

# Why Three and Why Twenty: The Capacity of a Chrysene Cluster, the Base-Four Intermediary, and the Production Operator $\Delta$

Charles D. Schaper, Ph.D.

*Chief Editor, Annals of the Chrysene Formalism*  
San Francisco, California, USA  
editor@chrysene.com

## Abstract

The predecessor classifies the marker assignments of a six-unit chrysene cluster into twenty frame classes and records, in a table caption, that the members of a class are one physical stack read under different conventions. We prove that assertion: the map from clusters to classes is a bijection, so a class is a cluster and not an equivalence class of descriptions. Frame invariance of any product is then a consequence rather than a hypothesis, and the assumptions move to where they can be attacked — that the product depends on the cluster's marker assignment alone, and that the pathway does not involve the initiator.

Under those two, a product takes at most  $2^{2N-2} + 2^{N-1}$  values, which is twenty at three vessels. That is a ceiling; exactly twenty requires in addition that every class produce an amino acid and that distinct classes produce distinct ones, and the constant map shows how much the saturation clause carries. The cluster is fixed at six units on chemical grounds independent of codons and of the alphabet size, so the triplet codon is derived from the carrier alongside the ceiling: two standing questions from one input, neither number justifying the other. The conventional derivation of the triplet, that sixteen codons will not carry twenty amino acids, is unavailable here, the class count satisfying  $|X|/4^N = \frac{1}{4} + 2^{-(N+1)} < 1$  at every length so that the alphabet never outruns the codon supply and the sufficiency condition distinguishes no length. The property that explains twenty also forbids something. Every class has size two or four, so every fibre of an assembly-level map is even, while the standard genetic code has two fibres of size one; no such operator produces it. The same argument places termination at the reading level, two codons of one class carrying a stop and an amino acid, which is where release factors are observed to act. The production operator is shown to be a function of five binary agreements read modulo end-to-end reversal and of nothing else: absolute orientation is excluded, a vessel-localised pathway must transform the central vessel and requires odd cluster length, and a mechanism may read the amino or keto character of a flanking vessel but not the identity of a flanking base. Two discriminations are worked in those terms.

The standard genetic code is used in no step of the capacity argument, and by a further proposition the count is independent of the nucleobase alphabet entirely. What is not claimed is stated before the constructions: the theorem bounds the products of a chrysene cluster and asserts nothing about the biological alphabet; and the count is twenty only if both frame data are re-descriptions of one cluster rather than data of it, failing which it is thirty-two.

**Keywords:** Chrysene Formalism; registry classes; frame group; orbit count; information capacity; alphabet size of the genetic code; triplet codon; production operator; parity obstruction; non-implication.

## 1 Introduction

### 1.1 The problem

The genetic code has twenty amino acids. Accounts of that number argue, almost without exception, from properties of the amino acids themselves: which were available under prebiotic conditions, which were added late, which are stereochemically

matched to their codons, which a code would need in order to fold a protein. Each explains why a particular member belongs. None explains the cardinal.

The corpus contains a classification that produces the cardinal and does not use the code. [Sch26e] counts the orbits of the frame group on the marker assignments of a chrysene stack of  $2N$  sites and finds  $2^{2N-2} + 2^{N-1}$  of them, which is twenty at  $2N = 6$ . That paper then refutes the coincidence: its Proposition 3.15 records

that the agreement of twenty classes with twenty amino acids implies no map, and that its surjection of sixty-four words onto twenty classes is not the genetic code, the fibre profiles differing. The companion paper [Sch26f] takes the other direction, building an adjacency from the code and selecting a coarsening of it, and is conditional on the code at every step.

Neither asks what the classification bounds *by itself*. The question is available and it is not the question either paper answers: not whether the twenty classes can be identified with the twenty amino acids, but how many distinguishable products a carrier of that shape can have at all. That question needs no codon table, and this paper shows it has an answer.

## 1.2 The objective

To establish that the frame classes are the physical clusters rather than descriptions of them, and hence that invariance of a product under the frame is a consequence and not a hypothesis; to bound the number of distinct products of a cluster, and to separate the bound from the assertion that it is attained; to fix the cluster length on chemical grounds independent of the code and derive the codon length from it; to determine what a production pathway may depend on, and what it may not; to state what the capacity forbids as sharply as what it permits; and, throughout, to keep the standard genetic code out of every hypothesis and every construction.

## 1.3 What is established

1. A class is a cluster: the map from clusters to frame classes is a bijection, so the two or four marker assignments of a class are one physical object described under the available frames (Proposition 4.4). In the predecessor this appears only as a table caption.
2. Consequently frame invariance of a product is a corollary of locality and initiator independence, not an assumption (Corollary 4.6), and the hypothesis register is restructured accordingly (Remark 4.7).
3. The capacity theorem: at most  $2^{2N-2} + 2^{N-1}$  distinct products, twenty at  $N = 3$ ; exactly twenty under totality and injectivity (Theorem 5.4). The two clauses are stated separately so that the first may be granted without the second.
4. The codon is a triplet, derived from the cluster and not assumed (Corollary 5.6); and the classical argument for the triplet cannot be run here,  $|X| < 4^N$  at every length (Proposition 5.2, Remark 5.3).
5. Parity: every class has even size, the standard code has two singleton fibres, and therefore no assembly-level operator produces it (Proposition 5.9, Corollary 5.10). Termination is placed at the reading level by the same argument (Corollary 5.12).

6. The production operator is a function of the registry class — five binary agreements read modulo reversal — and of nothing else (Proposition 6.1); absolute orientation is excluded (Corollary 6.2); a vessel-localised pathway must transform the central vessel and requires odd  $N$  (Proposition 7.2); and the amino or keto character of a flanking vessel is admissible where the identity of a base is not (Corollary 7.9).
7. The count is independent of the nucleobase alphabet (Proposition 4.12), and seven non-implications are recorded (Proposition 3.1).

## 1.4 What is not claimed

No axiom is introduced (Remark 2.2); A3, A4, A5, A7, A9, A10 and A11 are used nowhere (Proposition 2.5), so nothing here concerns sequence dependence; and no deferred target of any predecessor is discharged (Remark 8.9).

Three readings the construction invites are refuted rather than used, and they are stated at the outset because the body would otherwise be taken to support them.

The first concerns the word *count*. The theorem gives a ceiling. The constant map carrying every class to one amino acid satisfies both standing hypotheses and has an image of one element, so the passage from “at most twenty” to “exactly twenty” is carried entirely by totality and injectivity, which this paper does not establish.

The second concerns the number itself. Twenty is the orbit count of the full frame group; it is thirty-two if the projection reference alone acts and sixty-four if neither does. Which applies is not a mathematical question: the initiator supplies both frame data, so for the *codon* they are data of the assembly, and only for the *product* — under the hypothesis that the pathway does not involve the initiator — are they re-descriptions. A reader who declines that hypothesis should read this paper as a theorem about thirty-two (Remark 3.3).

The third concerns the title. The theorem bounds the products of a chrysene cluster. That the biological alphabet is that set is an identification of the kind [Sch26b, §9.2] records as the corpus’s largest unproved premise, and the paper declines it (Proposition 3.1(vii)). The question is answered for a carrier, not for biology.

## 1.5 Logical position

Every numbered statement carries one of the four status labels of Definition 2.1. Every result using a hypothesis beyond A1–A12 names it, and every hypothesis is registered in Section 2.5. Every molecular reading is confined to remarks marked [E] and is used in no proof. This paper falls within the framework scope condition of [Sch26b, §4.5] and is written to the orthogonal-constraint review bar of [Sch26b, §4.3].

Three features of the record are stated here rather than left to the appendix.

