p_\delta\}^1 \dots p_\omega\} \sim_3$$

$$p_x\{[p_1 \dots p_{\psi-1}\{^1 \ell_3(p_\psi)\{[p_{\psi+1} \dots \{[p_\alpha \dots \{xp_\beta \dots p_\gamma\}] \dots p_\chi\} \dots p_\delta\}^1 \dots p_\omega\} \sim_3$$

$$\begin{gathered} p_x\{[p_1 \dots p_{\psi-1}\{^1\{^2[q_1 \dots \{[q_{\alpha'} \dots \{xq_{\beta'} \dots q_{\gamma'}\}]q_{\gamma'+1} \dots q_{\omega'}\}^2 \\ \{[p_{\psi+1} \dots \{[p_\alpha \dots \{xp_\beta \dots p_\gamma\}] \dots p_\chi\} \dots p_\delta\}^1 \dots p_\omega\} \sim_3 \end{gathered}$$

$$\begin{gathered} \{[p_xp_1 \dots p_{\psi-1}q_1 \dots q_{\alpha'} \dots q_{\beta'-1}]q_{\beta'} \dots q_{\gamma'} \dots q_{\omega'} \\ \{^1[p_{\psi+1} \dots \{[p_\alpha \dots \{xp_\beta \dots p_\gamma\}] \dots p_\chi\}^1 \dots p_\delta \dots p_\omega\} \sim_3 \end{gathered}$$

$$\begin{gathered} \{[\{^1[p_xp_1 \dots p_{\psi-1}q_1 \dots q_{\alpha'} \dots q_{\beta'-1}]q_{\beta'} \dots q_{\gamma'} \dots q_{\omega'}\}^1 \\ p_{\psi+1} \dots p_\alpha \dots p_{\beta-1}]p_\beta \dots p_\gamma \dots p_\chi \dots p_\delta \dots p_\omega\} \sim_3 \end{gathered}$$

$$\{[p_x p_1 \dots p_{\psi-1} \ell_3(p_\psi) p_{\psi+1} \dots p_\alpha \dots p_{\beta-1}] p_\beta \dots p_\gamma \dots p_\chi \dots p_\delta \dots p_\omega\} \sim_3$$

$$\{[p_x p_1 \dots p_{\psi-1} p_\psi p_{\psi+1} \dots p_\alpha \dots p_{\beta-1}] p_\beta \dots p_\gamma \dots p_\chi \dots p_\delta \dots p_\omega\}.$$

Next, assume that the claim is true for $n-1$ terms. Suppose that n terms $p_{i_1}, \dots, p_{i_n}$ contain an occurrence of x, $1 \le i_1 < \dots < i_n \le \beta - 1$. Then

$$p_x\{[p_1 \dots ({}^1 p_{i_1} \dots p_{i_2} \dots p_{i_n} \dots \{[p_\alpha \dots \{x p_\beta \dots p_\gamma\}] p_{\gamma+1} \dots)^1 \dots p_\omega\} \sim_3$$

(by the inductive assumption)

$$p_x\{[p_1 \dots \{{}^1[p_{i_1} \dots p_{i_2} \dots p_{i_n} \dots p_\alpha \dots p_{\beta-1}] p_\beta \dots p_\gamma p_{\gamma+1} \dots\}^1 \dots p_\omega\} \sim_3$$

(by the rule for cancelling, $p_x \cdot x \sim_3 p_x$)

$$p_x\{[p_1 \dots \{{}^1[p_{i_1} \dots p_{i_2} \dots p_{i_n} \dots p_\alpha \dots p_{\beta-1}] x p_\beta \dots p_\gamma p_{\gamma+1} \dots\}^1 \dots p_\omega\} \sim_3$$

(by the base case from above)

$$\{[p_x p_1 \dots p_{i_1} \dots p_{i_2} \dots p_{i_n} \dots p_\alpha \dots p_{\beta-1}] p_\beta \dots p_\gamma \dots p_\omega\}.$$

□

Lemma 12.6. *Let x and y occur in the term p_x. Then*

$$p_x\{[p_1 \dots \{[p_\alpha \dots \{x p_\beta \dots p_\gamma\}] \dots p_\xi p_{\xi+1} \dots p_\omega\} \sim_3$$
$$p_x\{[p_1 \dots \{[p_\alpha \dots \{x p_\beta \dots p_\gamma\}] \dots p_\xi y p_{\xi+1} \dots p_\omega\}.$$

Proof. Using Lemma 12.5, we obtain the following:

$$p_x\{[p_1 \dots \{[p_\alpha \dots \{x p_\beta \dots p_\gamma\}] \dots p_\xi p_{\xi+1} \dots p_\omega\} \sim_3$$
$$\{[p_x p_1 \dots p_\alpha \dots p_{\beta-1}] p_\beta \dots p_\gamma \dots p_\xi p_{\xi+1} \dots p_\omega\} \sim_3$$
$$\{[p_x p_1 \dots p_\alpha \dots p_{\beta-1}] p_\beta \dots p_\gamma \dots p_\xi y p_{\xi+1} \dots p_\omega\} \sim_3$$
$$p_x\{[p_1 \dots \{[p_\alpha \dots \{x p_\beta \dots p_\gamma\}] \dots p_\xi y p_{\xi+1} \dots p_\omega\}.$$

□

Lemma 12.7. *Let q_x and p_x be terms such that $S(p_x) \subseteq S(q_x)$ and let x be the leftmost variable of the term p_x. Then*

$$q_x\{[p_1 \dots \{[p_\alpha \dots \{x p_\beta \dots p_\gamma\}] \dots p_\omega\} \sim_3 q_x\{[p_1 \dots \{[p_\alpha \dots \{p_x p_\beta \dots p_\gamma\}] \dots p_\omega\}.$$

Proof. A corollary of the previous Lemma. □

Lemma 12.8. *Let t be a term, x and y variables of t, and let $s = t^{y,i}$ for some i. Then $\overline{t} \sim_3 \overline{s}$, where the terms $\overline{t}$ and $\overline{s}$ are obtained from the terms t and s by substitution of the variable x with a term p.*

Proof. Let us consider two cases.

