Algebraic & Geometric Topology 12 (2012) 469–492 **469**

Associahedra and weak monoidal structures on categories

ZBIGNIEW FIEDOROWICZ
STEVEN GUBKIN
RAINER M VOGT

This paper answers the following question: what algebraic structure on a category corresponds to an A_n structure (in the sense of Stasheff) on the geometric realization of its nerve?

18D10, 55P48; 06A07

In his trailblazing paper [12], Stasheff constructed an infinite hierarchy of higher homotopy associativity conditions for an H–space X. These conditions are parametrized by a family $\{K_n\}_{n\geq 2}$ of polyhedra, which came to be known as associahedra. The vertices of K_n are in 1-1 correspondence with all possible ways of associating an n–fold product $x_1x_2\dots x_n$, and an H–space X is said to be an A_n–space if there is a map $K_n \times X^n \longrightarrow X$ whose restriction to the vertices enumerates all possible ways of associating the binary multiplication on X into an n–fold multiplication. An A_∞–space is known to be equivalent to a strict monoid MX and hence, up to group completion, to a loop space.

At the same time[1] Mac Lane [9] analyzed higher associativity conditions for monoidal structures on categories. He formulated analogs of Stasheff's A_n conditions for categories. For $n = 2, 3, 4$ the analogy is perfect. In particular Mac Lane's A_4 condition is that a pentagonal diagram commute, whereas Stasheff's K_4 is a pentagon. However for $n \geq 5$ the analogy breaks down. Mac Lane's coherence theorem states that the A_4 condition implies all the higher A_n conditions for $n \geq 5$. By contrast for any $n \geq 2$ one can construct H–spaces X which satisfy the A_n condition but not the A_{n+1} (or any higher) condition.

In this paper we show how Mac Lane's notion of a monoidal structure on a category can be weakened so as to obtain a full hierarchy of A_n conditions. The paper is similar in spirit to Balteanu, Fiedorowicz, Schwänzl and Vogt [1] where an E_n hierarchy of commutativity conditions on categories was considered, analogous to those on n–fold

[1] Stasheff informs us that, although his paper [12] and Mac Lane's [9] both appeared in 1963, Mac Lane's work preceded his and influenced his thinking. He further informs us (cf [14]) that the associahedra were implicitly defined in the even earlier work of Tamari [15; 16].

Published: 20 March 2012 DOI: 10.2140/agt.2012.12.469

loop spaces. Similarly to the case of associativity for categories, Joyal and Street [5, Proposition 5.4] showed that if these commutativity conditions are required to hold up to natural isomorphisms, then the E_3 condition implies all higher E_n conditions. In [1] we demonstrated that we could recover the entire E_n hierarchy for categories by weakening these commutativity conditions to hold up to natural transformations instead. This strategy does not work for associativity, since LaPlaza [7, Theorem 5] showed that even if the associativity conditions are weakened to hold up to natural transformations, instead of isomorphisms, this laxened form of Mac Lane's A_4 condition still implies all higher A_n conditions. Thus a different strategy for weakening Mac Lane's A_n conditions for categories is required.

In Sections 1 and 2 we develop this strategy: we define the category theoretical analogues of Stasheff's associahedra in Section 1 and A_n–monoidal categories in Section 2. In Section 3 we relate our work to that of LaPlaza (and implicitly to that of Tamari) and give a simpler proof of his coherence result. In Section 4 we prove a rectification result for A_∞–monoidal categories, similar in spirit to Mac Lane's rectification of a monoidal category to a strictly monoidal one, by translating the rectification of an A_∞–space to a monoid into category theory.

This paper presupposes some familiarity with the notion of operad and related concepts. A précis of the relevant definitions may be found in May [10] and some historical context in Stasheff [13]. Since we will be dealing exclusively with noncommutative operations, we will be using the non–Σ forms of operads throughout.

To forestall any possible misunderstanding, it should be pointed out that this paper is not related in any significant way to the notion of A_∞–category as developed by Fukaya, Kontsevich, Soibelman and others (cf [6] for an overview). Here we discuss ordinary categories with weak monoidal structures, not some notion of a weak higher category.

We would like to take this opportunity to thank Jim Stasheff and Stefan Forcey for some helpful suggestions and references to previous work in this area and the referee for helpful comments for the presentation of the paper.

1 The associahedra as an operad in *CAT*

In order to keep track of associativity data for our weakly monoidal categories, we will need a categorical equivalent of the associahedron K_m. To begin with we formalize the notion of a parenthesized word:

Definition 1.1 A *parenthesized word* (W, P) is a finite linear order W together with a (possibly empty) collection of closed intervals $P = \{p_i = [a_i, b_i]\}$ subject to the following requirements.

- The cardinality of each p_i is at least 2 and is strictly smaller than the cardinality of W.
- For any i, j, either $p_i \subset p_j$, $p_j \subset p_i$ or $p_i \cap p_j = \varnothing$.

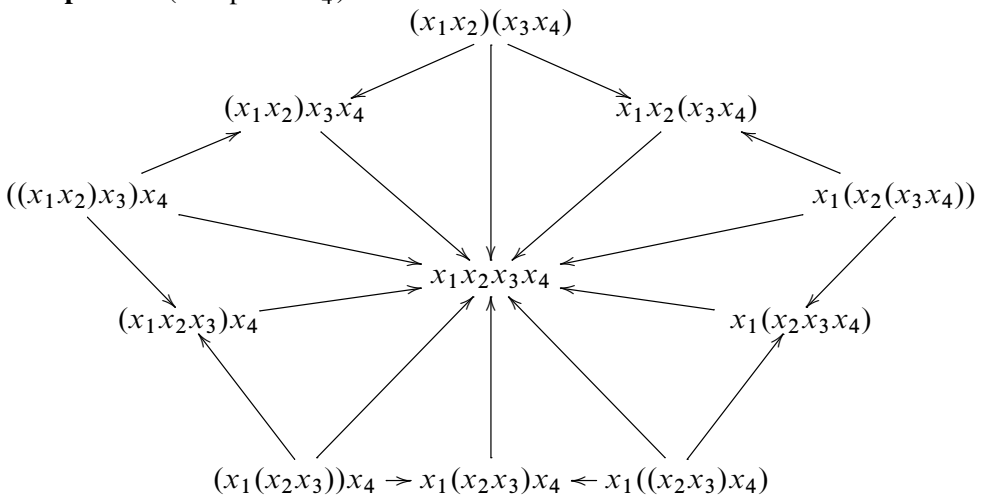
A parenthesized word (W, P) can be converted into a parenthesized string of characters by putting as many left parentheses in front of an element $a \in W$ as a is an initial element of some $p_i \in P$ and as many right parentheses after an element $b \in W$ as b is a final element in some $p_i \in P$, and concatenating the resulting characters, for instance.

$$\{x_1 < x_2 < x_3 < x_4 < x_5 < x_6, \{[x_2, x_6], [x_2, x_4], [x_5, x_6]\}\} \mapsto x_1((x_2x_3x_4)(x_5x_6)).$$

It is clear that (W, P) can be recovered from the parenthesized string and we will often find it convenient to represent (W, P) in this way. In most cases we will use the standard linear orders $W_m = \{x_1 < x_2 \cdots < x_m\}$. In some induction arguments however we will need to consider subintervals of the W_m.

Definition 1.2 We define $\mathfrak{K}_m$ to be the poset of parenthesized words on the linear order W_m, where $(W_m, P_2) \leq (W_m, P_1)$ if and only if $P_1 \subset P_2$. The minimal elements in this order are called the *fully parenthesized* words of length m. In the degenerate cases $m = 1$ and $m = 0$, the poset $\mathfrak{K}_1$ consists of the single parenthesized word $\mathrm{id} = (W_1, \varnothing)$, and $\mathfrak{K}_0$ consists of the single parenthesized word $0 = (\varnothing, \varnothing)$. The string $x_1x_2 \ldots x_m = (W_m, \varnothing)$ is the terminal object in $\mathfrak{K}_m$. As noted above, sometimes it will be convenient to use some other linear order W' of the same cardinality m. In that case the unique order isomorphism between W' and W_m specifies a canonical isomorphism between $\mathfrak{K}_m$ and the corresponding poset of parenthesized words on W'.

Example 1.3 (The poset $\mathfrak{K}_4$)



The m–th associahedron is defined to be the polytope which has one vertex for every fully parenthesized word of length m. Two vertices (W, P_i) and (W, P_j) are on the same k–dimensional face if they share at least $m-k$ parentheses, ie $P_i \cap P_j$ has cardinality at least $m-k$. Thus our poset $\mathfrak{K}_m$ is exactly the face poset of the m–th associahedron, and so the geometric realization of the nerve of $\mathfrak{K}_m$ is simply the barycentric subdivision of the m–th associahedron. Note that we are using a Fraktur font to distinguish the poset $\mathfrak{K}_m$ from the associahedron K_m which is the geometric realization of its nerve (as a topological space).

The following lemma will prove to be surprisingly useful:

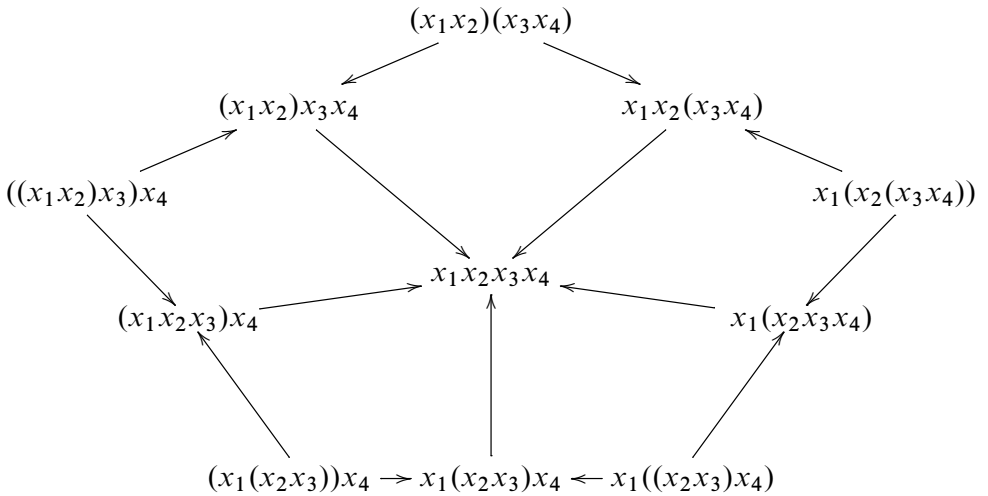
Lemma 1.4 *Let $(W_k, P) < (W_k, P')$ in $\mathfrak{K}_k$. Then the subposet*

$$[(W_k, P), (W_k, P')] = \{(W_k, P'') \in \mathfrak{K}_k | (W_k, P) \leq (W_k, P'') \leq (W_k, P')\}$$

is isomorphic to the poset $\mathcal{I}^m$ where $\mathcal{I}$ is the poset $1 < 0$ and m is the number of parentheses in (W_k, P) which are not in (W_k, P'). In other words, the factorizations of a fixed morphism in $\mathfrak{K}_k$ form a commutative cubical diagram.