First, and unlike the companion paper, *the standard genetic code is used in no hypothesis and no construction*. It appears three times, each time to show what the account does not reach, and by Proposition 4.12 the count is independent of the nucleobase alphabet as well. Remark 2.4 states this and Remark 2.21 lists the hypotheses of the companion papers that are deliberately absent, so that the claim can be checked rather than taken.

Second, the chemistry is confined to two registered hypotheses and a scope statement.  $H_{\text{clu}}$  and  $H_{\text{int}}$  are realization hypotheses and not [E] claims, because Corollary 5.6 is a proof and an empirical instantiation claim enters no proof. Section 7 identifies its reactants and outlines its process; the synthesis steps are forthcoming as a separate chemical treatise and no statement of Sections 4 to 6 is conditional on them (Remark 7.17).

Third, the defect ledger of Appendix B is populated from the pre-drafting pass onward, in the crit/rev/ver form of [Sch26b, Def. 3]. Sixteen items are recorded, of which fourteen were raised against the concept outline before any section was written. Three of the sixteen made the manuscript claim *more* rather than less, each by making an assumption explicit, which is a direction the ledgers of the two preceding manuscripts did not take (Remark B.3). Targeted verification re-runs the mathematical stage and does not employ an independent reviewing system; by [Sch26b, §9.4] the reader should discount it accordingly, the seven numerical checks excepted.

## 1.6 Organization

Section 2 fixes conventions, provenance, the carrier and the hypothesis register. Section 3 states the traps and non-implications, before the constructions they guard. Section 4 proves that a class is a cluster and records what the base-four intermediary costs. Section 5 proves the capacity theorem, the derivation of the triplet, and what capacity forbids. Section 6 determines what the production operator may depend on and states its obligations. Section 7 works two discriminations in those terms. Section 8 restates the non-implications and inventories the Targets. Appendix A tabulates the twenty classes; Appendix B is the concordance, the chemical scope, the defect ledger and the verification record.

## 2 Conventions, Logical Status, and the Hypothesis Register

### 2.1 Logical status

**Definition 2.1** (Status labels). Imported verbatim from [Sch26c, Defs. 1, 3, 5, 6] and used as in [Sch26a, Def. 2.7]: [A] axiomatic; [D] derived realization theorem; [T] target; [E] empirical instantiation claim.

**Remark 2.2** (No thirteenth axiom). [A] This paper introduces no axiom. Every structure below is an object imported from [Sch26d] or [Sch26e], a definition internal to  $\{A1, \dots, A12\}$ , or a realization hypothesis in the sense of [Sch26c, Def. 3] registered in Section 2.5. The restriction is that of [Sch26d, Rem. 2.2]: an additional axiom would reopen [Sch26c, Thm. 61] and [Sch26c, Prop. 39], neither of which is re-established here.

**Remark 2.3** (Two counting conventions). Statements about a purely finite object — a set of binary strings, a group of permutations of a four-element set — are proved without reference to any axiom and are labelled [A]. Each is instantiated once, by a corollary labelled [D], as a statement about clusters under the hypotheses named there. The two are kept apart throughout, in the manner of [Sch26d, Rem. 3.9] and [Sch26e, Rem. 2.3].

**Remark 2.4** (This paper uses no external datum). One convention differs from the companion paper and is stated here rather than left to inference. [Sch26f, Rem. 2.4] records that the standard genetic code enters that paper constructively: it defines the adjacency, and every statement of its Sections 5 to 7 depends on it.

*No statement of this paper uses the standard genetic code.* It is mentioned in three places — Corollary 5.10 and Corollary 5.12, where a property of the published table is used to show what the account does *not* reach, and Remark 5.3, where a conventional argument is set aside — and in none of them does it enter a hypothesis or a construction. The class count, the capacity theorem and the derivation of the triplet are independent of it, and by Proposition 4.12 they are independent of the nucleobase alphabet as well. This is the first result of the corpus of which that is true and it is the reason the paper is written.

### 2.2 Provenance

**Proposition 2.5** (Provenance). [A] Each constituent of the constructions of this paper is supplied by a named axiom: the site set and the translation action, with the stacking direction, by A6; the two attachment sets, hence the species alternation, by A1; the non-collapse of the marker by A2; the oriented external perimeter cycle read by the marker by A8; and the constraint functionals and the admissible domain by A12. No constituent is supplied by A3, A4, A5, A7, A9, A10 or A11.

*Proof.* By inspection of the cited definitions against the statements of the axioms, as in [Sch26e, Prop. 2.7], from which

the present list does not differ.  $\square$

**Remark 2.6** (Consequence).  $A7$  being unused, no statement below concerns the order dependence of an encoded word, and the sequence-dependence theorem [Sch26d, Thm. 4.17] is not invoked. The only ordering in this paper is the ordering of sites within a cluster, which is a datum of the reading and is quotiented away by Proposition 4.4.

from it, by  $H_{\text{int}}$  and Corollary 5.6; the ceiling of twenty follows from it, by Theorem 5.4. Neither number is used to justify the cluster, and neither is used to justify the other. The conventional derivation runs the other way — twenty amino acids require more than sixteen codons, so codons are triplets — and is in any case unavailable here, by Remark 5.3. This remark enters no proof; what the proofs use is Hypotheses 2.7 and 2.8.

## 2.3 Imported objects

| Object                                                                      | Source                              |
|-----------------------------------------------------------------------------|-------------------------------------|
| marker $\text{mk}$ , values $F = 0, B = 1$                                  | [Sch26d, Def. 3.3]                  |
| non-collapse of the marker                                                  | [Sch26d, Prop. 3.4]                 |
| placement, and the group $R$ of available operations                        | [Sch26d, Def. 3.2]                  |
| vessel; the binaries $b_1, b_2$ ; alphabet $\Sigma_p$                       | [Sch26d, Defs. 3.10, 3.12, 3.13]    |
| frame $(k_0, \varepsilon, f)$ ; frame-setters $\Phi$ ; the torsor structure | [Sch26d, Defs. 6.1, 6.2, Prop. 6.5] |
| decomposition of the frame action                                           | [Sch26d, Prop. 6.7]                 |
| initiator                                                                   | [Sch26d, Def. 9.1]                  |
| frame group $\text{Fr} = \langle \text{flip}, \text{rev} \rangle$           | [Sch26e, Def. 2.13]                 |
| the class set $X$ and its count                                             | [Sch26e, Lem. 3.1, Thm. 3.3]        |
| class sizes two and four, profile eight and twelve                          | [Sch26e, Prop. 3.6, Cor. 3.7]       |
| realization of every class under the constraint set of $A12$                | [Sch26e, Prop. 3.11]                |

## 2.4 The carrier

The two hypotheses of this subsection fix the physical object and are the only place where the chemistry enters a proof. Both are stated as realization hypotheses and not as [E] remarks, because Corollary 5.6 depends on them and an [E] claim may enter no proof.

**Hypothesis 2.7** ( $H_{\text{clu}}$ , the cluster length). The reacting unit of the chrysene architecture is a stack of six units alternating in species, hence a cluster in the sense of Definition 4.1, with  $N = 3$  vessels.

**Hypothesis 2.8** ( $H_{\text{int}}$ , one base per interface). Each of the three catalyst–substrate interfaces of a cluster produces exactly one nucleobase.