First case: $x = y$. Then

$$t_{x,i} = t'_{x,i}\{[t_1 \dots \{[t_\alpha \dots \{x t_\beta \dots t_\gamma\}] t_{\gamma+1} \dots t_\omega\} \sim_3$$
$$\{[t'_{x,i} t_1 \dots t_\alpha \dots t_{\beta-1}] t_\beta \dots t_\gamma t_{\gamma+1} \dots t_\omega\}.$$

On the other hand,

$$\overline{t_{x,i}} = \overline{t'_{x,i}}\{[\overline{t_1} \dots \{[\overline{t_\alpha} \dots \{p \overline{t_\beta} \dots \overline{t_\gamma}\}] \overline{t_{\gamma+1}} \dots \overline{t_\omega}\} \sim_3$$

(using Lemma 12.5, where z is the leftmost variable p)

$$\overline{t'_{x,i}}\{[\overline{t_1}\dots\{[\overline{t_\alpha}\dots\overline{t_{\beta-1}}\{z\overline{t_\beta}\dots\overline{t_\gamma}\}]\overline{t_{\gamma+1}}\dots\overline{t_\omega}\} \sim_3$$

$$\{[\overline{t'_{x,i}}\overline{t_1}\dots\overline{t_\alpha}\dots\overline{t_{\beta-1}}]\overline{t_\beta}\dots\overline{t_\gamma}\dots\overline{t_\omega}\}.$$

The term $\{[\overline{t'_{x,i}}\overline{t_1}\dots\overline{t_\alpha}\dots\overline{t_{\beta-1}}]\overline{t_\beta}\dots\overline{t_\gamma}\dots\overline{t_\omega}\}$ is exactly what we get from the term $\{[t'_{x,i}t_1\dots t_\alpha\dots t_{\beta-1}]t_\beta\dots t_\gamma t_{\gamma+1}\dots t_\omega\}$ by substitution of the variable x with the term p. Since the term s is obtained from t by replacing the subterm $t_{x,i}$ with $\{[t'_{x,i}t_1\dots t_\alpha\dots t_{\beta-1}]t_\beta\dots t_\gamma t_{\gamma+1}\dots t_\omega\}$, it follows that $\overline{t} \sim_3 \overline{s}$.

Second case: $x \neq y$. Then

$$t_{y,i} = t'_{y,i}\{[t_1\dots\{[t_\alpha\dots\{yt_\beta\dots t_\gamma\}]t_{\gamma+1}\dots t_\omega\} \sim_3$$

$$\{[t'_{y,i}t_1\dots t_\alpha\dots t_{\beta-1}]t_\beta\dots t_\gamma t_{\gamma+1}\dots t_\omega\}.$$

On the other hand,

$$\overline{t_{y,i}} = \overline{t'_{y,i}}\{[\overline{t_1}\dots\{[\overline{t_\alpha}\dots\{y\overline{t_\beta}\dots\overline{t_\gamma}\}]\overline{t_{\gamma+1}}\dots\overline{t_\omega}\}.$$

Using Lemma 12.5, we get

$$\overline{t_{y,i}} \sim_3 \{[\overline{t'_{y,i}}\overline{t_1}\dots\overline{t_\alpha}\dots\overline{t_{\beta-1}}]\overline{t_\beta}\dots\overline{t_\gamma}\dots\overline{t_\omega}\}.$$

The term $\{[\overline{t'_{y,i}}\overline{t_1}\dots\overline{t_\alpha}\dots\overline{t_{\beta-1}}]\overline{t_\beta}\dots\overline{t_\gamma}\dots\overline{t_\omega}\}$ is exactly the term that we get from $\{[t'_{y,i}t_1\dots t_\alpha\dots t_{\beta-1}]t_\beta\dots t_\gamma t_{\gamma+1}\dots t_\omega\}$ by substitution of the variable x with the term p. Since the term s is obtained from t by replacing the subterm $t_{y,i}$ with $\{[t'_{y,i}t_1\dots t_\alpha\dots t_{\beta-1}]t_\beta\dots t_\gamma t_{\gamma+1}\dots t_\omega\}$, it follows that $\overline{t} \sim_3 \overline{s}$. □

Lemma 12.9. *$\sim_3$ is a fully invariant congruence of* $\mathbf{T}$.

Proof. Let t and p be terms. By application of Lemma 12.8 finitely many times, we get $\ell_3(t_1) = \ell_3(t_2)$, where the terms t_1 and t_2 are obtained from the terms $\ell_3(t)$ and t by replacing all the occurrences of the variable x with the term p. Therefore, the substitution rule holds, i. e. $\sim_3$ is a fully invariant congruence. □

Theorem 12.10. *$\sim_3$ is a $*$-linear equational theory extending* $\mathbf{Q}_4$.

Proof. $\sim_3$ is an equational theory according to Lemma 12.9. It is $*$-linear by Lemma 12.3. And it can be checked that $\mathbf{Q}_4$ is in the corresponding variety (in fact, it is sufficient to check that neither $\mathbf{Q}_1$, nor $\mathbf{Q}_2$, is in the variety). □

Let $\mathcal{L}_3$ denote the corresponding variety.