Proof We can uniquely associate to each element (W_k, P'') in $[(W_k, P), (W_k, P')]$ a characteristic function on the set of parentheses in (W_k, P) which are not in (W_k, P') by giving the value 1 to each parenthesis which occurs in (W_k, P'') and 0 to any which do not so occur. But such a characteristic function is evidently the same thing as an object of $\mathcal{I}^m$ and it is clear that order relations match. □

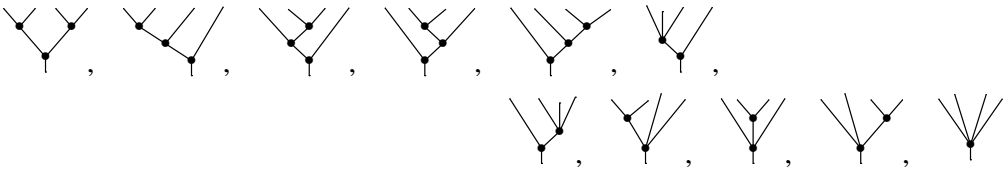
If we take $(W_k, P') = (W_k, \varnothing)$, then geometrically this gives a decomposition of the associahedra into cubes. The decomposition of K_4 into 5 squares looks like this:



The decomposition of K_5 into 14 cubes can be found on the arXiv with our paper [3] at `http://arxiv.org/src/1005.3979v4/anc/cubical.flv`.

There is a one-to-one correspondence between parenthesized words and *stable rooted trees*. Briefly these are planar rooted trees where each node has at least two input edges. Refer to Leinster [8] for a formal definition. The correspondence is given by labelling the leaves of such a tree with the labels $x_1, x_2, \ldots, x_n$ in left to right order. [In the degenerate cases $n = 1$ and $n = 0$, the identity $\mathrm{id} \in \mathfrak{K}_1$ corresponds to the tree with a single edge and no nodes and $\varnothing \in \mathfrak{K}_0$ corresponds to the empty tree with no edges and no nodes.] Then for each node of the tree, except for the bottom root node, one takes the set of labels sitting over that node as one of the intervals $p_i \in P$ in the collection P, thus giving us a parenthesized word (W_n, P). For example, here are all of the parenthesized words on the linear order W_4 and their corresponding stable rooted trees:

Example 1.5 The trees



represent $(x_1x_2)(x_3x_4)$, $((x_1x_2)x_3)x_4$, $(x_1(x_2x_3))x_4$, $x_1((x_2x_3)x_4)$, $x_1(x_2(x_3x_4))$, $(x_1x_2x_3)x_4$, $x_1(x_2x_3x_4)$, $(x_1x_2)x_3x_4$, $x_1(x_2x_3)x_4$, $x_1x_2(x_3x_4)$, $x_1x_2x_3x_4$ respectively.

The poset structure on $\mathfrak{K}_n$ of Definition 1.2 can be described in terms of trees as follows: $T < T'$ if T' can be obtained from T by shrinking some of the edges of T. It is of course more convenient to use parenthesized words when describing $\mathfrak{K}_n$ as a poset. However the language of trees is more convenient to describe the operad structure on the $\mathfrak{K}_n$.

The $\{\mathfrak{K}_i\}_{i\geq 0}$ form an operad in CAT, the category of small categories. Given stable rooted trees $S \in \mathfrak{K}_m$ and $T_i \in \mathfrak{K}_{k_i}$ for $i = 1, 2, \ldots, m$, we obtain a new tree by grafting the root of T_i to the i–th leaf of S. In terms of parenthesized words we are substituting the word for T_i in place of the i–th character of the word for S, and reindexing to insure that all characters in the resulting word are distinct. This only makes sense if $m \geq 2$ and all $k_i \geq 2$. If $S = \mathrm{id} \in \mathfrak{K}_1$, we define the composed tree to be T_1. If $T_i = \mathrm{id} \in \mathfrak{K}_1$, then we leave the i–th leaf of S unchanged. If $T_i = 0 \in \mathfrak{K}_0$, then we delete the i–th leaf of S. If this leaves only one input edge for the node below,

we delete that node as well. If it leaves no input edges for the node below, we delete both that node and the edge below. We apply this algorithm recursively: if the next node below receives only one input edge or no input edges, we delete that node or that node together with the edge below, and so on. In the special case when $\sum k_i = 1$ or $\sum k_i = 0$, the resulting degenerate trees are defined to be id or 0 respectively.

This process is clearly functorial in each of $\mathfrak{K}_m, \mathfrak{K}_{k_1}, \mathfrak{K}_{k_2}, \dots, \mathfrak{K}_{k_m}$, and so we obtain a functor

$$\gamma_{m,k_1,k_2,\dots,k_m}\colon \mathfrak{K}_m \times \prod_1^m \mathfrak{K}_{k_i} \to \mathfrak{K}_{\sum_1^m k_i}.$$

These functors define a categorical operad $\mathfrak{K} = \{\mathfrak{K}_i\}_{i\geq 0}$.

The associahedral operad $\mathfrak{K} = \{\mathfrak{K}_i\}_{i\geq 0}$ has an operadic filtration

$$\mathfrak{K}^{(2)} \subset \mathfrak{K}^{(3)} \subset \mathfrak{K}^{(4)} \subset \cdots,$$

where $\mathfrak{K}_i^{(n)}$ is the subposet of $\mathfrak{K}_i$ consisting of trees where each node has input valence $\leq n$ (ie has at most n incoming edges). We define $\mathfrak{K}^{(\infty)} = \mathfrak{K}$. We note for future reference that if an element (W, P) of $\mathfrak{K}_k$ lies in filtration n and $(W', P') < (W, P)$ then (W', P') also lies in filtration n. This follows from our description above of the poset structure in terms of trees.

Proposition 1.6 *The poset $\mathfrak{K}_i^{(n)}$ is the face poset of a subcomplex of the (unsubdivided) associahedron K_i. This subcomplex contains all cells of K_i of dimension $\leq n-2$. Consequently the nerve of $\mathfrak{K}_i^{(n)}$ is $(n-3)$–connected. In particular if $n \geq 4$ the nerve of $\mathfrak{K}_i^{(n)}$ is simply connected.*

Proof If an element (W, P) of $\mathfrak{K}_i$ is in $\mathfrak{K}_i^{(n)}$ and $(W', P') < (W, P)$ then (W', P') is also contained in $\mathfrak{K}_i^{(n)}$, since the tree representing (W, P) is obtained from the tree representing (W', P') by shrinking internal edges. Therefore $\mathfrak{K}_i^{(n)}$ is the face poset of a subcomplex of K_i. Now the vertices of K_i are parametrized by the elements of $\mathfrak{K}_i^{(2)}$, which are represented by binary trees. It follows that the cells of K_i of dimension j are obtained by shrinking j internal edges of a binary tree. It easily follows that the dimension of the cell parametrized by a given tree is the sum over all nodes of the incoming valence of that node minus 2. Thus the maximal possible incoming valence of a node in a tree parametrizing a cell of dimension j is $j+2$. Hence the subcomplex of K_i parametrized by $\mathfrak{K}_i^{(n)}$ contains all cells of K_i of dimension $\leq n-2$.

Now the nerve of $\mathfrak{K}_i^{(n)}$ is the barycentric subdivision of this subcomplex of K_i. Moreover K_i is obtained from this subcomplex by adding cells of dimensions $\geq n-1$. Since K_i is contractible, it follows that the complex and hence the nerve of $\mathfrak{K}_i^{(n)}$ is $(n-3)$–connected. □

Remark 1.7 $\mathfrak{K}_i^{(n)}$ is generally larger than the face poset of the $(n-2)$–skeleton of K_i. For instance $\mathfrak{K}_5^{(3)}$ is the face poset of the subcomplex of the 3–dimensional associahedron K_5 consisting of all the edges together with the three square faces.

Remark 1.8 Our categorical operad $\mathfrak{K}$ is almost the same as Leinster's *StTr* [8, pages 233–234]. The only difference is that he has $StTr(0) = \varnothing$, whereas we have $\mathfrak{K}_0 = \{0\}$. So our approach encodes the notion of a unit for algebras over $\mathfrak{K}$. Leinster expected that the nerve of $StTr(k) = \mathfrak{K}_k$ is homeomorphic to the associahedron, which we prove. Thus Leinster's topological operad is precisely the same as Stasheff's.

The tree description of a CAT–operad containing $\mathfrak{K}$ appears in Ginzburg and Kapranov [4].

Remark 1.9 Since the nerve of a product in CAT is a product in TOP, it follows that the nerve of a $\mathfrak{K}^{(n)}$ algebra is an A_n–space in the sense of Stasheff.

2 A_n–monoidal categories and coherence

Definition 2.1 For $n = 2, 3, \ldots, \infty$, an A_n–monoidal category is a category $\mathcal{C}$ together with multiplications $\mu_k\colon \mathcal{C}^k \to \mathcal{C}$ for $0 \le k < n+1$ such that

(1) $\mu_1\colon \mathcal{C} \to \mathcal{C}$ is the identity functor.