**Remark 2.9** (Provenance of the two, and why the order matters). [E] In the intended reading, the six-unit cluster is the reacting unit of the architecture claimed in US 12,195,423 and US 12,473,247, fixed by the geometry of the  $\pi$ -stack and the reach of the phosphoethoxy substituent, and each catalyst–substrate interface synthesises one nucleobase. The order in which these are established is the whole of the paper’s claim to non-circularity, and it is therefore stated explicitly.  $H_{\text{clu}}$  is a statement about the chrysene architecture. It makes no reference to codons, to the length of a codon, or to the number of amino acids, and it would be the same statement if no genetic code existed. The triplet follows

## 2.5 The production operator, and its register

**Definition 2.10** (The production operator).  $\Delta$  denotes the map assigning to a cluster the chemical product of the amino-acid pathway, that pathway being distinct from the nucleobase pathway of  $H_{\text{int}}$  and proceeding within the six units.

**Hypothesis 2.11** ( $H_{\text{prod}}(0)$ , locality). The product  $\Delta(\mathfrak{R})$  is a function of the marker assignment of  $\mathfrak{R}$  alone: not of temperature, of concentration, of neighbouring clusters, or of reaction history.

**Hypothesis 2.12** ( $H_{\text{prod}}(0')$ , initiator independence). The amino-acid pathway does not involve the initiator of [Sch26d, Def. 9.1]; consequently  $\Delta$  is independent of the direction sign  $\varepsilon$  and of the projection reference  $f$ .

**Hypothesis 2.13** ( $H_{\text{prod}}(1)$ , totality). Every cluster produces an amino acid.

**Hypothesis 2.14** ( $H_{\text{prod}}(2)$ , injectivity). Clusters of distinct classes produce distinct amino acids.

**Remark 2.15** (Which clause does what). The ceiling of Theorem 5.4(i) uses  $H_{\text{prod}}(0)$  and  $(0')$  and nothing else. The count of clause (ii) uses all four.  $H_{\text{clu}}$  is what instantiates either at  $N = 3$ .  $H_{\text{int}}$  is used only in Corollary 5.6 and nowhere in Section 5.

The four are not of equal standing.  $(0)$  is the load-bearing one: it is what makes  $\Delta$  a function of the carrier at all, and a pathway sensitive to conditions outside the cluster would defeat the theorem without being in any way chemically implausible.  $(0')$  decides between twenty and thirty-two, by Remark 5.15.  $(1)$  and  $(2)$  together are the assertion that the system uses its capacity, which is substantive and which this paper does not establish.

**Remark 2.16** (Frame-invariance is not among them). Drafts before Pass 0 registered frame-invariance of  $\Delta$  as the load-bearing hypothesis. It is not a hypothesis at all: by Proposition 4.4 a class is a cluster, so a function of the cluster is automatically constant on the class, and Corollary 4.6 derives the invariance from  $H_{\text{prod}}(0)$  and  $(0')$ . The substitution is recorded in the defect ledger; the content of the withdrawn hypothesis is now carried by  $(0)$  and  $(0')$ , where a reader can attack it.

## 2.6 Imported hypotheses

**Hypothesis 2.17** ( $H_{\text{mk}}$ , marker realization). As [Sch26d, Hyp. 2.6].

**Hypothesis 2.18** ( $H_{\text{stack}}$ ,  $z$ -constraint). As [Sch26d, Hyp. 2.7].

**Hypothesis 2.19** ( $H_{\text{frm}}$ , external frame datum). As [Sch26d, Hyp. 2.8]. Used in Proposition 4.4 to make  $\text{Fr}$  the group of frame changes at the level of markers, by [Sch26d, Prop. 6.7].

**Hypothesis 2.20** ( $H_{\text{ind}}$ , vessel independence). As [Sch26d, Hyp. 2.11]. Used only for the surjectivity clause of Proposition 4.4(i), through [Sch26d, Prop. 3.19].

**Remark 2.21** (Hypotheses deliberately not used). The following are registered in the companion papers and are used nowhere here, which is worth recording because their absence is what makes Remark 2.4 true.  $H_{\text{bij}}$ ,  $H_{\text{sec}}$  and  $H_{\text{base}}$  concern the transformation into the nucleobase alphabet, which by Proposition 4.12 no statement of Section 5 requires.  $H_{\text{crys}}$ ,  $H_{\text{fit}}$  and  $H_{\text{var}}$  concern the filter scores and the selection of a coarsening, and no selection is made here.  $H_{\text{stop}}$  concerns the designation of three codons as non-coding; by Corollary 5.12 termination is a reading-level datum and this paper neither needs nor uses that designation.

## 2.7 Naming, and the concordance substitutions

**Remark 2.22** (Two substitutions and one reservation). Three names require care and all three are recorded in the concordance as *restated*, in the manner of [Sch26e, Rem. 2.12].

$H_{\text{prod}}(0')$  and  $H_{\text{prod}}(2)$  are, respectively, a strengthening and a restatement of clauses (i) and (ii) of  $H_{\text{dual}}$  of [Sch26f, Hyp. 2.30]. The renaming is made because this paper asserts no double-sided correspondence: it concerns one pathway

and not the relation between two. No clause of  $H_{\text{dual}}$  is weakened or strengthened by the renaming, except that what that paper obtained as clause (i) is here a corollary and the assumption has moved to  $H_{\text{prod}}(0)$  and  $(0')$ .

$\Delta$  is the production operator of Definition 2.10 and must not be read as the coarsening  $\varphi$  of [Sch26f, Thm. 5.10], which is a bijection from classes to amino acids selected among one hundred and twenty-eight by five filters. The two have the same source and target and are otherwise unrelated:  $\varphi$  is chosen by a criterion,  $\Delta$  is produced by a pathway, and whether they coincide is [Sch26f, Target 8.16] and is not addressed here.

$\Delta$  is the production operator of Definition 2.10 and is not the localized cipher involution  $\Delta_{\mathbf{k}}$  of A8, which [Sch26d, §2.2] carries under that letter. The two are kept apart by the same discipline [Sch26d, Rem. 2.20] applies to the three senses of registry: the operator of this paper is never subscripted and is never localized at a site, while  $\Delta_{\mathbf{k}}$  is always both. No statement of this paper uses A8's involution — A8 enters only through the oriented perimeter cycle read by the marker (Proposition 2.5) and, in one remark used in no proof, through the cipher pairing  $\delta_{\text{dc}}$  (Remark 7.7) — so the two never occur in one argument. The substitution is recorded in the concordance.

Finally, *cluster* is reserved for the six-site object, as in [Sch26f, Def. 2.14], and *vessel* for the adjacent pair of [Sch26d, Def. 3.10]. No fourth sense of *registry* is introduced; the three of [Sch26d, Rem. 2.20] stand.

## 2.8 Notation concordance

| Symbol                               | Meaning                                                            |
|--------------------------------------|--------------------------------------------------------------------|
| $\mathfrak{R}$                       | a cluster, Definition 4.1                                          |
| $\text{desc}(\mathfrak{R}, \varphi)$ | its marker assignment read under $\varphi$                         |
| $X$                                  | the class set; by Prop. 4.4 the set of clusters                    |
| $\Delta$                             | the production operator, Definition 2.10                           |
| $W$                                  | the marker-to-letter bijection, Prop. 4.8                          |
| $L$                                  | a letter assignment $\Sigma_{\text{p}} \rightarrow \{A, C, G, U\}$ |
| $\varphi$                            | reserved: the coarsening of [Sch26f], not used here                |

## 3 Traps and Non-Implications

The formalism's governing restraint is that restoring a foundational object never licenses restoring an unproved implication of that object [Sch26b, §3.1]. This section states what the constructions of Sections 4 to 7 do not establish, and states it before them, for the reason given at Remark 3.5. Each clause is refuted by a witness named in its proof; the section adds no argument and exists so that the list can be checked against the completed text.

**Proposition 3.1** (Non-implications). **[D]** Assume  $H_{mk}$ ,  $H_{stack}$ ,  $H_{frm}$  and  $H_{ind}$ . None of the following implications holds; each is refuted by the witness named in its proof.