13. All $*$-linear theories

Theorem 13.1. *There are precisely six $*$-linear varieties of groupoids: $\mathcal{L}_1$, $\mathcal{L}_2$, $\mathcal{L}_3$ and their duals.*

Proof. It follows from the results of Sections 2, 3, 4, 5 and 6 that the groupoids $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_4$ and their duals are the only candidates for a 3-generated free groupoid of a $*$-linear equational theory. Theorems 8.3, 10.1 and 12.10 show that in each case there is at least one extending $*$-linear theory. And according to Theorem 7.3 and Corollary 9.2, the extensions are unique. □

14. $\mathcal{L}_3$ is inherently non-finitely based

In this section, t always denotes a term in variables $x_1, \dots, x_n$.

We start with several technical notions. Let $\varphi(t)$ denote the semigroup word obtained from a term t by deleting all parentheses and cancelling all exponents. E.g., $\varphi(x(y(y(z(xy)y)x))) = xyzxyx$.

We say that a term t has the *property* B_k (we write shortly $B_k(t)$), if

$$\varphi(t) = x_1 \dots x_n x_1 \dots x_n \dots x_1 \dots x_n x_1 \dots x_l w,$$

where w is an arbitrary word, $1 \leq l \leq n$ and $k = |\varphi(t)| - |w|$. The prefix of the length k is called the *head of* $\varphi(t)$. We say that an occurrence of a variable in the term t is a *head occurrence*, if the corresponding occurrence in $\varphi(t)$ is in its head. The key notion in further text is the *separator*. This is the leftmost occurrence of x_l in t such that the corresponding occurrence in $\varphi(t)$ is the rightmost letter of the head. E.g., the term $x(y(y(z(xy)y)x))$ has the property B_5 and the separator is y at the sixth position.

We say that a term t has the *property* A_k (shortly $A_k(t)$), if it has the property B_k and the property C_k saying that the left sequence of the separator contains only terms in a single variable. Note that C_k is equivalent to the fact that every subterm of t containing an occurrence left of the separator, either contains only one variable, or contains the separator. Also, note that $A_k(t)$ implies $A_j(t)$ for all $j \leq k$. E.g., the term $x(y(y(z(xy)y)z))$ has the property A_5, but it does not have the property A_6. Of course, all of the above properties are relative to the (linearly ordered) set of variables. We will mention which set of variables we are referring to, whenever it is not obvious.

In the sequel, we will use the notation $p_{x,i} = p'_{x,i}p''_{x,i}$ from the definition of $\sim_3$. By cancellation of the i-th occurrence of a variable x in a term t we mean application of the identity $t \approx t^{x,i}$. Again, $[y_1 y_2 \dots y_k]$ will stand for $y_1(y_2(\dots(y_{k-1}y_k)))$.

Lemma 14.1. *Let u be a subterm of a term t. If u contains only the leftmost occurrences of variables in t, then u is a subterm of $\ell_3(t)$.*

Proof. Consider cancellation of the i-th occurrence of a variable x in t ($i \geq 2$). Since u does not contain the i-th occurrence of x, either u is not a subterm of $p_{x,i}$, or it is a subterm of $p'_{x,i}$, or it is a subterm of some member of the left or right sequence. In all cases, u is also a subterm of $t^{x,i}$. □

Lemma 14.2. *If $k \leq n$, then $A_k(\ell_3(t))$ implies $A_k(t)$.*

Proof. First, we prove $B_k(t)$. Assume the opposite. There exists a variable x_j that occurs between x_i and x_{i+1} in the head of the word $\varphi(t)$ for some $i < k$. Indeed, $j < i$, because $\sim_3$ is left non-permutational. Let s be a term obtained from t by cancelling all non-first occurrences of variables left of this occurrence of x_j. Again, x_j occurs between x_i and x_{i+1} in the head of the word $\varphi(s)$. Assume that the left sequence of this occurrence in s is $s_1, \dots, s_m$ and that the first occurrence of x_j is in s_{m_0}. Then $s^{x_j,2}$ contains the subterm $[s_{m_0} s_{m_0+1} \dots s_m]$ and so does $\ell_3(s^{x_j,2}) = \ell_3(t)$ according to Lemma 14.1 (recall that all variables left of x_j occur at most once in s). This is a contradiction with the fact that $\ell_3(t)$ satisfies C_k, because this subterm contains more than one variable, but not the separator.

Next, we prove $C_k(t)$. Assume that there is a subterm u of t with more than one variable, containing an occurrence left of the separator, but not the separator. Let

s be a term obtained from t by replacing u with $\ell_3(u)$ and by cancelling all non-first occurrences of variables left of the subterm u. Either the first variable of $\ell_3(u)$ is different from its left neighbour in s, then $\ell_3(u)$ contains only first occurrences and thus, according to Lemma 14.1, $\ell_3(u)$ is a subterm of $\ell_3(s) = \ell_3(t)$, a contradiction with $A_k(\ell_3(t))$. Or this is not true, it means the first variable of $\ell_3(u)$, let us call it x, is identical with its left neighbour. Consider cancelling the second occurrence of x in s. The cancellation appears in the subterm $p_{x,2} = p'_{x,2}p''_{x,2}$ and x is the first variable of $p''_{x,2}$. So the left sequence of x in $p''_{x,2}$ is empty and its right sequence is non-empty; let q be its first member. Hence $p'_{x,2}q$ is a subterm of $p^{x,2}$ and thus also of $s^{x,2}$. It contains only leftmost occurrences, so, according to Lemma 14.1, it is a subterm of $\ell_3(s^{x,2}) = \ell_3(t)$ too. However, it does not contain the separator, a contradiction. □

Lemma 14.3. *If $k \leq n$, then $A_k(t)$ implies $A_k(t^{x,i})$ for any occurrence of x in t.*

Proof. $B_k(t^{x,i})$ follows from the fact that either $\varphi(t^{x,i}) = \varphi(t)$ (if one of the neighbours of the i-th occurrence of x is also x), or $\varphi(t^{x,i})$ results from $\varphi(t)$ by removing a non-first occurrence of the variable x.