(2) $\mu_0\colon * \to \mathcal{C}$ is an object $0 \in \mathcal{C}$ that acts as a strict unit in the sense that

$$\mu_k(\mathrm{Id}_{\mathcal{C}}^i \times 0 \times \mathrm{Id}_{\mathcal{C}}^j) = \mu_{k-1}$$

for any i, j such that $i + j = k - 1$.

$\mathcal{C}$ is also equipped with natural transformations (associators)

$$\alpha^{i,j,k}\colon \mu_{i+1+k} \circ (\mathrm{Id}_{\mathcal{C}}^i \times \mu_j \times \mathrm{Id}_{\mathcal{C}}^k) \longrightarrow \mu_{i+j+k},$$

for $0 \le i + j + k < n + 1$, satisfying

(i) $\alpha^{i,0,k}, \alpha^{i,1,k}$ and $\alpha^{0,j,0}$ are the identity

and the coherence conditions specified by the following commutative diagrams:

(ii)
$$\begin{array}{ccc}
\mu_{a+b+d+2}(\bar{A}, 0, \bar{B}, \mu_c(\bar{C}), \bar{D}) & = & \mu_{a+b+d+1}(\bar{A}, \bar{B}, \mu_c(\bar{C}), \bar{D}) \\
\downarrow{\scriptstyle \alpha^{a+b+1,c,d}_{(\bar{A},0,\bar{B}),\bar{C},\bar{D}}} & & \downarrow{\scriptstyle \alpha^{a+b,c,d}_{(\bar{A},\bar{B}),\bar{C},\bar{D}}} \\
\mu_{a+b+c+d+1}(\bar{A}, 0, \bar{B}, \bar{C}, \bar{D}) & = & \mu_{a+b+c+d}(\bar{A}, \bar{B}, \bar{C}, \bar{D})
\end{array}$$

(iii)
$$\begin{array}{ccc}
\mu_{a+c+d+2}(\bar{A},\mu_b(\bar{B}),\bar{C},0,\bar{D}) & = & \mu_{a+c+d+1}(\bar{A},\mu_b(\bar{B}),\bar{C},\bar{D}) \\
\downarrow{\scriptstyle \alpha^{a,b,c+d+1}_{\bar{A},\bar{B},(\bar{C},0,\bar{D})}} & & \downarrow{\scriptstyle \alpha^{a,b,c+d}_{\bar{A},\bar{B},(\bar{C},\bar{D})}} \\
\mu_{a+b+c+d+1}(\bar{A},\bar{B},\bar{C},0,\bar{D}) & = & \mu_{a+b+c+d}(\bar{A},\bar{B},\bar{C},\bar{D})
\end{array}$$

(iv)
$$\begin{array}{ccc}
\mu_{a+d+1}(\bar{A},\mu_{b+c+1}(\bar{B},0,\bar{C}),\bar{D}) & = & \mu_{a+d+1}(\bar{A},\mu_{b+c}(\bar{B},\bar{C}),\bar{D}) \\
\downarrow{\scriptstyle \alpha^{a,b+c+1,d}_{\bar{A},(\bar{B},0,\bar{C}),\bar{D}}} & & \downarrow{\scriptstyle \alpha^{a,b+c,d}_{\bar{A},(\bar{B},\bar{C}),\bar{D}}} \\
\mu_{a+b+c+d+1}(\bar{A},\bar{B},0,\bar{C},\bar{D}) & = & \mu_{a+b+c+d}(\bar{A},\bar{B},\bar{C},\bar{D})
\end{array}$$

(v)
$$\begin{array}{ccc}
\mu_{a+c+e+2}(\bar{A},\mu_b(\bar{B}),\bar{C},\mu_d(\bar{D}),\bar{E}) & \xrightarrow{\alpha^{a,b,c+e+1}_{\bar{A},\bar{B},(\bar{C},\mu_d(\bar{D}),\bar{E})}} & \mu_{a+b+c+e+1}(\bar{A},\bar{B},\bar{C},\mu_d(\bar{D}),\bar{E}) \\
\downarrow{\scriptstyle \alpha^{a+c+1,d,e}_{(\bar{A},\mu_b(\bar{B}),\bar{C}),\bar{D},\bar{E}}} & & \downarrow{\scriptstyle \alpha^{a+c+1,d,e}_{(\bar{A},\bar{B},\bar{C}),\bar{D},\bar{E}}} \\
\mu_{a+c+d+e+1}(\bar{A},\mu_b(\bar{B}),\bar{C},\bar{D},\bar{E}) & \xrightarrow{\alpha^{a,b,c+d+e}_{\bar{A},\bar{B},(\bar{C},\bar{D},\bar{E})}} & \mu_{a+b+c+d+e}(\bar{A},\bar{B},\bar{C},\bar{D},\bar{E})
\end{array}$$

(vi)
$$\begin{array}{ccc}
\mu_{a+e+1}(\bar{A},\mu_{b+d+1}(\bar{B},\mu_c(\bar{C}),\bar{D}),\bar{E}) & \xrightarrow{\mu_{a+e+1}(\mathrm{id}_{\bar{A}},\alpha^{b,c,d}_{\bar{B},\bar{C},\bar{D}},\mathrm{id}_{\bar{E}})} & \mu_{a+e+1}(\bar{A},\mu_{b+c+d}(\bar{B},\bar{C},\bar{D}),\bar{E}) \\
\downarrow{\scriptstyle \alpha^{a,b+d+1,e}_{\bar{A},(\bar{B},\mu_c(\bar{C}),\bar{D}),\bar{E}}} & & \downarrow{\scriptstyle \alpha^{a,b+c+d,e}_{\bar{A},(\bar{B},\bar{C},\bar{D}),\bar{E}}} \\
\mu_{a+b+d+e+1}(\bar{A},\bar{B},\mu_c(\bar{C}),\bar{D},\bar{E}) & \xrightarrow{\alpha^{a+b,c,d+e}_{(\bar{A},\bar{B}),\bar{C},(\bar{D},\bar{E})}} & \mu_{a+b+c+d+e}(\bar{A},\bar{B},\bar{C},\bar{D},\bar{E})
\end{array}$$

Here $\bar{A}$, $\bar{B}$, $\bar{C}$, $\bar{D}$, $\bar{E}$ are taken to be objects of $\mathcal{C}^a$, $\mathcal{C}^b$, $\mathcal{C}^c$, $\mathcal{C}^d$, $\mathcal{C}^e$, respectively.

Essentially coherence conditions (i)–(iv) require the associators to be compatible with the strict unit 0, while (v) and (vi) just say that if we are removing two pairs of matching parentheses in a multiplication, it doesn't matter which we remove first.

Remark 2.2 In [8, pages 93–94], Leinster defines the notion of a lax monoidal category, which is similar in spirit to the above definition, but there are some crucial differences. A lax monoidal category in his sense, has multiplications

$$\mu_k(A_1, A_2, \ldots, A_k) = (A_1 \otimes A_2 \otimes \cdots \otimes A_k)$$

for all $k \in \mathbb{N}$ together with natural transformations

$$\gamma^{k_1,\ldots,k_n}\colon \mu_n \circ (\mu_{k_1} \otimes \ldots \otimes \mu_{k_n}) \to \mu_{k_1+\ldots+k_n}$$

and a natural transformation

$$\iota_A\colon A \longrightarrow \mu_1(A) = (A).$$

The natural transformations γ satisfy a coherence condition which is essentially our coherence conditions (v) and (vi) combined into a single diagram. There is no unit condition for μ_0 (so one might as well require the existence of μ_k for $k>0$ only). Moreover his natural transformation ι is not the identity. Thus a lax monoidal category in his sense possesses arbitrarily long nondegenerate strings of composable natural transformations between unary multiplications

$$A \xrightarrow{\iota_A} (A) \xrightarrow{\iota(A)} ((A)) \xrightarrow{\iota((A))} \cdots$$

It follows that the operad controlling such a structure has an infinite dimensional nerve.

The main result of this paper is:

Theorem 2.3 *A category $\mathcal{C}$ is a $\mathfrak{K}^{(n)}$–algebra if and only if it is an A_n–monoidal category.*

Proof Given an action, $\theta_i\colon \mathfrak{K}_i^{(n)}\times\mathcal{C}^i \longrightarrow \mathcal{C}$, define $\mu_i\colon \mathcal{C}^i \longrightarrow \mathcal{C}$ to be the restriction of this action to $\{x_1x_2\dots x_i\}\times\mathcal{C}^i$, where $x_1x_2\dots x_i=(W_i,\varnothing)$ is the terminal object of $\mathfrak{K}_i$. This makes sense for $0\le i<n+1$, since in those cases $x_1x_2\dots x_i$ is contained in the n–th filtration $\mathfrak{K}^{(n)}$. We then define $\alpha^{i,j,k}$ to be the restriction of θ_{i+j+k} to $\{(W_{i+j+k},\{[x_{i+1},x_{i+j}]\})\longrightarrow(W_{i+j+k},\varnothing)\}\times\mathcal{C}^{i+j+k}$, for $0\le i+j+k<n+1$.