- (i) Twenty classes  $\not\Rightarrow$  twenty amino acids. The capacity theorem is a ceiling; the count requires saturation.
- (ii) The capacity theorem  $\not\Rightarrow$  any assignment of amino acids to codons. No operator satisfying  $H_{prod}(0)$  and  $(0')$  produces the standard genetic code, or any table with a fibre of odd cardinality.
- (iii) The partition result of [Sch26f, Thm. 5.2] — that a perfect matching to the twenty amino acids exists under the amino/keto partition and fails by one and by two under the other two —  $\not\Rightarrow$  anything about capacity, and conversely. The two are different functions of different data.
- (iv) Twenty at  $N = 3$   $\not\Rightarrow$  twenty at any other cluster length. The counts are 2, 6, 20, 72, 272 at  $N = 1, \dots, 5$ .
- (v) The capacity theorem  $\not\Rightarrow$  anything about the nucleobase alphabet, and no statement about the nucleobases  $\not\Rightarrow$  the capacity. The count is proved over  $\Sigma_0$  and does not use the letter assignment  $L$ .
- (vi) Sufficiency of the codon supply  $\not\Rightarrow$  any constraint on the cluster length.  $|X| < 4^N$  at every  $N$ , so the alphabet never outruns the codons and the classical argument for the triplet cannot be run here.
- (vii) A bound on the products of a chrysene cluster  $\not\Rightarrow$  that the biological amino-acid alphabet is that set. No such identification is asserted here.

*Proof.* (i) By witness. The constant map carrying every class to alanine satisfies  $H_{prod}(0)$  and  $(0')$  — it is a function of the marker assignment and independent of the initiator — and its image has one element. The ceiling of Theorem 5.4(i) is therefore attained nowhere in general, and the passage from “at most twenty” to “exactly twenty” is carried entirely by  $H_{prod}(1)$  and  $(2)$ . Clause (ii) of that theorem names them for this reason.

(ii) Corollary 5.10, whose witness is Proposition 5.9 together with the two singleton fibres of the standard code, methionine at AUG and tryptophan at UGG. The obstruction is parity and needs no adjacency, no matching and no coarsening; it is strictly stronger than the fibre-profile argument of [Sch26e, Prop. 3.15(ii)].

(iii) By exhibition of the two quantities. The class count is invariant under all twenty-four letter assignments, by Proposition 4.12, and is computed without reference to any codon table. The maximum matching of [Sch26f, Thm. 5.2] takes the three values 20, 19, 18 according to the partition and is

computed from the standard genetic code. A quantity constant in  $L$  and a quantity varying in  $L$  cannot determine one another, and the second in addition requires a datum the first does not use. In particular the partition result is *a posteriori*: it could not have been stated before the codon table was consulted, and the capacity could have been stated without it.

(iv) The two sequences are those of Theorem 5.1, computed at check V3.

(v) The forward implication is Proposition 4.12. For the converse,  $|X|$  is determined by the site number and the frame group and is unchanged if  $\Sigma'_0$  is replaced by any other four-element set, or if no secondary alphabet is given at all; so no statement about the nucleobases constrains it.

(vi) Proposition 5.2:  $|X|/4^N = \frac{1}{4} + 2^{-(N+1)} < 1$  for every  $N \geq 1$ . The sufficiency condition of the classical argument is therefore satisfied at  $N = 1$  as at  $N = 3$  and distinguishes no length.

(vii) Is a statement about what this paper asserts, and is verified by inspection of its statements. Theorem 5.4 quantifies over clusters and its conclusion is about the products of clusters. An assertion that the biological alphabet is the image of  $\Delta$  would be an empirical instantiation claim in the sense of [Sch26c, Def. 6], of the kind [Sch26b, §9.2] records as the corpus’s largest unproved premise, and none is made or used.  $\square$

**Remark 3.2** (Status of the individual clauses). Clauses (i), (iv), (v) and (vi) are refuted by witnesses using no hypothesis and could carry the label **[A]** separately. The proposition is labelled **[D]** because (ii) and (iii) refer to the standard genetic code as a published table and (vii) refers to the hypotheses of Section 2.5.

**Remark 3.3** (Trap: the count belongs to the full frame group). **[A]** The orbit count is twenty only for  $Fr = \langle \text{flip, rev} \rangle$ ; it is thirty-two for the projection reference alone and sixty-four for no frame change at all (Proposition 5.14). So

a six-unit cluster  $\not\Rightarrow$  twenty classes,

absent the statement that both frame data are re-descriptions of one cluster rather than data of it. For the *codon* they are not, the initiator supplying both; for the *product* they are, by  $H_{prod}(0')$ . The difference between twenty and thirty-two is therefore the difference between the two branches of the cluster and nothing else, and a reader who declines  $H_{prod}(0')$  should read this paper as a theorem about thirty-two. Remark 5.15 states this where the count is proved, and Remark 4.2 where the carrier is defined.

**Remark 3.4** (Trap:  $\Delta$  is not the coarsening of the companion paper). **[A]**  $\Delta$  and the coarsening  $\varphi$  of [Sch26f, Thm. 5.10]

have the same source and the same target: both assign an amino acid to each of the twenty classes. They are otherwise unrelated.  $\varphi$  is *selected*, among one hundred and twenty-eight bijections compatible with the standard genetic code, by five physical filters and a margin of 0.17%;  $\Delta$  is *produced*, by a pathway, and this paper offers no criterion by which it is chosen. Accordingly

$$\Delta \text{ exists} \not\Rightarrow \Delta = \varphi,$$

and the agreement of the two is [Sch26f, Target 8.16], not a result here. A reader who assumes the identification imports the whole of that paper's dependence on the standard genetic code into this one, which Remark 2.4 exists to prevent.

**Remark 3.5** (Why the section is placed here). A reader who arrives at Section 5 holding the number twenty will read the capacity theorem as a derivation of the genetic code, and will do so whether or not one is offered; the reading is invited by the reading-frame analogy of [Sch26d, Rem. 6.12], by the biogenesis thesis of [Sch26b, §7.2], and by the construction of the companion paper, none of which this paper uses. The refutations are therefore placed before the constructions, as [Sch26e, Rem. 3.17] places its two, and are stated again in Section 8.

**Remark 3.6** (The three that matter most). Clauses (i), (ii) and (vii) are the ones on which a reader is most likely to over-read.

Clause (i) states where the paper's substance ends. The mathematics gives a ceiling of twenty; that the ceiling is attained is the assertion that the system uses its capacity, and it is a hypothesis. The abstract says ceiling before it says count, and the theorem is written in two clauses so that the first may be granted without the second.

Clause (ii) states what the capacity costs. The property that explains twenty — that the product is attributable to the assembly, hence constant on classes — is the same property that forbids an assembly-level operator from producing the standard code. The two cannot be separated, and a reader who wants an operator reaching all sixty-four assignments is asking for the capacity theorem to be false.

Clause (vii) states the gap between the theorem and the question in the title. The theorem bounds what a chrysene cluster can produce. That the biological twenty is that set is an identification the paper declines to make, and the title asks a question the paper answers for chrysene clusters and not for biology.

A reader who grants all three retains what the paper proves: that a six-unit carrier admits exactly twenty distinguishable states, that its output is therefore a base triplet and its product

alphabet is bounded by twenty, and that neither conclusion uses the genetic code.

**Remark 3.7** (Numerical verification). Every count and every witness of Proposition 3.1 was recomputed from the definitions by the supplementary script, which imports nothing from the tables of this paper. The relevant checks are V1 (orbit counts under every subgroup of Fr), V2 (the class count under all twenty-four letter assignments), V3 (the counts at  $N = 1, \dots, 8$  and the ratio  $|X|/4^N$ ) and V4 (the parity statement, and the two singleton fibres of the standard code).