Let us denote $q_1, \dots, q_m$ the left sequence of the separator in t. By assumptions, every q_j is a term in a single variable. To prove $C_k(t^{x,i})$, we consider two cases.

Case 1: the i-th occurrence of x precedes the separator. So there is j such that this occurrence is in q_j. We have two subcases. Either $p_{x,i}$ is a subterm of q_j. Then $t^{x,i}$ results from t be replacement of the term q_j by a different term, in the same single variable, therefore $C_k(t^{x,i})$ holds. Or the i-th occurrence of x is the first variable of q_j. Then $p'_{x,i} = q_{j-1}$ and the left sequence of the separator in $t^{x,i}$ is $q_1, \dots, q_{j-2}, q', q_{j+1}, \dots, q_m$, where q' is a term containing only the variable x (in fact, $q' = q_{j-1}r_1 \dots r_{m'}$, where $r_1, \dots, r_{m'}$ is the right sequence of the first occurrence in q_j). Hence $C_k(t^{x,i})$ holds too.

Case 2: the separator precedes the i-th occurrence of x. Let r be the member of the right sequence of the separator in t containing the i-th occurrence of x and let $r_1, \dots, r_{m_0}$ and $s_1, \dots, s_{m_1}$ be the left and right sequences of the occurrence in r. Let q denote the largest subterm of t containing the separator and not containing r. We have three subcases. First, the $(i-1)$-th occurrence of x in t is in r. Then $t^{x,i}$ results from t by replacement of the subterm r with another term, hence the left sequence of the separator remains unchanged a thus $C_k(t^{x,i})$ holds. Second, the $(i-1)$-th occurrence of x in t is in q. Then $p'_{x,i} = q$ and thus $p_{x,i} = qr$ is replaced for $[p'_{x,i}r_1r_2 \dots r_{m_0}]s_0 \dots s_{m_1}$. So the left sequence of the separator in $t^{x,i}$ is the same as in t and thus $C_k(t^{x,i})$ holds. If none of the two cases takes place, then $p'_{x,i} = q_j$ for some $j \leq m_2$, where m_2 is the greatest number such that q_{m_2} is not contained in the subterm q. In this case, $p_{x,i} = q_jp''_{x,i}$ is replaced for $[q_jq_{j+1} \dots q_{m_2}qr_1 \dots r_{m_0}]s_0 \dots s_{m_1}t_{m_3} \dots t_{m_4}$, where $t_{m_3}, \dots, t_{m_4}$ is a part of the right sequence of the separator in t. Consequently, the left sequence of the separator in $t^{x,i}$ is the same as in t and thus $C_k(t^{x,i})$ holds. □

Corollary 14.4. *Let t, s be terms in variables $x_1, \dots, x_n$ such that $\mathcal{L}_3$ satisfies $t \approx s$. If $k \leq n$, then $A_k(t)$ if and only if $A_k(s)$.*

Proof. Lemmas 14.2 and 14.3 yield $A_k(\ell_3(t))$ iff $A_k(t)$. The claim thus follows from the fact that $\mathcal{L}_3$ satisfies $t \approx s$ iff $\ell_3(t) = \ell_3(s)$. □

Lemma 14.5. *Let t, s be terms in variables $x_1, \dots, x_n$ such that $t = \alpha(t')$ and $s = \alpha(s')$ for a substitution α and some terms t', s' of length at most n. Assume that $\mathcal{L}_3$ satisfies $t' \approx s'$. Then, for every k, $A_k(t)$ if and only if $A_k(s)$.*

Proof. For $k \leq n$ the claim follows from Corollary 14.4, so suppose $k > n$. Assume that $A_k(t)$ holds, we prove $A_k(s)$. Let $q_1, \dots, q_m$ denote the left sequence of the separator in t and consider the least i such that q_i contains the variable x_n. Let $r = q_i q$ be the minimal subterm of t containing q_i as a proper subterm. Indeed, r contains the separator.

Since t' has at most n letters, we conclude that r is a subterm of $\alpha(x)$ for some variable x (because $i \geq n$). Consequently, r is a subterm of s, because $\sim_3$ is regular. Moreover, all variables occurring left of the leftmost occurrence of x in t' are substituted by a term in a single variable different from x_n (since $q_1, \dots, q_{i-1}$ are such terms). Since $\sim_3$ is left non-permutational, the set of variables occurring left of the leftmost occurrence of x is the same in both t' and s'. So left of the leftmost occurrence of the subterm r in s there is no occurrence of the variable x_n; it means, the first occurrence of x_n in s is the leftmost variable of the subterm r. However, according to Corollary 14.4, $A_n(s)$ holds. Particularly, $B_n(s)$ says that the variables left of the leftmost subterm r are in the ascending order. Since the rest of the head occurrences is in r (and thus untouched), $B_k(s)$ holds. So, we have a separator in s and we denote $q'_1, \dots, q'_{m'}$ its left sequence. Let j be the least number such that q'_j contains the variable x_n. Again, since r is a subterm of both s and t, we have $q'_j = q_i$, $q'_{j+1} = q_{i+1}$, $\dots$, $q'_{m'} = q_m$ and it follows from $C_n(s)$ that $q'_1, \dots, q'_{j-1}$ are also terms in a single variable. Hence $C_k(s)$ holds too. $\square$

Lemma 14.6. *Let Σ be a finite set of identities of $\mathcal{L}_3$ with lengths of terms at most n and let $\Sigma \vdash t \approx s$, where t and s are terms in variables $x_1, \dots, x_n$. Then, for every k, $A_k(t)$ if and only if $A_k(s)$.*

Proof. We first notice the (rather obvious) fact that there exists a finite set of identities $\Sigma' \supseteq \Sigma$ over the set of variables $\{x_1, \dots, x_n\}$ used in some proof of $\Sigma \vdash t \approx s$, which is obtained from Σ using only the Substitution rule, such that we need not use the Substitution rule in proving $\Sigma' \vdash t \approx s$. We also may assume (and do) that Σ' is closed under substitutions that permute variables.