Conditions (1), (2) and (i) follow from the fact that $(x_1)\in\mathfrak{K}_1^{(n)}$ is the identity of the operad and composing the constant $0\in\mathfrak{K}_0^{(n)}$ into any input of $\{x_1x_2\dots x_i\}\in\mathfrak{K}_i^{(n)}$ gives $\{x_1x_2\dots x_{i-1}\}\in\mathfrak{K}_{i-1}^{(n)}$. Conditions (ii)–(iv) also follow from the latter fact. Finally conditions (v) and (vi) follow from the restriction of $\theta_{a+b+c+d+e}$ to $\mathcal{D}\times\mathcal{C}^{a+b+c+d+e}$ and $\mathcal{D}'\times\mathcal{C}^{a+b+c+d+e}$, where $\mathcal{D}$ and $\mathcal{D}'$ are the following commutative diagrams in $\mathfrak{K}^{(n)}$ with $r=a+b+c+d+e$

$$\begin{array}{ccc}
(W_r,\{[x_{a+1},x_{a+b}],[x_{a+b+c+1},x_{a+b+c+d}]\}) & \longrightarrow & (W_r,\{[x_{a+1},x_{a+b}]\}) \\
\downarrow & & \downarrow \\
(W_r,\{[x_{a+b+c+1},x_{a+b+c+d}]\}) & \longrightarrow & (W_r,\varnothing),
\end{array}$$

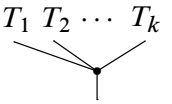
$$\begin{array}{ccc}
(W_r,\{[x_{a+1},x_{a+b+c+d}],[x_{a+b+1},x_{a+b+c}]\}) & \longrightarrow & (W_r,\{[x_{a+b+1},x_{a+b+c}]\}) \\
\downarrow & & \downarrow \\
(W_r,\{[x_{a+1},x_{a+b+c+d}]\}) & \longrightarrow & (W_r,\varnothing),
\end{array}$$

respectively.

Conversely suppose that $\mathcal{C}$ is an A_n–monoidal category. Then we define

$$\theta_i\colon \mathrm{Obj}(\mathfrak{K}_i^{(n)}) \times \mathcal{C}^i \longrightarrow \mathcal{C}$$

by induction on i as follows. We define θ_0 to be μ_0 and θ_1 to be $\mu_1 = \mathrm{id}_{\mathcal{C}}$. Having defined θ_j for $j < i$, consider an object T in $\mathfrak{K}_i^{(n)}$ represented by a tree



with $k < n+1$ and where T_j has m_j input edges, so that $m_1+m_2+\cdots+m_k = i$. Let $(\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_k)$ represent an object in $\mathcal{C}^i$, with $\bar{A}_j \in \mathrm{Obj}\,(\mathcal{C}^{m_j})$, $j = 1, 2, \ldots, k$. By induction $\theta_{m_j}(T_j, \bar{A}_j)$ are already defined for $j = 1, 2, \ldots, k$. We then define

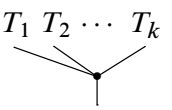
$$\theta_i\left(T, \bar{A}_1, \bar{A}_2, \ldots, \bar{A}_k\right) = \mu_k(\theta_{m_1}(T_1, \bar{A}_1), \theta_{m_2}(T_2, \bar{A}_2), \ldots, \theta_{m_k}(T_k, \bar{A}_k))$$

[Here we use implicitly the canonical isomorphisms between the associahedral posets based on subintervals of W_i with the associahedral posets based on the standard linear orders W_{m_i} of the same cardinality; cf Definition 1.2.] We define θ_i for morphisms in $\mathcal{C}^i$ similarly. This completes the induction.

Next we extend the definition of θ_i to define natural transformations

$$\theta_i\colon \mathrm{IMor}(\mathfrak{K}_i^{(n)}) \times \mathrm{Obj}(\mathcal{C}^i) \longrightarrow \mathrm{Mor}(\mathcal{C})$$

where $\mathrm{IMor}(\mathfrak{K}_i^{(n)})$ are the indecomposable morphisms in $\mathfrak{K}_i^{(n)}$, ie morphisms which can't be factored nontrivially (or equivalently morphisms given by dropping a single pair of matching parentheses in a parenthesized word). Again we proceed by induction on i, starting with $i = 0$ and $i = 1$ where these are vacuously defined. Now consider an indecomposable morphism $\lambda\colon T \to T'$ in $\mathfrak{K}_i^{(n)}$, where T has the form



with $k < n+1$ and where T_j has m_j input edges, so that $m_1+m_2+\cdots+m_k = i$. Then λ is obtained by shrinking a single interior edge in T. There are two possibilities: (1) an interior edge of some tree T_j is shrunk or (2) an edge below some T_j is shrunk. Now let $(\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_k)$ represent an object in $\mathcal{C}^i$, with $\bar{A}_j \in \mathrm{Obj}\,(\mathcal{C}^{m_j})$, $j = 1, 2, \ldots, k$.

In the first case we define

$$\begin{aligned}\theta_i(\lambda, \bar{A}_1, \bar{A}_2, \dots, \bar{A}_k) \\ = \mu_k\big(&\mathrm{id}_{\theta_{m_1}(T_1,\bar{A}_1)}, \mathrm{id}_{\theta_{m_2}(T_2,\bar{A}_2)}, \dots, \mathrm{id}_{\theta_{m_{j-1}}(T_{j-1},\bar{A}_{j-1})}, \\ &\theta_{m_j}(\lambda', \bar{A}_j), \mathrm{id}_{\theta_{m_{j+1}}(T_{j+1},\bar{A}_{j+1})}, \dots, \mathrm{id}_{\theta_{m_k}(T_k,\bar{A}_k)}\big)\end{aligned}$$

where λ' is the indecomposable morphism in $\mathfrak{K}^{(n)}_{m_j}$ given by shrinking that particular edge. In the second case we define

$$\theta_i(\lambda, \bar{A}_1, \bar{A}_2, \dots, \bar{A}_k) = \alpha^{m_1+m2+\dots m_{j-1}, m_j, m_{j+1}+m_{j+2}+\cdots+m_k}_{(\bar{A}_1,\bar{A}_2,\dots,\bar{A}_{j-1}),\bar{A}_j,(\bar{A}_{j+1},\dots,\bar{A}_k)}.$$

This completes the induction.

Finally to extend θ_i to all morphisms in $\mathfrak{K}_i$, we must show that for any factorization of a morphism in $\mathfrak{K}^{(n)}_i$ into indecomposable morphisms, the corresponding composition of natural transformations defines the same morphism in $\mathcal{C}$. But according to Lemma 1.4, the factorizations of any morphism in $\mathfrak{K}_i$ give rise to a cubical diagram in $\mathcal{C}$. According to coherence conditions (v) and (vi) of an A_n–monoidal category, all the 2–dimensional faces of this cubical diagram commute. It is an elementary consequence that the entire cubical diagram in $\mathcal{C}$ commutes; cf Lemma 2.4 below. It follows that there are well defined functors:

$$\theta_i\colon \mathfrak{K}^{(n)}_i \times \mathcal{C}^i \longrightarrow \mathcal{C}$$

for all $i \geq 0$. The fact that θ_i are compatible with the operadic compositions

$$\gamma_{m,k_1,k_2,\dots,k_m}\colon \mathfrak{K}^{(n)}_m \times \prod_1^m \mathfrak{K}^{(n)}_{k_i} \to \mathfrak{K}^{(n)}_{\sum_1^m k_i}$$

follows from the inductive construction of θ_i if all the $k_i > 1$. If $k_i \leq 1$ or $m = 1$, the compatibility follows from conditions (1), (2) and (i)–(iv) of the definition of an A_n–monoidal category. □

Lemma 2.4 *A cubical diagram in any category commutes if and only if each of its 2–dimensional faces commutes.*

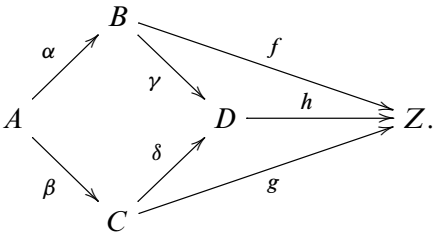
Proof We proceed by induction on the dimension of the cube. The statement is vacuously true if the dimension is ≤ 2. Suppose it is true for all cubical diagrams of dimension $< m$, and suppose we are given an m–dimensional cubical diagram. Consider two edge paths from the initial object A of the diagram to Z, the terminal object. Let these edge paths factor as

$$A \xrightarrow{\alpha} B \xrightarrow{f} Z, \qquad A \xrightarrow{\beta} C \xrightarrow{g} Z,$$

respectively, where α and β are edges of the diagram and f and g are composites of the remainders of these edge paths. If $\alpha = \beta$, then by induction $f = g$ and we are done. Otherwise let

$$\begin{array}{ccc} A & \xrightarrow{\alpha} & B \\ \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} \\ C & \xrightarrow{\delta} & D \end{array}$$

be the 2–dimensional face spanned by α and β. Pick any edge path $h\colon D \longrightarrow Z$, and consider the diagram



By hypothesis the square commutes and by induction the two triangles commute. Hence $f\alpha = g\beta$. This completes the induction and proof. □

3 Relation to coherence theorems for monoidal categories

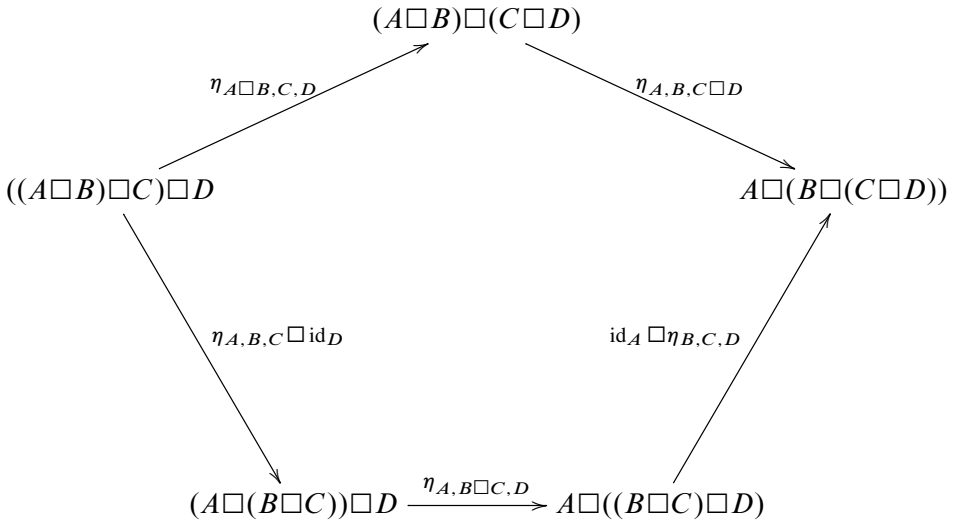
Definition 3.1 We say that an A_n–monoidal category is *undirected* if all the associativity natural transformations $\alpha^{i,j,k}$ are isomorphisms.