## 4 The Carrier and the Base-Four Intermediary

This section fixes the physical object the paper is about, proves that the frame classes of [Sch26e] *are* those objects rather than descriptions of them, and records what the passage to the base-four intermediary does and does not cost. Section 5 rests on Proposition 4.4 throughout.

### 4.1 The cluster

**Definition 4.1** (Cluster). A *cluster* is a stack of six sites  $n_1, \dots, n_6$  in the sense of [Sch26d, Def. 3.1], alternating in species, read at the offset  $o = 0$  of [Sch26e, Def. 2.9], together with the placement datum of [Sch26d, Def. 3.2] at each site. Its *marker assignment* is  $m = (\text{mk}(n_1), \dots, \text{mk}(n_6)) \in \{0, 1\}^6$ .

The initiator of [Sch26d, Def. 9.1] is *not* a site of the cluster. Two clusters are identified when their placements are related site by site by an element of the group  $R$  of [Sch26d, Def. 3.2], that is, when no operation available to a unit distinguishes them.

**Remark 4.2** (Why the initiator is excluded, and what it costs). The exclusion is a definition and not an oversight, and everything in Section 5 turns on it. By [Sch26d, Def. 6.1] the initiator supplies all three frame data — the origin, the direction sign  $\varepsilon$  and the projection reference  $f$  — and by [Sch26d, Def. 9.1] it occupies a distinguished end of the stack. An assembly taken *with* its initiator therefore has a distinguished end and a determined reading direction, and reversal is not a re-description of it but a different situation; the group acting would be  $\langle \text{flip} \rangle$  alone and the count of Proposition 4.4 would be thirty-two rather than twenty (Proposition 5.14).

What licenses the exclusion is  $H_{\text{prod}}(0')$ : the synthesis

pathway of Section 6 proceeds within the six units, from the phosphoethoxy arm at C12 of a substrate unit, and does not involve the initiator. The product is therefore blind to a datum the reading of a codon requires. That asymmetry is the central empirical commitment of this paper, and it is stated as a hypothesis in Section 2 rather than assumed here.

## 4.2 A class is a cluster

The following is the statement on which the capacity theorem rests. In [Sch26e] the corresponding assertion appears in the caption of its Table 1 and nowhere as a numbered statement; a caption cannot carry a hypothesis register, and the statement is therefore proved here.

**Definition 4.3** (Description). For a cluster  $\mathfrak{R}$  and a frame-setter  $\varphi \in \Phi$  of [Sch26d, Def. 6.2],  $\text{desc}(\mathfrak{R}, \varphi) \in \{0, 1\}^6$  denotes the marker assignment of  $\mathfrak{R}$  read under  $\varphi$ .

**Proposition 4.4** (A class is a cluster). *[D] Assume  $H_{\text{mk}}$ ,  $H_{\text{stack}}$ ,  $H_{\text{frm}}$  and  $H_{\text{ind}}$ . Then*

- (i) *for each fixed  $\varphi$ , the map  $\mathfrak{R} \mapsto \text{desc}(\mathfrak{R}, \varphi)$  is a bijection from clusters onto  $\{0, 1\}^6$ ;*
- (ii) *for each fixed  $\mathfrak{R}$ , the set  $\{\text{desc}(\mathfrak{R}, \varphi) : \varphi \in \Phi\}$  is exactly one Fr-orbit;*
- (iii) *consequently  $\mathfrak{R} \mapsto \text{Fr} \cdot \text{desc}(\mathfrak{R}, \varphi)$  is a bijection from clusters onto the class set  $X$ , independent of  $\varphi$ .*

*A class is therefore a cluster, and the four (or two) marker assignments it contains are one physical object described under the available frames.*

*Proof.* (i) Injectivity: by  $H_{\text{mk}}$  the marker of a site is determined by the placement there, and by [Sch26d, Prop. 3.4] its two values are identified by no element of  $R$  and by no element of the labelled symmetry stabiliser of  $A_2$ ; so two clusters with the same marker assignment under one frame are related site by site by an element of  $R$  and are identified by Definition 4.1. Surjectivity is  $H_{\text{ind}}$  with [Sch26d, Prop. 3.19]: no constraint functional couples distinct vessels, so every element of  $\Sigma_v^3$ , hence by [Sch26e, Lem. 2.15] every element of  $\{0, 1\}^6$ , is realized.

(ii) By [Sch26d, Prop. 6.5] the frame-setters form a torsor under  $\mathbb{Z}/2 \times \mathbb{Z}/2$ , generated by reversal of the side binary and of the registry binary. By [Sch26d, Prop. 6.7], at the level of markers the second acts as flip, the change  $f \mapsto f + 1$  incrementing every marker, and the first acts as rev, the reversal of the site index; the third frame datum, the offset, induces no action on markers at all, a marker being a function of a

single site and of no decomposition. Hence the descriptions of a fixed cluster are the images of any one of them under  $\langle \text{flip}, \text{rev} \rangle = \text{Fr}$ , which is one orbit.

(iii) By (ii) the assignment is well defined and independent of  $\varphi$ , by (i) it is injective on clusters, and by (i) with (ii) every orbit is attained.  $\square$

**Remark 4.5** (What the proposition does and does not say). It says that a marker assignment is a *description* and a class is the thing described. It does not say that the four descriptions are indistinguishable to an observer holding a frame: by [Sch26d, Prop. A.7] the four frame-setters read four distinct words from one and the same stack, and that is precisely why the codon is not an assembly quantity. The two statements are consistent and are the two halves of [Sch26e, Thm. 3.3(iii)].

Nor does it say that a cluster is physically realized. That every class is occupied under the constraint set of [Sch26d, Def. 9.2] is [Sch26e, Prop. 3.11]; that the marker is information-bearing at all remains [Sch26d, Prop. 3.8] and is [T].

**Corollary 4.6** (Frame-invariance is not a hypothesis). *[D] Under  $H_{\text{prod}}(0)$  and  $H_{\text{prod}}(0')$ , the product of a cluster is a function of the cluster, hence by Proposition 4.4 a function on  $X$ . No further invariance assumption is required.*

*Proof.*  $H_{\text{prod}}(0)$  makes the product a function of the marker assignment;  $H_{\text{prod}}(0')$  makes it independent of  $\varepsilon$  and  $f$ , hence constant on the descriptions of one cluster, which by Proposition 4.4(ii) is a class.  $\square$

**Remark 4.7** (Where the content went). Drafts before Pass 0 registered frame-invariance of the product as the load-bearing hypothesis. Corollary 4.6 shows it is not one. What is assumed is  $H_{\text{prod}}(0)$ , that the product depends on the marker assignment *alone* and not on temperature, concentration, neighbouring clusters or reaction history; and  $H_{\text{prod}}(0')$ , that the pathway does not involve the initiator. Those two carry the whole weight, and the change is recorded in the defect ledger.