Let M_k be the set of all identities in variables $x_1, \dots, x_n$ provable from Σ' without using the Substitution rule such that A_k holds for one side of the identity and fails for the other one. We prove by induction that M_k is empty for every k. Particularly, we get that $A_k(t)$ if and only if $A_k(s)$.

For contradiction, let m be the smallest number such that M_m is non-empty. According to Corollary 14.4, we have $m > n$.

Pick an identity $p \approx q \in M_m$ with the shortest proof from Σ' without using the Substitution rule and let $p_1 \approx q_1$, $p_2 \approx q_2$, $\dots$, $p_l \approx q_l$ be the shortest proof; hence $p_l = p$ and $q_l = q$. Because of Lemma 14.5, the identity $p \approx q$ is not in Σ' (it means $l \neq 1$). Also, $p \approx q$ is not obtained from the previous identities by symmetry, as otherwise $q \approx p \in M_m$ would have a shorter proof. Similarly, $p \approx q$ cannot be obtained from the previous identities by transitivity on $p_i \approx q_i$ and $p_j \approx q_j$ with $q_i = p_j$.

So $p \approx q$ must be obtained by the Replacement rule, i.e., there is an identity $p_i \approx q_i$ from the proof such that q is obtained from p by replacing its subterm p_i with q_i. In the rest of the proof, we will only speak of *this* occurrence of p_i in

p, the one which is being replaced by q_i. So when we mention a subterm p_i of p, we mean, in fact, "the occurrence of p_i in p that is replaced in the l-th step of the proof." Without loss of generality, suppose that $A_m(p)$ holds and $A_m(q)$ fails. Hence the subterm p_i of p (the one which is being replaced) contains some head occurrences of variables. If p_i is composed of only one variable, then q_i is a term composed of the same variable and $A_m(q)$ is a clear consequence of $A_m(p)$. If p_i has no occurrences of a variable to the left of the separator, then $A_m(p)$ implies $A_m(q)$, too. Therefore, p_i must contain two head occurrences of different variables and, by $A_m(p)$, the subterm p_i contains the separator in p. We have two cases.

First case: The subterm p_i contains a head occurrence of x_1 such that no head occurrences, other than possibly some more occurrences of x_1, lie to the left of p_i in p. Then the identity $p_i \approx q_i$ is in M_m and it has a shorter proof than $p \approx q$, a contradiction.

Second case: The subterm p_i contains the separator x_s of p, but p_i does not satisfy A_m with the same occurrence of x_s as separator. Let the left sequence of the separator in p be $r_1, \dots, r_\alpha$. Then the left sequence in p_i of the same occurrence of x_s which is the separator of p is $r_\beta, r_{\beta+1}, \dots, r_\alpha$. Obviously, each r_j has to have exactly one variable. Now, let $\varphi(r_\beta) = x_\beta$ and let ψ be the substitution $x_\beta \mapsto x_1$, $x_{\beta+1} \mapsto x_2$, $\dots$, $x_n \mapsto x_{1+n-\beta}$, $x_1 \mapsto x_{2+n-\beta}$, $\dots$, $x_{\beta-1} \mapsto x_n$. Then $\Sigma' \vdash \psi(p_i) \approx \psi(q_i)$ without using the Substitution Rule (just use the sequence $\psi(p_1) \approx \psi(q_1)$, $\psi(p_2) \approx \psi(q_2), \dots,$ $\psi(p_i) \approx \psi(q_i)$ and the fact that Σ' is closed under ψ). Now, as $A_h(\psi(p_i))$ holds for some $h < m$ with the separator $\psi(x_s)$ (the same occurrence which serves as the separator in p), then by the inductive assumption $A_h(\psi(q_i))$ holds, as well. But that means that q must satisfy at least B_m. Consider the occurrence of x_s which is the separator of q and a subterm r in its left sequence. This subterm is either in q_i, or is equal to some r_γ, $\gamma < \beta$. In the second case, it obviously has only one variable. In the first case, $\psi(r)$ is in the left sequence of the separator $\psi(x_s)$ in $\psi(q_i)$, and so contains exactly one variable. But then so does r, as ψ just renames the variables. In both cases, $A_m(q)$ holds, a contradiction. □

Theorem 14.7. *The variety $\mathcal{L}_3$ is inherently non-finitely based.*

Proof. Let $\mathcal{L}_3^n$ denote the variety based by the identities of $\mathcal{L}_3$ in at most n variables. We prove that $\mathcal{L}_3^n$ is not locally finite for any n and thus that $\mathcal{L}_3$ is inherently non-finitely based.

Note that $\mathcal{L}_3^n$ has a base Σ_n of identities of length at most $2n$: it can be obtained from the multiplication table of the n-generated free groupoid by setting $rs \approx \ell_3(rs)$ where r, s runs through all linear terms in n variables. Consider the terms

$$t_i = [\underbrace{x_0 \dots x_{2n} x_0 \dots x_{2n} \dots x_0 \dots x_{2n} x_0 \dots x_{i-1 \bmod 2n+1}}_{i \text{ letters}}],$$

for every $i \geq 2n+1$. Clearly, $A_k(t_i)$ holds, if and only if $k \leq i$. Therefore, by Lemma 14.6, all t_i are pairwise inequivalent in Σ_n, hence the free $(2n+1)$-generated groupoid in the variety $\mathcal{L}_3^n$ is infinite. □