Proposition 3.2 *An undirected A_n–monoidal category is a monoidal category if $n \geq 4$.*

Proof If $\mathcal{C}$ is an undirected A_n–monoidal category, then the corresponding action functors $\theta_i\colon \mathfrak{K}_i^{(n)} \times \mathcal{C}^i \longrightarrow \mathcal{C}$ extend to $\theta_i\colon \overline{\mathfrak{K}}_i^{(n)} \times \mathcal{C}^i \longrightarrow \mathcal{C}$, where $\overline{\mathfrak{K}}_i^{(n)}$ is obtained from $\mathfrak{K}_i^{(n)}$ by formally inverting all the morphisms. By Proposition 1.6 the nerve of $\mathfrak{K}_i^{(n)}$ is simply connected. Now recalling that inverting all the morphisms in a connected category has the effect of killing off the higher homotopy groups of its nerve (cf [11, Proposition 1]), we see that the nerve of $\overline{\mathfrak{K}}_i^{(n)}$ is contractible, and it follows that the objects of $\mathfrak{K}_i^{(2)}$ are connected to each other by uniquely defined isomorphisms in $\overline{\mathfrak{K}}_i^{(n)}$. The images of these isomorphisms under θ_i specify uniquely defined natural isomorphisms connecting all possible different ways of associating the binary product $\mu_2\colon \mathcal{C}^2 \longrightarrow \mathcal{C}$ into an i–fold product $\mathcal{C}^i \longrightarrow \mathcal{C}$ so that all diagrams involving them commute. Thus $\mathcal{C}$ is a monoidal category, in the classical sense of Mac Lane. □

Next we derive LaPlaza's coherence theorem [7], which generalizes Mac Lane's coherence theorem to the case where the associativity natural transformation for a monoidal structure on a category is not required to be an isomorphism. We begin with a preliminary version of this result.

Theorem 3.3 *Let $(\mathcal{C}, \Box, 0, \eta)$ be a directed monoidal category with a strict unit. That is, $\Box\colon \mathcal{C}\times\mathcal{C}\longrightarrow\mathcal{C}$ is a bifunctor and 0 is an object of $\mathcal{C}$ which serves as a strict unit for $\Box$, ie the restrictions of $\Box$ to $0\times\mathcal{C}$ and $\mathcal{C}\times 0$ are the identity. Finally $\eta_{A,B,C}\colon (A\Box B)\Box C\longrightarrow A\Box(B\Box C)$ is a natural transformation (not necessarily an isomorphism) such that $\eta_{A,B,C}$ is the identity whenever one of A, B, C is 0 and such that the pentagonal diagram*



commutes. Then $\mathcal{C}$ can be endowed with the structure of an A_∞–monoidal category.

Proof We define $\mu_0(*)=0$, μ_1 to be the identity, $\mu_2=\Box$ and then we inductively define μ_i to be the composite

$$\mathcal{C}^i=\mathcal{C}\times\mathcal{C}^{i-1}\xrightarrow{\mathrm{id}_{\mathcal{C}}\times\mu_{i-1}}\mathcal{C}\times\mathcal{C}\xrightarrow{\Box}\mathcal{C}$$

Thus $\mu_i(A_1,A_2,A_3,\dots,A_i)=A_1\Box(A_2\Box(A_3\Box(\cdots\Box(A_{i-1}\Box A_i)\cdots)))$ and

$$(*)\qquad \mu_{a+b}(\bar{A},\bar{B})=\mu_{a+1}(\bar{A},\mu_b(\bar{B}))$$

for any objects $\bar{A}\in\mathcal{C}^a$, $\bar{B}\in\mathcal{C}^b$.

Now let $\bar{B}\in\mathcal{C}^b$ and $\bar{C}\in\mathcal{C}^c$. We define the associativity $\alpha^{0,b,c}_{\bar{B},\bar{C}}$ inductively on b. We assume $c>0$, since $\alpha^{0,b,0}$ is required by definition to be the identity. For $b=1$, we

also require $\alpha^{0,1,c}$ to be the identity. So suppose $b>1$ and $\bar{B}=(B_1,\bar{B}')$. Then we define $\alpha^{0,b,c}$ to be the composite

$$\mu_{c+1}(\mu_b(\bar{B}),\bar{C})=\mu_2(\mu_b(\bar{B}),\mu_c(\bar{C}))=(B_1\square\mu_{b-1}(\bar{B}'))\square\mu_c(\bar{C})$$
$$\xrightarrow{\eta_{B_1,\mu_{b-1}(\bar{B}'),\mu_c(\bar{C})}} B_1\square(\mu_{b-1}(\bar{B}')\square\mu_c(\bar{C}))$$

followed by the composite

$$B_1\square(\mu_{b-1}(\bar{B}')\square\mu_c(\bar{C}))=B_1\square\mu_2(\mu_{b-1}(\bar{B}'),\mu_c(\bar{C}))=B_1\square\mu_{c+1}(\mu_{b-1}(\bar{B}'),\bar{C})$$
$$\xrightarrow{\mathrm{id}_{B_1}\square\,\alpha^{0,b-1,c}_{\bar{B}',\bar{C}}} B_1\square\mu_{b+c-1}(\bar{B}',\bar{C})=\mu_{b+c}(\bar{B},\bar{C}).$$

This completes the inductive definition of $\alpha^{0,b,c}_{\bar{B},\bar{C}}$. We then define $\alpha^{a,b,c}_{\bar{A},\bar{B},\bar{C}}$ to be the composite

$$\mu_{a+c+1}(\bar{A},\mu_b(\bar{B}),\bar{C})=\mu_{a+1}(\bar{A},\mu_{c+1}(\mu_b(\bar{B}),\bar{C}))$$
$$\xrightarrow{\mu_{a+1}(\mathrm{id}_{\bar{A}},\alpha^{0,b,c}_{\mu_b(\bar{B}),\bar{C}})}\mu_{a+1}(\bar{A},\mu_{b+c}(\bar{B},\bar{C}))=\mu_{a+b+c}(\bar{A},\bar{B},\bar{C}).$$

Note that this implies that $\alpha^{a,b,c}_{\bar{A},\bar{B},\bar{C}}$ is the identity if $c=0$, and that

$$(**)\qquad \alpha^{a_1+a_2,b,c}_{(\bar{A}_1,\bar{A}_2),\bar{B},\bar{C}}=\mu_{a_1+1}(\mathrm{id}_{A_1},\alpha^{a_2,b,c}_{\bar{A}_2,\bar{B},\bar{C}}).$$

Conditions (1), (2), (i)–(iv) for an A_∞–monoidal category are either true by construction or follow by a straight forward induction argument using the hypotheses that 0 is a strict unit for $\square$ and that $\eta_{A,B,C}$ is the identity whenever one of A, B or C is 0.

By $(**)$ the verification of condition (v) reduces to the special case of the diagram

$$\begin{array}{ccc}
\mu_{c+e+2}(\mu_b(\bar{B}),\bar{C},\mu_d(\bar{D}),\bar{E}) & \xrightarrow{\alpha^{0,b,c+e+1}_{\bar{B},(\bar{C},\mu_d(\bar{D}),\bar{E})}} & \mu_{b+c+e+1}(\bar{B},\bar{C},\mu_d(\bar{D}),\bar{E}) \\
\downarrow{\scriptstyle \alpha^{c+1,d,e}_{(\mu_b(\bar{B}),\bar{C}),\bar{D},\bar{E}}} & & \downarrow{\scriptstyle \alpha^{c+1,d,e}_{(\bar{B},\bar{C}),\bar{D},\bar{E}}} \\
\mu_{c+d+e+1}(\mu_b(\bar{B}),\bar{C},\bar{D},\bar{E}) & \xrightarrow{\alpha^{0,b,c+d+e}_{\bar{B},(\bar{C},\bar{D},\bar{E})}} & \mu_{b+c+d+e}(\bar{B},\bar{C},\bar{D},\bar{E})
\end{array}$$

since the general diagram for (v) can be obtained from this one by applying the functor $\mu_{a+1}(\bar{A},-)$ to it. By $(*)$ and $(**)$, this diagram in turn is the same as the diagram

$$\begin{array}{ccc}
\mu_2(\mu_b(\bar{B}),\mu_{c+e+1}(\bar{C},\mu_d(\bar{D}),\bar{E})) & \xrightarrow{\alpha^{0,b,1}_{\bar{B},\mu_{c+e+1}(\bar{C},\mu_d(\bar{D}),\bar{E})}} & \mu_{b+1}(\bar{B},\mu_{c+e+1}(\bar{C},\mu_d(\bar{D}),\bar{E})) \\
\downarrow{\scriptstyle \mu_2(\mathrm{id}_{\mu_b(\bar{B})},\alpha^{c,d,e}_{\bar{C},\bar{D},\bar{E}})} & & \downarrow{\scriptstyle \mu_{b+1}(\mathrm{id}_{\bar{B}},\alpha^{c,d,e}_{\bar{C},\bar{D},\bar{E}})} \\
\mu_2(\mu_b(\bar{B}),\mu_{c+d+e}(\bar{C},\bar{D},\bar{E})) & \xrightarrow{\alpha^{0,b,1}_{\bar{B},\mu_{c+d+e}(\bar{C},\bar{D},\bar{E})}} & \mu_{b+1}(\bar{B},\mu_{c+d+e}(\bar{C},\bar{D},\bar{E}))
\end{array}$$

This last diagram in turn commutes because

$$\alpha^{0,b,1}_{\bar{B},X}\colon \mu_2(\mu_b(\bar{B}),X)\longrightarrow \mu_{b+1}(\bar{B},X)$$

is a natural transformation.