## 4.3 The base-four intermediary

**Proposition 4.8** (The recoding is a bijection). *[A] The assignment  $(\text{mk}(\mathbf{c}), \text{mk}(\mathbf{s})) \mapsto (b_1, b_2)$  of [Sch26d, Def. 3.12] is an affine bijection of  $(\mathbb{Z}/2)^2$ , with linear part  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  and translation part  $(0, f)$ . Applied vesselwise it carries the six binary markers of a cluster to three letters of  $\Sigma_v$ , and*

$$2^6 = 4^3 = 64.$$

*Sch26d, Prop. 3.14.* for the affine bijection; the vesselwise statement is [Sch26e, Lem. 2.15], the offset-0 vessels partitioning the six sites.  $\square$

**Lemma 4.9** (The recoding intertwines the frame actions). *[D] Assume  $H_{mk}$ ,  $H_{stack}$  and  $H_{frm}$ . The bijection  $W$  of Proposition 4.8 satisfies  $W \circ \text{flip} = \rho^{\times 3} \circ W$  and  $W \circ \text{rev} = (\text{reversal}) \circ \tau^{\times 3} \circ W$ , so it is Fr-equivariant and the two actions have the same orbits.*

*Sch26e, Lem. 2.15.*  $\square$

**Corollary 4.10** (Nothing is lost in the recoding). *[D] The class set computed over the six binary markers and the class set computed over the three base-four letters are canonically identified. In particular the count of Section 5 may be taken in either presentation, and the two-step passage — six binary faces, then three base-four letters — is information-preserving at each step. Whatever is lost is lost at the quotient by the frame and not in the recoding.*

**Remark 4.11** (The intermediary is not a nucleobase). *[E] A reader meeting a four-letter alphabet in this setting will take it for  $\{A, C, G, U\}$ . It is not. The intermediary is  $\Sigma_v$ , the vessel alphabet of [Sch26d, Def. 3.13], whose four letters are the orientational states of a catalyst–substrate pair. In the intended reading it is a construct of greater complexity than a nucleobase, and it is a common intermediary complex to both products of the cluster: the nucleobase on one branch and the L-amino acid on the other. It therefore sits upstream of the fork and not on the nucleic-acid side of it. This remark enters no proof.*

**Proposition 4.12** (The classification does not use the nucleobases). *[A] Let  $L : \Sigma_v \rightarrow \{A, C, G, U\}$  be any bijection. The class set  $X$  and its cardinality are independent of  $L$ , and no statement of Section 5 invokes  $L$ . In particular  $|X| = 20$  at  $N = 3$  under each of the twenty-four assignments.*

*Proof.*  $X$  is the orbit set of Fr on  $\{0, 1\}^6$  by Proposition 4.4, and  $L$  relabels the codon written on a class without altering the class. The count was recomputed under all twenty-four assignments; this is check V2 of Appendix B.  $\square$

**Remark 4.13** (Why this is a strengthening, and what it separates). Two consequences, and the second is the more useful.

The capacity theorem is independent of the nucleobase identification, hence of the chemistry of that branch, hence

not hostage to it. The result stands whether or not the letter assignment of [Sch26f, Def. 2.16] is the right one.

And it is exactly what separates this paper's result from the partition result of [Sch26f, Thm. 5.2], that a perfect matching to the twenty amino acids exists precisely under the amino/keto partition, with Hall deficiencies one and two for the other two. That result depends on  $L$  and on the standard genetic code, and it could not have been stated in advance of either. This paper's count depends on neither. The two should not be read as the same kind of statement, and Proposition 3.1(iii) records the non-implication.

**Remark 4.14** (Three four-element objects). The text keeps apart  $\Sigma_v$ , the intermediary;  $\Sigma'_0 = \{A, C, G, U\}$ , the nucleobases; and  $\Phi$ , the four frame-setters. No two are canonically isomorphic. That  $\Sigma_v$  and  $\Phi$  are both torsors under  $\mathbb{Z}/2 \times \mathbb{Z}/2$  carrying exactly one canonically pointed coordinate, and are nevertheless not canonically identified, is [Sch26d, Rem. 6.6]; the map  $z : \Phi \rightarrow \Sigma_v$  carrying a frame-setter to the zero it designates is two-to-one onto the aligned class, by [Sch26d, Cor. A.8].

## 5 The Capacity Theorem

### 5.1 The count

**Theorem 5.1** (The class count). *[A] For a stack of  $2N$  sites the frame group has*

$$|X| = 2^{2N-2} + 2^{N-1}$$

*orbits; at  $2N = 6$  the value is twenty. Eight classes have size two and twelve size four.*

*Sch26e, Lem. 3.1. and [Sch26e, Cor. 3.7], where three independent proofs are given. The statement is imported and not reproved; by Corollary 4.10 it may be read over the six binary markers or over the three base-four letters indifferently.  $\square$*

**Proposition 5.2** (The codon supply is never the binding constraint). *[A] For every  $N \geq 1$ ,*

$$\frac{|X|}{4^N} = \frac{1}{4} + 2^{-(N+1)} < 1,$$

*so  $|X| < 4^N$  at every cluster length.*

*Proof.*  $|X| = 2^{2N-2} + 2^{N-1}$  and  $4^N = 2^{2N}$ ; divide. Both terms are positive and the sum is at most  $1/4 + 1/4 = 1/2$  at  $N = 1$ , decreasing to  $1/4$ . Verified for  $N = 1, \dots, 8$  at check V3 of Appendix B.  $\square$

**Remark 5.3** (The classical argument for the triplet cannot be run here). [A] The conventional account of the codon length is that  $4^2 = 16 < 20 \leq 64 = 4^3$ , so a doublet code cannot carry twenty amino acids. Inside this formalism that inequality is automatic and therefore distinguishes nothing: the classes are a quotient of the words, so by Proposition 5.2 the alphabet cannot outrun the codon supply at *any* length, and the sufficiency condition holds at  $N = 1$  as at  $N = 3$ . The cluster length must therefore come from elsewhere, and it does:  $N = 3$  is fixed by the chemistry of Section 2, independently of codons and of the number of amino acids. The causation in this paper runs from the cluster to both numbers and not from either number to the cluster. The non-implication is Proposition 3.1(vi).

## 5.2 Capacity

**Theorem 5.4** (Capacity). [D] *Let a cluster of  $2N$  sites produce a chemical product, and assume  $H_{\text{prod}}(0)$  and  $H_{\text{prod}}(0')$ . Then*

- (i) (ceiling) *the product takes at most  $2^{2N-2} + 2^{N-1}$  distinct values; under  $H_{\text{clu}}$ , which fixes  $N = 3$ , at most **twenty**;*
- (ii) (saturation) *if in addition  $H_{\text{prod}}(1)$  and  $H_{\text{prod}}(2)$  hold, the products are amino acids and, under  $H_{\text{clu}}$ , the number of distinct ones is **exactly twenty**.*

*Proof.* (i) By Corollary 4.6 the product is a function on  $X$ , and the image of a function on a finite set has at most  $|X|$  elements; Theorem 5.1 supplies  $|X|$ . (ii)  $H_{\text{prod}}(1)$  makes the function total and amino-acid valued, so the image is a set of amino acids;  $H_{\text{prod}}(2)$  makes it injective, so the image has exactly  $|X|$  elements.  $\square$

**Remark 5.5** (The two clauses are not of the same kind). Clause (i) is a bound and rests on two hypotheses whose content is physical but whose form is modest: that the product depends on the cluster's marker assignment alone, and that the pathway does not involve the initiator. Clause (ii) asserts that the system uses its capacity, and that is a substantive claim about the chemistry which this paper does not establish. The statement is written in two clauses so that a reader may grant the first without the second, and the abstract says twenty is a ceiling before it says it is a count.

**Corollary 5.6** (The codon is a triplet). [D] *Assume  $H_{\text{clu}}$  and  $H_{\text{int}}$ . A cluster has three catalyst–substrate interfaces, each producing one nucleobase, so the output of a cluster is a base triplet. The length of the codon is therefore a consequence of the length of the cluster and is not assumed.*

**Remark 5.7** (Why these are hypotheses and not [E]. remarks) Corollary 5.6 is a proof, so the chemical facts it uses may not be [E] claims: by the convention of [Sch26b, §4.3] an empirical instantiation claim enters no proof.  $H_{\text{clu}}$  and  $H_{\text{int}}$  are therefore registered as realization hypotheses at Section 2.4, with their molecular reading confined to Remark 2.9. An earlier draft of this section cited the empirical claim directly and is corrected; the item is in the defect ledger.