15. The lattices of subvarieties of $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$

Lemma 15.1. *In $\mathcal{L}_i$, $i \in \{1, 2, 3\}$, each of the identities*

(a) $xy \approx yx$,

(b) $yx \approx x$,
(c) $xy \approx x$,
(d) $(xy)z \approx (xz)y$,
(e) $(xy)z \approx x(zy)$

implies $(xy)z \approx x(yz)$. *In* $\mathcal{L}_2$ *and* $\mathcal{L}_3$*, each of the identities*

(f) $x(yz) \approx x(zy)$,
(g) $w((xy)z) \approx w(x(yz))$

implies $(xy)z \approx x(yz)$. *In* $\mathcal{L}_i$*, the identity (f) implies the indentity (g).*

Proof. (a) $(xy)z \approx z(xy) \approx_{\mathcal{L}_i} z(x(yz)) \approx (x(yz))z \approx_{\mathcal{L}_i} x(yz)$.
(b) $yx \approx x(yx) \approx_{\mathcal{L}_i} xy$ and then use (a).
(c) $(xy)z \approx xy \approx x \approx x(yz)$.
(d) $(xy)z \approx_{\mathcal{L}_i} (xy)(yz) \approx (x(yz))y \approx_{\mathcal{L}_i} x(yz)$.
(e) $(xy)z \approx_{\mathcal{L}_i} (xy)(yz) \approx x(yzy) \approx_{\mathcal{L}_i} x(yz)$.
(f) $(xy)z \approx_{\mathcal{L}_2} x(xyz) \approx x(z(xy)) \approx_{\mathcal{L}_2} x(zy) \approx x(yz)$ and $(xy)z \approx_{\mathcal{L}_3} x((yx)z) \approx x(z(yx)) \approx_{\mathcal{L}_3} x(zy) \approx x(yz)$.
(g) $(xy)z \approx_{\mathcal{L}_i} x(xyz) \approx x(x(yz)) \approx_{\mathcal{L}_i} x(yz)$ for $i = 2, 3$.
The last claim can be proven analogously to (a). □

For a term t we denote by $\Phi(t)$ the sequence of the variables (possibly with repetitions) from $S(t)$, written in the order of their occurrences in t from the left to the right. So, $\Phi(u) = \Phi(v)$ if and only if $u \sim_a v$, where $\sim_a$ denotes the equational theory of semigroups.

Lemma 15.2. *If* $u \sim_a v$ *then* $\ell_1(u) \sim_a \ell_1(v)$.

Proof. It is easy to see that $\Phi\ell_1(u)$ is obtained from the sequence $\Phi(u)$ by deleting all the non-first occurrences of variables. So, if $\Phi(u) = \Phi(v)$ then $\Phi\ell_1(u) = \Phi\ell_1(v)$. □

Lemma 15.3. *Let* E_1 *consist of equations* $u \approx v$ *such that* $\ell_1(u) = xu_1 \dots u_n$ *and* $\ell_1(v) = xv_1 \dots v_n$ *for a variable* x*, a nonnegative integer* n *and terms* u_i, v_i *such that* $u_i \sim_a v_i$*. Then* E_1 *is the equational theory generated by* $\sim_1$ *and the equation* $w(xy \cdot z) \approx w(x \cdot yz)$.

Proof. We are going to prove that E_1 is an equational theory; the rest is easy. Clearly, E_1 is an equivalence containing $\sim_1$.

Let $u \approx v$ belong to E_1, $\ell_1(u) = xu_1 \dots u_n$, $\ell_1(v) = xv_1 \dots v_n$.

Let t be a term. We have $\ell_1(ut) = \ell_1(u)\delta_u\ell_1(t) = xu_1 \dots u_n\delta_u\ell_1(t)$ and $\ell_1(vt) = \ell_1(v)\delta_v\ell_1(t) = xv_1 \dots v_n\delta_v\ell_1(t)$ where $\delta_u = \delta_v$, hence $ut \approx vt$ in E_1. We have $\ell_1(tu) = \ell_1(t)\delta_t\ell_1(u)$ and $\ell_1(tv) = \ell_1(t)\delta_t\ell_1(v)$; since $\ell_1(u) \sim_a \ell_1(v)$ obviously implies $\delta_t\ell_1(u) \sim_a \delta_t\ell_1(v)$, we get $tu \approx tv$ in E_1. So, E_1 is a congruence.

Let f be a substitution. Denote by g the endomorphism of $\mathbf{L}$ such that $g(x) = \ell_1 f(x)$ for all $x \in X$. Then $\ell_1 f$ and $g\ell_1$ are two homomorphisms of $\mathbf{T}$ into $\mathbf{L}$ coinciding on X, and hence $\ell_1 f = g\ell_1$. So,

$$\begin{aligned}\ell_1 f(u) =& g\ell_1(u) = g(xu_1 \dots u_n) = g(x) \circ g(u_1) \circ \dots \circ g(u_n)\\ =& g(x) \cdot \delta_{g(x)} g(u_1) \cdot \dots \cdot \delta_{g(xu_1 \dots u_{n-1})} g(u_n)\end{aligned}$$

and similarly $\ell_1 f(v) = g(x) \cdot \delta_{g(x)} g(v_1) \cdot \dots \cdot \delta_{g(xv_1 \dots v_{n-1})} g(v_n)$. For every i we have $g(u_i) = g\ell_1(u_i) = \ell_1 f(u_i) \sim_a \ell_1 f(v_i) = g\ell_1(v_i) = g(v_i)$, since $u_i \sim_a v_i$ implies $f(u_i) \sim_a f(v_i)$ and hence $\ell_1 f(u_i) \sim_a \ell_1 f(v_i)$ by Lemma 15.2. Since