By similar reasoning, the verification of condition (vi) reduces to the special case of the diagram

$$\begin{array}{ccc}
\mu_{e+1}(\mu_{b+d+1}(\bar{B},\mu_c(\bar{C}),\bar{D}),\bar{E}) & \xrightarrow{\mu_{e+1}(\alpha^{b,c,d}_{\bar{B},\bar{C},\bar{D}},\mathrm{id}_{\bar{E}})} & \mu_{e+1}(\mu_{b+c+d}(\bar{B},\bar{C},\bar{D}),\bar{E}) \\
\downarrow{\scriptstyle \alpha^{0,b+d+1,e}_{(\bar{B},\mu_c(\bar{C}),\bar{D}),\bar{E}}} & & \downarrow{\scriptstyle \alpha^{0,b+c+d,e}_{(\bar{B},\bar{C},\bar{D}),\bar{E}}} \\
\mu_{b+d+e+1}(\bar{B},\mu_c(\bar{C}),\bar{D},\bar{E}) & \xrightarrow{\alpha^{b,c,d+e}_{\bar{B},\bar{C},(\bar{D},\bar{E})}} & \mu_{b+c+d+e}(\bar{B},\bar{C},\bar{D},\bar{E}).
\end{array}$$

By $(*)$ it follows that this diagram is unchanged if we replace E throughout by $\mu_e(E)$. Hence we may as well suppose that $e=1$ and $\bar{E}=E$ is an object of $\mathcal{C}$. Then by the inductive definition of $\alpha^{0,i,j}$ and $(**)$, we can factor the diagram as follows:

$$\begin{array}{ccc}
(B_1\square\mu_{b+d}(\bar{B}',\mu_c(\bar{C}),\bar{D}))\square E & \xrightarrow{(\mathrm{id}_{B_1}\square\alpha^{b-1,c,d}_{\bar{B}',\bar{C},\bar{D}})\square\,\mathrm{id}_E} & (B_1\square\mu_{b+c+d-1}(\bar{B}',\bar{C},\bar{D}))\square E \\
\downarrow{\scriptstyle \eta_{B_1,\mu_{b+d}(\bar{B}',\mu_c(\bar{C}),\bar{D}),E}} & & \downarrow{\scriptstyle \eta_{B_1,\mu_{b+c+d-1}(\bar{B}',\bar{C},\bar{D}),E}} \\
B_1\square(\mu_{b+d}(\bar{B}',\mu_c(\bar{C}),\bar{D})\square E) & \xrightarrow{\mathrm{id}_{B_1}\square(\alpha^{b-1,c,d}_{\bar{B}',\bar{C},\bar{D}}\square\,\mathrm{id}_E)} & B_1\square(\mu_{b+c+d-1}(\bar{B}',\bar{C},\bar{D})\square E) \\
\downarrow{\scriptstyle \mathrm{id}_{B_1}\square\alpha^{0,b+d,1}_{(\bar{B}',\mu_c(\bar{C}),\bar{D}),\bar{E}}} & & \downarrow{\scriptstyle \mathrm{id}_{B_1}\square\alpha^{0,b+c+d-1,1}_{(\bar{B},\bar{C},\bar{D}),E}} \\
B_1\square\mu_{b+d+1}(\bar{B}',\mu_c(\bar{C}),\bar{D},E) & \xrightarrow{\mathrm{id}_{B_1}\square\alpha^{b-1,c,d+1}_{\bar{B},\bar{C},(\bar{D},E)}} & B_1\square\mu_{b+c+d}(\bar{B}',\bar{C},\bar{D},E)
\end{array}$$

The upper square commutes by naturality of η. The commutativity of the lower square corresponds to a reduction of the problem from b to $b-1$. Recursing on this reduction we reduce to the case $b=0$, ie showing that the diagram

$$\begin{array}{ccc}
\mu_{d+1}(\mu_c(\bar{C}),\bar{D})\square E & \xrightarrow{\alpha^{0,c,d}_{\bar{C},\bar{D}}\square\,\mathrm{id}_E} & \mu_{c+d}(\bar{C},\bar{D})\square E \\
\downarrow{\scriptstyle \alpha^{0,d+1,1}_{(\mu_c(\bar{C}),\bar{D}),E}} & & \downarrow{\scriptstyle \alpha^{0,c+d,1}_{(\bar{C},\bar{D}),E}} \\
\mu_{d+2}(\mu_c(\bar{C}),\bar{D},E) & \xrightarrow{\alpha^{0,c,d+1}_{\bar{C},(\bar{D},E)}} & \mu_{c+d+1}(\bar{C},\bar{D},E)
\end{array}$$

commutes. By $(*)$ and the inductive definition of $\alpha^{0,i,j}$ this diagram can be replaced

and expanded into the following diagram:

$$\begin{array}{ccc}
\mu_2(\mu_c(\bar{C}),\mu_d(\bar{D}))\square E & \xrightarrow{\alpha^{0,c,1}_{\bar{C},\mu_d(\bar{D})}\square\,\mathrm{id}_E} & \mu_{c+1}(\bar{C},\mu_d(\bar{D}))\square E \\
\downarrow{\scriptstyle \alpha^{0,2,1}_{(\mu_c(\bar{C}),\mu_d(\bar{D})),E}} & & \downarrow{\scriptstyle \alpha^{0,c+1,1}_{(\bar{C},\mu_d(\bar{D})),E}} \\
\mu_3(\mu_c(\bar{C}),\mu_d(\bar{D}),E) & \xrightarrow{\alpha^{0,c,2}_{\bar{C},(\mu_d(\bar{D}),E)}} & \mu_{c+2}(\bar{C},\mu_d(\bar{D}),E) \\
\downarrow{\scriptstyle \mu_2(\mathrm{id}_{\mu_c(\bar{C})},\alpha^{0,d,1}_{\bar{D},E})} & & \downarrow{\scriptstyle \mu_{c+1}(\mathrm{id}_{\bar{C}},\alpha^{0,d,1}_{\bar{D},E})} \\
\mu_2(\mu_c(\bar{C}),\mu_{d+1}(\bar{D},E)) & \xrightarrow{\alpha^{0,c,1}_{\bar{C},\mu_{d+1}(\bar{D},E)}} & \mu_{c+1}(\bar{C},\mu_{d+1}(\bar{D},E))
\end{array}$$

The lower square commutes by naturality of α. So it suffices to show the upper square commutes. This is just the previous diagram with $\bar{D}$ replaced by $\mu_d(\bar{D})$. Thus we have reduced to the case $d = 1$. We will find it convenient to display this diagram in reflected form:

$$\begin{array}{ccc}
\mu_2(\mu_c(\bar{C}),D)\square E & \xrightarrow{\alpha^{0,2,1}_{(\mu_c(\bar{C}),D),E}} & \mu_3(\mu_c(\bar{C}),D,E) \\
\downarrow{\scriptstyle \alpha^{0,c,1}_{\bar{C},D}\square\,\mathrm{id}_E} & & \downarrow{\scriptstyle \alpha^{0,c,2}_{\bar{C},(D,E)}} \\
\mu_{c+1}(\bar{C},D)\square E & \xrightarrow{\alpha^{0,c+1,1}_{(\bar{C},D),E}} & \mu_{c+2}(\bar{C},D,E)
\end{array}$$

We have to show this diagram commutes, where D and E are objects of $\mathcal{C}$ and $\bar{C}$ is an object of $\mathcal{C}^c$.

This diagram commutes trivially if $c \le 1$. So assume $c > 1$ and $\bar{C} = (C_1, \bar{C}')$. Again using $(*)$ and the inductive definition of $\alpha^{0,i,j}$, we can expand this diagram into:

$$\begin{array}{ccc}
\mu_2(\mu_2(C_1,\mu_{c-1}(\bar{C}')),D)\square E & \xrightarrow{\alpha^{0,2,1}_{(\mu_2(C_1,\mu_{c-1}(\bar{C}')),D),E}} & \mu_3(\mu_2(C_1,\mu_{c-1}(\bar{C}')),D,E) \\
\downarrow{\scriptstyle \alpha^{0,2,1}_{\mu_2(C_1,\mu_{c-1}(\bar{C}')),D}\square\,\mathrm{id}_E} & & \downarrow{\scriptstyle \alpha^{0,2,2}_{(C_1,\mu_{c-1}(\bar{C}')),(D,E)}} \\
\mu_3(C_1,\mu_{c-1}(\bar{C}'),D)\square E & \xrightarrow{\alpha^{0,3,1}_{(C_1,\mu_{c-1}(\bar{C}'),D),E}} & \mu_4(C_1,\mu_{c-1}(\bar{C}'),D,E) \\
\| & & \| \\
\mu_2(C_1,\mu_2(\mu_{c-1}(\bar{C}'),D))\square E & & \mu_2(C_1,\mu_3(\mu_{c-1}(\bar{C}'),D,E)) \\
\downarrow{\scriptstyle \mu_2(\mathrm{id}_{C_1},\alpha^{0,1,1}_{\mu_{c-1}(\bar{C}'),D})\square\,\mathrm{id}_E} & & \downarrow{\scriptstyle \mu_2(\mathrm{id}_{C_1},\alpha^{0,1,2}_{\mu_{c-1}(\bar{C}'),(D,E)})} \\
\mu_2(C_1,\mu_c(\bar{C}',D))\square E & & \mu_2(C_1,\mu_{c+1}(\bar{C}',D,E)) \\
\| & & \| \\
\mu_{c+1}(\bar{C},D)\square E & \xrightarrow{\alpha^{0,c+1,1}_{(\bar{C},D),E}} & \mu_{c+2}(\bar{C},D,E)
\end{array}$$

The top square in this diagram is the original diagram with $\bar{C}$ replaced by $(C_1, \mu_{c-1}(\bar{C}'))$, thus reducing it to the case $c = 2$. This top square can be expanded into the pentagonal diagram of the hypothesis of the theorem and thus commutes. It remains to show that the bottom square commutes.