**Remark 5.8** (Two questions, one input). Corollary 5.6 and Theorem 5.4 answer *why three* and *why twenty* from the single chemical fact that the cluster is six units. Neither number is used to justify the other, and the standard genetic code is used in neither. The conventional account derives the triplet from the alphabet size; here both descend from the carrier, and by Remark 5.3 the conventional derivation is unavailable in any case.

## 5.3 What capacity forbids

**Proposition 5.9** (Parity). [A] *Every class has size two or four, so every fibre of a function on  $X$ , pulled back to the sixty-four marker assignments, has even cardinality.*

*Sch26e, Prop. 3.6(i).* : no class has size one, the element flip having no fixed point. A fibre of a class function is a union of classes.  $\square$

**Corollary 5.10** (No assembly-level operator produces the standard code). [A] *The standard genetic code assigns methionine to AUG alone and tryptophan to UGG alone, two fibres of cardinality one. By Proposition 5.9 it is therefore not a function on  $X$ , and no operator satisfying  $H_{\text{prod}}(0)$  and  $H_{\text{prod}}(0')$  produces it.*

**Remark 5.11** (This is the boundary of the account, stated at its widest). Corollary 5.10 is one line and it is the honest limit of the paper. The property that explains twenty — that the product is attributable to the assembly, hence constant on classes — is the same property that forbids the production of the standard code. The two cannot be separated, and a reader who wants an assembly-level operator reaching all sixty-four assignments is asking for the capacity theorem to be false.

It is also a strictly stronger obstruction than the fibre-profile argument of [Sch26e, Prop. 3.15(ii)] and of [Sch26f], which compares the multiset of fibre sizes. Parity alone suffices, and it needs no adjacency, no matching and no coarsening.

**Corollary 5.12** (Termination is not an assembly quantity). *[A] The codons UAG and GCU lie in one class of size two (class 4 of [Sch26f, Tab. 1], the two marker assignments FFBBFBB and BBFBFF, which are exchanged by flip). A function on  $X$  cannot assign termination to the first and an amino acid to the second. Termination is therefore supplied by the reading and not by the cluster.*

**Remark 5.13** (The corollary is not a defect). *[E] In the intended reading termination is recognised by release factors, which are proteins of the translating apparatus and not a cognate chemistry carried by the codon. Corollary 5.12 derives that allocation rather than assuming it: the formalism requires termination to live at the reading level, and that is where it is observed. The remark enters no proof, and no claim is made about any organism.*

## 5.4 The count belongs to the full frame group

**Proposition 5.14** (Sensitivity). *[A] The number of orbits on the sixty-four marker assignments of a six-site cluster, under each subgroup of  $\text{Fr}$ , is*

| group acting                                              | orbits |
|-----------------------------------------------------------|--------|
| trivial                                                   | 64     |
| $\langle \text{flip} \rangle$ (projection reference only) | 32     |
| $\langle \text{rev} \rangle$ (reading direction only)     | 36     |
| $\langle \text{flip} \cdot \text{rev} \rangle$            | 36     |
| $\text{Fr} = \langle \text{flip}, \text{rev} \rangle$     | 20     |

*Only the full frame group gives twenty.*

*Proof.* By Burnside's lemma on each subgroup, or by exhaustive enumeration; check V1 of Appendix B.  $\square$

**Remark 5.15** (The proposition is a conditional, not a defence). Proposition 5.14 is stated because it makes the paper's dependence visible, not because it supports the count. Twenty is available only if both generators act, which by [Sch26d, Prop. 6.7] is to say that both the projection reference and the reading direction are re-descriptions of one cluster rather than data of it. For the *codon* they are not: the initiator supplies  $\varepsilon$  and  $f$ , and with the initiator in view the acting group is  $\langle \text{flip} \rangle$  and the count is thirty-two. For the *product* they are, by  $H_{\text{prod}}(0')$ .

So the difference between twenty and thirty-two is exactly the difference between the two branches of the cluster, and the paper's central empirical commitment is the one that decides it. A reader who rejects  $H_{\text{prod}}(0')$  should read this paper as a theorem about thirty-two, and Remark 4.2 says

so.

**Remark 5.16** (The family at other lengths). *[A] The capacity is not a property of the number twenty but of the cluster length: 2, 6, 20, 72, 272, 1056 at  $N = 1, \dots, 6$ . The claim of this paper is therefore a structural law with a free parameter that the chemistry sets, and not an assertion that twenty is distinguished among integers. A four-interface cluster would carry a ceiling of seventy-two, which is a falsifiable corollary and is Target 8.6.*

## 6 The Production Operator

Section 5 bounds the number of products without saying what determines one. This section determines what the formalism permits  $\Delta$  to depend on, which turns out to be a short and specific list, and states the obligations a pathway argument must discharge. It establishes no chemistry; Section 7 illustrates one.

### 6.1 What $\Delta$ may depend on

**Proposition 6.1** ( $\Delta$  is a function of the registry class). *[D] Assume  $H_{\text{mk}}$ ,  $H_{\text{stack}}$ ,  $H_{\text{frm}}$ ,  $H_{\text{prod}}(0)$  and  $H_{\text{prod}}(0')$ . Then there is a unique  $\widehat{\Delta}$  with*

$$\Delta(\mathfrak{R}) = \widehat{\Delta}(\overline{D}(m)),$$

*$m$  being any description of  $\mathfrak{R}$  and  $\overline{D}$  the registry sequence modulo reversal of [Sch26e, Def. 2.11]. Explicitly,  $\Delta$  is a function of five binary entries — the agreements and disagreements of the five adjacent pairs of units — read modulo end-to-end reversal, and of nothing else. There are*

$$\frac{2^5 + 2^3}{2} = 20$$

*such, in agreement with Theorem 5.1.*

*Proof.* By Corollary 4.6,  $\Delta$  is a function on  $X$ ; by [Sch26e, Thm. 3.3(i)] the orbit map of  $\text{Fr}$  is  $m \mapsto \overline{D}(m)$ , and every function on  $X$  therefore factors uniquely through it. The count is the number of binary sequences of length five modulo reversal:  $2^5$  sequences, of which  $2^3$  are palindromic, giving  $(2^5 + 2^3)/2 = 20$  by Burnside's lemma for the two-element group. This is the third proof of [Sch26e, Lem. 3.1] read as a statement about  $\Delta$ .  $\square$

**Corollary 6.2** (Absolute orientation is not an admissible argument). *[D] Under the hypotheses of Proposition 6.1, no pathway variable referring to the absolute orientation of a single unit can determine  $\Delta$ : the marker of any fixed site*

is inverted by flip, which leaves  $\overline{D}$  and hence  $\Delta$  unchanged. Nor can any variable distinguishing one end of the cluster from the other: rev reverses  $D$  and leaves  $\overline{D}$  unchanged.

*Proof.*  $D_n = m_n + m_{n+1}$  is invariant under  $m \mapsto m + 1$  and is reversed by reversal of the site index, by [Sch26e, Thm. 6.14(i)] of the predecessor's numbering;  $\overline{D}$  is the class modulo reversal. A function of a datum that changes under an operation fixing  $\Delta$  cannot determine  $\Delta$ .  $\square$

**Remark 6.3** (The chemical content of the corollary). [E] Corollary 6.2 is the sharpest constraint this paper places on a synthesis mechanism, and it is stated before any mechanism is proposed so that it cannot be accommodated after the fact. The discriminating variable must be a *relation* between consecutive units — whether two adjacent aromatic planes present the same or opposite sense — and must be read symmetrically from either end. A mechanism whose outcome is fixed by the absolute sense of a particular unit, or by which end the reaction starts from, is excluded, and excluded on structural grounds rather than chemical ones. This remark enters no proof.