$S(xu_1 \dots u_{i-1}) = S(xv_1 \dots v_{i-1})$, the terms $g(xu_1 \dots u_{i-1})$ and $g(xv_1 \dots v_{i-1})$ contain the same variables, the corresponding δ-operators are equal and we get

$$\delta_{g(xu_1 \dots u_{i-1})} g(u_i) \sim_a \delta_{g(xv_1 \dots v_{i-1})} g(v_i)$$

(these are either both empty or both nonempty). Hence $f(u) \approx f(v)$ in E_1. $\square$

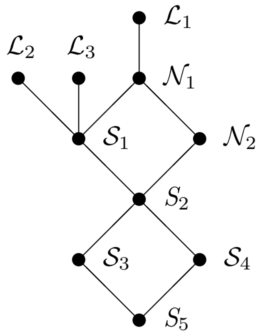
Lemma 15.4. *Let E_2 consist of equations $u \approx v$ such that $\ell_1(u) = xu_1 \dots u_n$ and $\ell_1(v) = xv_1 \dots v_n$ for a variable x, a nonnegative integer n and terms u_i, v_i such that $S(u_i) = S(v_i)$. Then E_2 is the equational theory generated by $\sim_1$ and the equation $w \cdot xy \approx w \cdot yx$.*

Proof. It is similar to the proof of Lemma 15.3. $\square$

Let us denote

- $\mathcal{N}_1$ the variety of $\mathcal{L}_1$-algebras satisfying $w(xy \cdot z) \approx w(x \cdot yz)$;
- $\mathcal{N}_2$ the variety of $\mathcal{L}_1$-algebras satisfying $w \cdot xy \approx w \cdot yx$;
- $\mathcal{S}_1$ the variety of idempotent semigroups satisfying $xyx \approx xy$;
- $\mathcal{S}_2$ the variety of idempotent semigroups satisfying $wxy \approx wyx$;
- $\mathcal{S}_3$ the variety of semigroups satisfying $xy \approx x$;
- $\mathcal{S}_4$ the variety of semilattices;
- $\mathcal{S}_5$ the trivial variety.

Theorem 15.5. *The following diagram shows a lower part of the lattice of subvarieties of groupoids:*



Proof. Let $\mathcal{L} \in \{\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3\}$ and ℓ be the corresponding normal form function. One can easily see that the intersection of $\mathcal{L}$ with the variety of semigroups is the variety $\mathcal{S}_1$. Since there is a full description of the lattice of varieties of idempotent semigroups (e.g., [3]), it is sufficient to focus on non-associative subvarieties of $\mathcal{L}$ only. According to Lemma 15.1, $w(xy \cdot z) \approx w(x \cdot yz)$ is a consequence of $w \cdot xy \approx w \cdot yx$ and the equations of $\mathcal{L}_1$, so we have all the inclusions listed above; it follows from Lemmas 15.3 and 15.4 that they are proper inclusions, and we do not have any other ones.

Let E be an equational theory containing the equational theory of $\mathcal{L}$. It is easy to see that if E contains an equation $u \approx v$ such that $S(u) \neq S(v)$, then E contains $xy \approx x$ or $xy \approx y$; and if E contains an equation $u \approx v$ where u, v have different first variables, then E contains $xy \approx yx$. In both cases, 15.1 yields associativity. So, it remains to consider the case when all equations of E are regular and both sides of any equation from E start with the same variable.

Let $u \approx v$ in E, so that $\ell(u) \approx \ell(v)$ and we can write $\ell(u) = xu_1 \dots u_k$ and $\ell(v) = xv_1 \dots v_m$ for a variable x, two nonnegative integers k, m and some terms

u_i, v_j. If it is possible to choose $u \approx v$ in such a way that there is an index i with $i \leq k$, $i \leq m$ and $S(u_i) \neq S(v_i)$, then (where i is the minimal index with this property) modify $\ell(u) \approx \ell(v)$ by a substitution sending x and all the variables of $S(u_1) \cup \dots \cup S(u_{i-1})$ to x, one fixed variable $y \in S(u_i) - S(v_i)$ to itself (we can assume without loss of generality that there is such a y) and all the other variables to a fixed variable $z \in S(v_i)$ to obtain in E one of these three equations: either $x \cdot zy \approx xz \cdot y$ or $x \cdot yz \approx xz \cdot y$ or $xy \cdot z \approx xz \cdot y$. By 15.1, each of them implies (together with the equations of $\mathcal{L}$) the associative law, and we are in the semigroup case. So, we can now assume that for any $u \approx v$ in E we have $k = m$ and $S(u_i) = S(v_i)$ for all i (thus, in the case of $\mathcal{L} = \mathcal{L}_1$, E is contained in the equational theory of $\mathcal{N}_2$). If it is possible to choose $u \approx v$ in such a way that $u_i \not\sim_a v_i$ for some i, then take two distinct variables y, z of $S(u_i)$ such that y occurs before z in $\Phi(u_i)$ but after z in $\Phi(v_i)$ and modify $\ell(u) \approx \ell(v)$ by the substitution sending y, z to themselves and all the other variables to x; we get $x \cdot yz \approx x \cdot zy$, thus, in the case of $\mathcal{L} = \mathcal{L}_1$, E is equal to the equational theory of $\mathcal{N}_2$, and in the other cases, by 15.1, E contains associativity. Now we can assume that for any $u \approx v$ in E we have $k = m$ and $u_i \sim_a v_i$ for all i (thus, in the case of $\mathcal{L} = \mathcal{L}_1$, E is contained in the equational theory of $\mathcal{N}_1$). If it is possible to choose $u \approx v$ in such a way that $u_i \neq v_i$ for some i, then it is again easy to set up a substitution to obtain the equation $w(xy \cdot z) \approx w(x \cdot yz)$ in E. Thus, in the case of $\mathcal{L} = \mathcal{L}_1$, E is the equational theory of $\mathcal{N}_1$, in the other cases, by 15.1, E contains associativity. Finally, if any $u \approx v$ in E satisfies $u_i = v_i$ for every i, E is the equational theory of $\mathcal{L}$. □

16. Generators for the varieties $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$

Denote by $\mathbf{F}_{\mathcal{V}}(n)$ the free n-generated groupoid in a variety $\mathcal{V}$.