After rewriting the bottom square in reflected form and applying the inductive definition of $\alpha^{0,i,j}$ we obtain the following expanded diagram:

$$\begin{array}{ccc}
(C_1\Box(\mu_{c-1}(\bar{C}')\Box D))\Box E & \xrightarrow{(\mathrm{id}_{C_1}\Box\alpha^{0,1,1}_{\mu_{c-1}(\bar{C}'),D})\Box\,\mathrm{id}_E} & (C_1\Box\mu_c(\bar{C}', D))\Box E \\
\downarrow{\scriptstyle \eta_{C_1,\mu_{c-1}(\bar{C}')\Box D,E}} & & \downarrow{\scriptstyle \eta_{C_1,\mu_c(\bar{C}',D),E}} \\
C_1\Box((\mu_{c-1}(\bar{C}')\Box D)\Box E) & \xrightarrow{\mathrm{id}_{C_1}\Box(\alpha^{0,1,1}_{\mu_{c-1}(\bar{C}'),D}\Box\,\mathrm{id}_E)} & C_1\Box(\mu_c(\bar{C}', D)\Box E) \\
\downarrow{\scriptstyle \mathrm{id}_{C_1}\Box\alpha^{0,2,1}_{(\mu_{c-1}(\bar{C}'),D),E}} & & \downarrow{\scriptstyle \mathrm{id}_{C_1}\Box\alpha^{0,c,1}_{(\bar{C}',D),E}} \\
C_1\Box\mu_3(\mu_{c-1}(\bar{C}'), D, E) & \xrightarrow{\mathrm{id}_{C_1}\Box\alpha^{0,c-1,2}_{\bar{C}',(D,E)}} & C_1\Box\mu_{c+1}(\bar{C}', D, E)
\end{array}$$

The top square commutes by naturality of η and the bottom square by induction hypothesis.

This completes the verification that we have constructed an A_∞–monoidal structure on $\mathcal{C}$. □

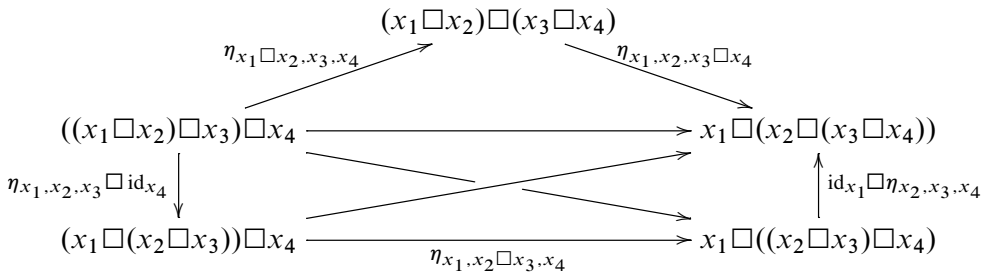
To obtain the full version of LaPlaza's coherence theorem, we start with an operadic reformulation of Theorem 3.3.

Definition 3.4 The LaPlaza operad $\mathcal{L} = \{\mathcal{L}_m\}_{m\geq 0}$ is the operad in *CAT* which acts on directed monoidal categories as in the hypothesis of Theorem 3.3. Specifically $\mathcal{L}_m$ can be described as a full subcategory of the free directed monoidal category on m generating objects $\{x_1, x_2, \ldots, x_m\}$, whose objects look like $x_1\Box x_2\Box\ldots\Box x_m$ after removing all parentheses. Thus $\mathcal{L}_0 = \{0\}$, $\mathcal{L}_1 = \{x_1\}$, and for $m \geq 2$ the objects of $\mathcal{L}_m$ are in bijective correspondence with planar binary trees with m input edges.

Remark 3.5 $\mathcal{L}_2$ is the trivial poset $\{x_1\Box x_2\}$, $\mathcal{L}_3$ is the poset

$$\eta_{x_1,x_2,x_3}\colon (x_1\Box x_2)\Box x_3 \longrightarrow x_1\Box(x_2\Box x_3),$$

isomorphic to $\mathcal{I}$, while $\mathcal{L}_4$ is the pentagonal poset generated by the labelled arrows shown below.



LaPlaza's coherence theorem states that $\mathcal{L}_m$ is a poset for all m.

Remark 3.6 LaPlaza works with natural transformations $\eta_{A,B,C}\colon A\square(B\square C)\longrightarrow(A\square B)\square C$. Moreover, he does not consider units. So, if $\mathcal{L}^*\subset\mathcal{L}$ is the suboperad obtained from $\mathcal{L}$ by dropping the unit, the operad $\mathcal{L}^*$ is dual to LaPlaza's original one. Note also that $\mathcal{L}_m=\mathcal{L}^*_m$ for $m\geq 1$. Our poset $\mathcal{L}_m$, $m\geq 3$ is precisely the poset considered by Tamari [16], and is now commonly called the *Tamari lattice* [17].

With this notation, we can reformulate Theorem 3.3 as follows.

Theorem 3.7 *There is a map of CAT–operads*

$$\Lambda\colon\mathfrak{K}\longrightarrow\mathcal{L}$$

which is a surjection.

The existence of Λ is clear from the statement of Theorem 3.3. Surjectivity follows from the proof of Theorem 3.3, where it is shown that

$$(x_1x_2)x_3\longrightarrow x_1x_2x_3$$

maps via Λ to

$$\eta_{x_1,x_2,x_3}\colon(x_1\square x_2)\square x_3\longrightarrow x_1\square(x_2\square x_3),$$

and the fact that η_{x_1,x_2,x_3} generates $\mathcal{L}$ as a CAT–operad.

LaPlaza's coherence theorem is not immediately apparent from Theorem 3.7, since a quotient category of a poset need not be a poset. We need the following additional observation.

Lemma 3.8 *For any object $T\in\mathcal{L}_m$, the inverse image under Λ of the subcategory $\{T\}$ is a subposet of $\mathfrak{K}_m$ containing both a minimal and a maximal object.*

Proof For $m = 0, 1, 2$, the functor Λ is an isomorphism and there is nothing to prove. For $m \geq 3$, we may regard T as a planar binary tree. Clearly the minimal object of $\Lambda^{-1}\{T\}$ is T regarded as an object of $\mathfrak{K}_m$. The maximal object of $\Lambda^{-1}\{T\}$ is obtained from T by successively shrinking the rightmost incoming edge to every node of T, with the exception of those edges which are leaves, till the rightmost edge of each node is a leaf. □

Example 3.9 The inverse images in Lemma 3.8 for $\Lambda\colon \mathfrak{K}_4 \longrightarrow \mathcal{L}_4$ are as follows:

$$\Lambda^{-1}\{((x_1\square x_2)\square x_3)\square x_4\} = \Big\{((x_1x_2)x_3)x_4 = \Big\},$$

$$\Lambda^{-1}\{(x_1\square x_2)\square(x_3\square x_4)\} = \Big\{(x_1x_2)(x_3x_4) = \longrightarrow (x_1x_2)x_3x_4 = \Big\},$$

$$\Lambda^{-1}\{(x_1\square(x_2\square x_3))\square x_4\} = \Big\{(x_1(x_2x_3))x_4 = \longrightarrow (x_1x_2x_3)x_4 = \Big\},$$

$$\Lambda^{-1}\{x_1\square((x_2\square x_3)\square x_4)\} = \Big\{x_1((x_2x_3)x_4) = \longrightarrow x_1(x_2x_3)x_4 = \Big\},$$

$$\Lambda^{-1}\{x_1\square(x_2\square(x_3\square x_4))\}$$

$$= \Big\{ x_1(x_2(x_3x_4)) \quad = \quad x_1(x_2x_3x_4) = \quad x_1x_2x_3x_4 = \quad x_1x_2(x_3x_4) = \Big\}$$

Corollary 3.10 (LaPlaza Coherence Theorem) *For all m, the category $\mathcal{L}_m$ is a poset (known as the Tamari lattice for $m \geq 3$).*

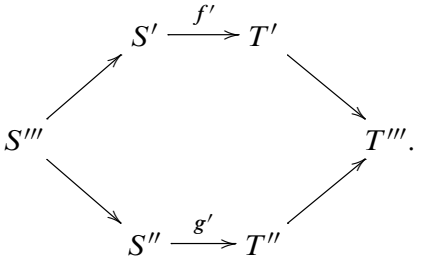
Proof Let

$$S \underset{g}{\overset{f}{\rightrightarrows}} T$$

be morphisms in $\mathcal{L}_m$. Since Λ is surjective, we may find preimages

$$f'\colon S' \longrightarrow T', \quad g'\colon S'' \longrightarrow T''$$

under Λ of f and g respectively. Now let S''' be the minimal element of $\Lambda^{-1}\{S\}$ and let T''' be the maximal element of $\Lambda^{-1}\{T\}$. Then since $\mathfrak{K}_m$ is a poset, we have a commutative diagram



Applying Λ to this diagram and noting that Λ sends the unlabelled arrows to identities, we obtain $f = g$. □

We will now give an explicit description of the posets $\mathcal{L}_m$ for $m \geq 4$. Similar considerations apply to $\mathfrak{K}_m$ and give an alternative description of those posets.

Definition 3.11 Let $\mathcal{O}$ be a CAT–operad with a single nullary operation $\mathcal{O}_0 = \{0\}$ (such as $\mathfrak{K}$ or $\mathcal{L}$). Suppose $m \geq 4$ and let $\{a < b < c\} \subset \{1, 2, 3, \ldots, m\}$. We define the functor $\pi_{a,b,c}\colon \mathcal{O}_m \longrightarrow \mathcal{O}_3$ to be the composite

$$\mathcal{O}_m \longrightarrow \mathcal{O}_m \times \prod_{i=1}^{m} \mathcal{O}_{k_i} \longrightarrow \mathcal{O}_3.$$

Here $k_i = 0$ if $i \notin \{a, b, c\}$, $k_a = k_b = k_c = 1$, the first map takes $\phi \in \mathcal{O}_m$ to $(c; \epsilon_1, \epsilon_2, \ldots, \epsilon_m)$, where $\epsilon_a = \epsilon_b = \epsilon_c = \mathrm{id} \in \mathcal{O}_1$ with all other $\epsilon_i = 0 \in \mathcal{O}_0$, and the second map is composition in $\mathcal{O}$.