Theorem 16.1. *The variety $\mathcal{L}_1$ is generated by $\mathbf{F}_{\mathcal{L}_1}(4)$, but not by $\mathbf{F}_{\mathcal{L}_1}(3)$ (it belongs to $\mathcal{N}_1$); it is generated by the groupoid $\mathbf{F}_{\mathcal{L}_1}(3)$ extended by the unit element. Also, $\mathcal{L}_1$ is generated by the five-element subdirectly irreducible groupoid with the following multiplication table:*

	a	b	c	d	e
a	a	b	d	d	a
b	b	b	c	c	b
c	c	c	c	c	c
d	d	d	d	d	d
e	a	b	c	d	e

Proof. Using Theorems 9.1 and 15.5, it is easy to check if a given groupoid generates $\mathcal{L}_1$. □

Theorem 16.2. *The variety $\mathcal{L}_i$ is generated by $\mathbf{F}_{\mathcal{L}_i}(3)$, $i \in \{2, 3\}$. Also, $\mathcal{L}_2$ is generated by the five-element subdirectly irreducible groupoid with the following multiplication table:*

	a	b	c	d	e
a	a	d	c	d	e
b	b	b	e	b	e
c	c	c	c	c	c
d	d	d	c	d	c
e	e	e	e	e	e

and $\mathcal{L}_3$ is generated by the four-element subdirectly irreducible groupoid with the following multiplication table:

	a	b	c	d
a	a	c	c	d
b	c	b	c	d
c	c	c	c	c
d	d	d	d	d

Proof. The free 3-generated groupoids are not semigroups, hence, by Theorem 15.5, they generate the respective variety. The smaller groupoids are quotients of the free ones and are not semigroups too. □

17. Quasi-$*$-linear theories of semigroups

In the last section, we discuss *quasi-$*$-linear varieties of semigroups.* This is a variety of semigroups such that in the corresponding equational theory every word is equivalent to a unique linear word. (It means, quasi-$*$-linearity is $*$-linearity modulo associativity.) We show that $\mathcal{S}_1$ and its dual are the only quasi-$*$-linear varieties of semigroups.

Lemma 17.1. *There are precisely three sharply 2-linear theories of semigroups. Their 2-generated free semigroups are $\mathbf{G}_1$, $\mathbf{G}_6$ and its dual, respectively.*

Proof. In idempotent semigroups, $x \cdot yx \approx xy \cdot x \approx xy \cdot yx$ and $x \cdot xy \approx xy \cdot y \approx xy$. A groupoid $\mathbf{G}_i$ satisfies these conditions, iff $i \in \{1, 6\}$. It is easy to check that both are semigroups, hence they serve as the 2-generated free semigroup for a 2-linear theory of semigroups. □

Lemma 17.2. *We cannot have $\mathbf{G}_1$ as the free two-generated groupoid for a quasi-3-linear theory of semigroups.*

Proof. From $\mathbf{G}_1$ we have $xyx \approx x$. Consequently, $xyz \approx xyzxz \approx xz$, a contradiction. □

Theorem 17.3. *There are precisely two quasi-$*$-linear varieties of semigroups: $\mathcal{S}_1$ and its dual. $\mathcal{S}_1$ is generated by $\mathbf{G}_6$ extended by a unit element and it is also generated by the following three-element semigroup:*

	a	b	c
a	a	b	c
b	b	b	b
c	c	c	c

Proof. $\mathbf{G}_6$ and its dual are the only candidates for the two-generated free groupoid. Any quasi-$*$-linear theory of semigroups extending $\mathbf{G}_6$ must contain the equation $xyx \approx xy$, hence it must contain $\mathcal{S}_1$. It is easy to see that $\mathcal{S}_1$ is quasi-$*$-linear, so it is the unique quasi-$*$-linear extension of $\mathbf{G}_6$. □

References

[1] J. Dudek, *Small idempotent clones I.* Czechoslovak Math. J. **48** (1998), 105–118.

[2] R. Freese, J. Ježek, P. Jipsen, P. Marković, M. Maróti and R. McKenzie, *The variety generated by order algebras.* Algebra Universalis **47/2** (2002), 103–138.

[3] J.A. Gerhard, *The lattice of equational classes of idempotent semigroups.* J. Algebra **15** (1970), 195–224.

[4] J. Ježek, *Three-variable equations of posets.* Czechoslovak Math. J. **52/4** (2002), 811–816.
[5] W.W. McCune, *Otter: An Automated Deduction System.*
Available at `http://www-unix.mcs.anl.gov/AR/otter/`
[6] R. McKenzie, G. McNulty and W. Taylor, *Algebras, Lattices, Varieties, Volume I.* Wadsworth & Brooks/Cole, Monterey, CA, 1987.

Petar Đapić, Petar Marković, Department of Mathematics, Faculty of Science, University of Novi Sad, trg Dositeja Obradovića 3, 21000 Novi Sad, Serbia and Montenegro
E-mail address: `pera@im.ns.ac.yu`

Jaroslav Ježek, David Stanovský, Department of Algebra, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 18600 Praha 8, Czech Republic
E-mail address: `jezek@karlin.mff.cuni.cz, stanovsk@karlin.mff.cuni.cz`

Ralph McKenzie, Department of Mathematics, Vanderbilt University, 1326 Stevenson Center, Nashville, TN 37240, U.S.A.
E-mail address: `mckenzie@math.vanderbilt.edu`