Proposition 3.12 *There is a commutative diagram*

$$\begin{array}{ccccc}
\mathfrak{K}_m & \xrightarrow{\{\pi_{a,b,c}\}} & \prod_{\{1\leq a<b<c\leq m\}} \mathfrak{K}_3 & & \\
\downarrow{\scriptstyle \Lambda} & & \downarrow{\scriptstyle \prod \Lambda} & & \\
\mathcal{L}_m & \xrightarrow{\{\pi_{a,b,c}\}} & \prod_{\{1\leq a<b<c\leq m\}} \mathcal{L}_3 & \xrightarrow{\cong} & \prod_{\{1\leq a<b<c\leq m\}} \mathcal{I},
\end{array}$$

with the horizontal arrows being full imbeddings of posets.

The proof is straight forward and left as an exercise for the reader.

4 Rectification of A_∞–monoidal categories

It is well known that a monoidal category is equivalent to a strictly monoidal category; cf [9, pages 257–259]. (Recall that a monoidal category is strict if the associativity natural transformations $\eta_{A,B,C}$ of Theorem 3.3 are the identities.) We establish an analogous result for A_∞–monoidal categories.

We first need a preliminary construction.

Definition 4.1 For $k \geq 2$ we define the poset $\widehat{\mathfrak{K}}_k$ to have as objects combinatorial trees as defined in [8, Appendix E] with k input edges. All nodes except the root node, ie the node at the output of the tree, are required to have more than one incoming edge. The root node may have zero, one, or more incoming edges. We define $T < T'$ if T' can be obtained from T by shrinking some internal edges. We define $\widehat{\mathfrak{K}}_1$ to consist of the single tree:

and $\widehat{\mathfrak{K}}_0$ to consist of the single tree:

Example 4.2 The tree

is allowed in $\widehat{\mathfrak{K}}_2$, while

is not allowed.

The collection $\widehat{\mathfrak{K}} = \{\widehat{\mathfrak{K}}_k\}_{k\geq 0}$ is a right module over the associahedral operad, that is there are maps of posets

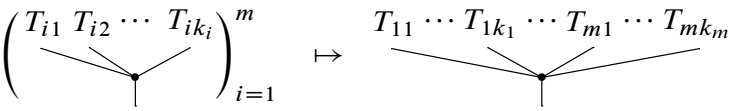
$$\widehat{\mathfrak{K}}_m \times \prod_{i=1}^{m} \mathfrak{K}_{k_i} \longrightarrow \widehat{\mathfrak{K}}_{k_1+k_2+\cdots+k_m}$$

satisfying the usual associativity and unit conditions. This right action is defined in exactly the same way as we defined the operad structure on $\mathfrak{K}$, with the single exception that when we compose with $0 \in \mathfrak{K}_0$, we never delete the root node. Moreover, $\widehat{\mathfrak{K}}$ is also

a left module over Ass, the trivial operad parametrizing strictly monoidal structures. The left action

$$\mathrm{Ass}(m)\times\prod_{i=1}^{m}\widehat{\mathfrak{K}}_{k_i}\cong\prod_{i=1}^{m}\widehat{\mathfrak{K}}_{k_i}\longrightarrow\widehat{\mathfrak{K}}_{k_1+k_2+\cdots+k_m}$$

is given by:



It is clear that the left and right actions commute with each other, so $\widehat{\mathfrak{K}}$ is an Ass–$\mathfrak{K}$–bimodule.

Theorem 4.3 *There is a functorial construction $\mathcal{C}\mapsto\mathcal{MC}$ together with functors $I\colon\mathcal{C}\longrightarrow\mathcal{MC}$ and $E\colon\mathcal{MC}\longrightarrow\mathcal{C}$ which associates to each A_∞–monoidal category $\mathcal{C}$ a strictly monoidal category $\mathcal{MC}$ such that*

(1) *the induced maps by I and E on the nerves of the categories are mutually inverse homotopy equivalences,*

(2) *the functor I induces a lax homomorphism of A_∞–spaces in the sense of* [2],

(3) *if $\mathcal{C}$ is strictly monoidal then E is a strictly monoidal functor.*

Proof Let

$$\mathcal{MC}=\widehat{\mathfrak{K}}\otimes_{\mathfrak{K}}\mathcal{C}=\Big(\coprod_{k\geq 0}\widehat{\mathfrak{K}}_k\times\mathcal{C}^k\Big)\Big/\approx$$

where the equivalence relation is given by

$$\big(T\circ(S_1,S_2,\dots,S_m),(\bar{A}_1,\bar{A}_2,\dots,\bar{A}_m)\big)\approx\big(T,S_1(\bar{A}_1),S_2(\bar{A}_2),\dots,S_m(\bar{A}_m)\big)\,,$$

where $T\in\widehat{\mathfrak{K}}_m$, $S_i\in\mathfrak{K}_{k_i}$, $\bar{A}_i\in\mathcal{C}^{k_i}$, for $i=1,2,\dots,m$. The left action of Ass on $\widehat{\mathfrak{K}}$ then induces a strict monoidal structure on $\mathcal{MC}$.

There are functors $I\colon\mathfrak{K}\longrightarrow\widehat{\mathfrak{K}}$ and $E\colon\widehat{\mathfrak{K}}\longrightarrow\mathfrak{K}$. The functor I takes a tree $S\in\mathfrak{K}$ to the tree:

The functor E takes a tree in $\widehat{\mathfrak{K}}_k$, deletes the root vertex if it has only one or no incoming edges, and regards it as a tree in $\mathfrak{K}$. The composite EI is the identity of $\mathfrak{K}$. There is a natural transformation from the composite IE to the identity of $\widehat{\mathfrak{K}}$,

given by shrinking the edge above the root vertex. The functors I, E and this natural transformation are compatible with the right actions of $\mathfrak{K}$ on $\widehat{\mathfrak{K}}$ and on itself. Hence there are induced functors

$$I\colon \mathcal{C} \longrightarrow \mathcal{MC}, \quad E\colon \mathcal{MC} \longrightarrow \mathcal{C}$$

such that EI is the identity of $\mathcal{C}$ and we have an induced natural transformation from IE to the identity of $\mathcal{MC}$. It follows that the maps induced by I and E on the nerves of these categories are mutually inverse equivalences.

Moreover the nerve of $\mathcal{MC}$ is homeomorphic to the topological construction described in [2, Theorem 1.26] applied to the nerve of $\mathcal{C}$. There it is shown that the map induced by I is a lax homomorphism of A_∞–spaces. The fact that E is strictly monoidal, when $\mathcal{C}$ is strictly monoidal, is straight forward. □

References

[1] **C Balteanu**, **Z Fiedorowicz**, **R Schwänzl**, **R M Vogt**, *Iterated monoidal categories*, Adv. Math. 176 (2003) 277–349 MR1982884

[2] **J M Boardman**, **R M Vogt**, *Homotopy invariant algebraic structures on topological spaces*, Lecture Notes in Math. 347, Springer, Berlin (1973) MR0420609

[3] **Z Fiedorowicz**, **S Gubkin**, **R M Vogt**, *Associahedra and weak monoidal structures on categories* `arXiv:1005.3979v4`

[4] **V Ginzburg**, **M Kapranov**, *Koszul duality for operads*, Duke Math. J. 76 (1994) 203–272 MR1301191

[5] **A Joyal**, **R Street**, *Braided tensor categories*, Adv. Math. 102 (1993) 20–78 MR1250465

[6] **B Keller**, *Introduction to A–infinity algebras and modules*, Homology Homotopy Appl. 3 (2001) 1–35 MR1854636

[7] **M L LaPlaza**, *Coherence for associativity not an isomorphism*, J. Pure Appl. Algebra 2 (1972) 107–120 MR0297833

[8] **T Leinster**, *Higher operads, higher categories*, London Math. Soc. Lecture Note Series 298, Cambridge Univ. Press (2004) MR2094071

[9] **S Mac Lane**, *Natural associativity and commutativity*, Rice Univ. Studies 49 (1963) 28–46 MR0170925

[10] **J P May**, *Definitions: operads, algebras and modules*, from: "Operads: Proceedings of Renaissance Conferences (Hartford, CT/Luminy, 1995)", (J-L Loday, J D Stasheff, A A Voronov, editors), Contemp. Math. 202, Amer. Math. Soc. (1997) 1–7 MR1436912

[11] **D Quillen**, *Higher algebraic K–theory. I*, from: "Algebraic K–theory, I: Higher K-theories (Proc. Conf., Battelle Memorial Inst., Seattle, WA, 1972)", (H Bass, editor), Lecture Notes in Math. 341, Springer, Berlin (1973) 85–147 MR0338129

[12] **J D Stasheff**, *Homotopy associativity of H–spaces. I, II*, Trans. Amer. Math. Soc. 108 (1963) 275–292, 293–312 MR0158400

[13] **J D Stasheff**, *The pre-history of operads*, from: "Operads: Proceedings of Renaissance Conferences (Hartford, CT/Luminy, 1995)", (J-L Loday, J D Stasheff, A A Voronov, editors), Contemp. Math. 202, Amer. Math. Soc. (1997) 9–14 MR1436913

[14] **J D Stasheff**, *How I "met" Dov Tamari*, from: "Associahedra, Tamari Lattices and Related Structures Associahedra, Tamari Lattices and Related Structures", (F Müller-Hoissen, J M Pallo, J Stasheff, editors), Progress in Math. 299, Springer, Basel (2012)

[15] **D Tamari**, *Monoïdes préordonnés et chaînes de Malcev*, PhD thesis, Université de Paris (1951) MR0051833

[16] **D Tamari**, *The algebra of bracketings and their enumeration*, Nieuw Arch. Wisk. 10 (1962) 131–146 MR0146227

[17] **Wikipedia**, *Tamari lattice*, Wikipedia, The Free Encyclopedia (2012)

ZF, SG: *Department of Mathematics, Ohio State University*
Columbus OH 43210-1174, USA

RMV: *Fachbereich Mathematik/Informatik, Universität Osnabrück*
Albrechtstr. 28a, D-49069 Osnabrück, Germany

fiedorow@math.ohio-state.edu, gubkin@math.ohio-state.edu,
rainer@mathematik.uni-osnabrueck.de

http://www.math.osu.edu/~fiedorowicz.1/,
http://www.math.osu.edu/people/gubkin/view

Received: 3 June 2010 Revised: 7 November 2011