## 6.2 The invariants available

**Proposition 6.4** (The derived invariants and their resolution). [A] The following are functions on  $X$  and are the scalars naturally derived from  $\overline{D}$ .

- (i) The weight  $|D| \in \{0, \dots, 5\}$ , the number of disagreeing adjacent pairs. Its fibres have sizes 1, 3, 6, 6, 3, 1, so  $|D|$  separates  $X$  into six parts and cannot by itself determine  $\Delta$ .
- (ii) The class size  $|O| \in \{2, 4\}$ . A class has size two exactly when  $\overline{D}$  is palindromic, that is when the agreement pattern reads the same from both ends; eight classes are of size two and twelve of size four.
- (iii) The central entry  $\overline{D}_3$ , which on the size-two classes distinguishes the stabiliser  $\{\text{id}, \text{rev}\}$  from  $\{\text{id}, \text{fliprev}\}$ , four classes each.

No combination of (i)–(iii) separates  $X$ :  $\overline{D}$  itself is the complete invariant.

*Proof.* (i) and (ii) by enumeration of the twenty classes; check V5 of Appendix B. The equivalence of size two with palindromic  $\overline{D}$  is [Sch26e, Prop. 3.6(iii)], and (iii) is [Sch26e, Prop. 3.6(iv)]. That (i)–(iii) do not separate  $X$  follows from (i): six weight values cannot separate twenty classes, and refining by the two remaining binaries gives at most twenty-four cells but does not in fact separate, as the enumeration shows.  $\square$

**Remark 6.5** (What a pathway argument may therefore use). A mechanism may be parametrised by  $|D|$ , by the class size, by the central entry, or by any other function of  $\overline{D}$ , and by no datum outside that list. Proposition 6.4(i) says that a mechanism smooth in  $|D|$  alone cannot determine the product, since six values cannot separate twenty classes; so a complete account must be sensitive to the *pattern* of agreements and not only to their number. That is a testable structural claim about any candidate pathway and is the form in which Section 7's illustration should be read.

## 6.3 The obligations

**Remark 6.6** (Three obligations, and which is load-bearing). Proposition 6.1 is conditional on  $H_{\text{prod}}(0)$  and  $(0')$ , and Theorem 5.4(ii) additionally on (1) and (2). The obligations a chemical account must discharge are therefore these, in order of weight.

*Locality*,  $H_{\text{prod}}(0)$ . That the product depends on the marker assignment alone. A pathway whose outcome varies with temperature, concentration, the presence of neighbouring clusters, or the order in which units assembled would defeat Proposition 6.1 without being in any way chemically implausible. This is the load-bearing obligation and it is the one a synthesis programme is least likely to test, because a single successful determination under fixed conditions does not bear on it.

*Initiator independence*,  $H_{\text{prod}}(0')$ . That the pathway does not involve the initiator. By Remark 3.3 this is what decides between twenty and thirty-two.

*Totality and injectivity*,  $H_{\text{prod}}(1)$  and (2). That every class produces an amino acid and that distinct classes produce distinct ones.

**Proposition 6.7** (Injectivity is not reachable class by class). [A] A family of determinations, one for each class, does not entail  $H_{\text{prod}}(2)$ . Nineteen determinations leave the twentieth unconstrained by them, and a complete family of twenty determinations may assign one amino acid to two classes, in which case the image has at most nineteen elements and Theorem 5.4(ii) fails while clause (i) stands.

*Proof.* By witness: the constant family of Proposition 3.1(i) is a determination for every class and is not injective.  $\square$

**Remark 6.8** (Injectivity as a falsification test). Proposition 6.7 is worth stating positively. Because the determinations are made independently, their joint injectivity is a non-trivial prediction rather than a construction: a programme that determines the twenty products one at a time

and finds them distinct has tested  $H_{\text{prod}}(2)$  and not assumed it. A collision — two classes with the same product — refutes the saturation clause and leaves the ceiling intact, which is the most informative negative result available, and it is why the twenty determinations should be carried out and recorded even where the answer is expected.

## 6.4 What this section does not supply

**Remark 6.9** (No criterion selects  $\Delta$ ). This paper exhibits no rule determining  $\hat{\Delta}$  on the twenty classes. Proposition 6.1 says what its arguments are; Proposition 6.4 says which scalars are available; neither says what the function is. In particular  $\Delta$  is not constructed here, and by Remark 3.4 it is not the coarsening of the companion paper, which is selected by a criterion this paper does not use and which depends on a datum this paper does not use. Determining  $\hat{\Delta}$  is Target 8.4.

**Remark 6.10** (What Section 7 imports, and what it does not). [A] A disclosure is owed here, because Section 7 displays two particular discriminations and the reader is entitled to know how they were chosen.

The two cases of Section 7 — the largest and one of the smallest separations available — were selected for expositional contrast, and the selection was informed by the candidate sets of [Sch26f, Prop. 5.8], which are computed from the standard genetic code. The choice of what to display therefore has a provenance in that paper.

No statement of Section 7 depends on it. What Section 7 asserts is that the pathway, applied to a cluster of a named class, yields a named amino acid. That assertion is chemical, is checked against no codon table, and would stand if the companion paper did not exist. The distinction is between the order of exposition and the order of proof, and Remark 2.4 is unaffected: the standard genetic code enters no hypothesis and no construction of this paper. The item is recorded in the defect ledger so that a reader need not reconstruct it.

## 7 Two Discriminations

Section 6 determines what  $\Delta$  may depend on. This section exhibits the architecture of the pathway in those terms and works two cases.

The objective is mathematical and is stated at the outset so that the section is not read as something it is not. What is claimed here is that two specific discriminations are *expressible in the invariants Section 6 permits*, and that so expressed they are consistent with the capacity theorem. Whether the chemistry is correct is a separate question, not addressed and not required:

the reactants are identified — they are the complex intermediary products described in the synthesis of US 12,195,423 — and the process is outlined to the level the mathematical claim needs and no further. A full chemical development, with the synthesis steps, the intermediates and the liberating step, is forthcoming as a separate treatise and is not presupposed here.

The section is therefore an illustration and not a method in two senses: two determinations are not twenty, and by Proposition 6.7 twenty would still not establish injectivity. The provenance of the two cases is disclosed at Remark 6.10.

## 7.1 The architecture in invariant terms

**Definition 7.1** (Transforming vessel and flanking environment). A pathway is *vessel-localised* when the chemical transformation yielding the product occurs at one distinguished vessel of the cluster, the remaining vessels acting on the outcome without themselves transforming. The distinguished vessel is the *transforming vessel* and the others are its *flanking environment*.

**Proposition 7.2** (A vessel-localised pathway requires odd  $N$ ). [A] *Reversal carries the  $i$ -th vessel of a cluster of  $N$  vessels to the  $(N + 1 - i)$ -th. It therefore fixes a vessel if and only if  $N$  is odd, in which case the fixed vessel is unique and central. Consequently a vessel-localised pathway whose product satisfies  $H_{\text{prod}}(0')$  can distinguish a vessel only for odd  $N$ , and only the central one: a designation of any other vessel is reversed by a change of reading direction and is not a datum of the cluster.*

*Proof.* The map  $i \mapsto N + 1 - i$  on  $\{1, \dots, N\}$  has a fixed point iff  $N + 1 - i = i$  for some  $i$ , that is iff  $N$  is odd, and then  $i = (N + 1)/2$  is unique. By [Sch26d, Prop. 6.7] reversal of the reading direction is a change of frame, so by Corollary 6.2 no frame-dependent designation determines  $\Delta$ . Verified at check V6.  $\square$

**Corollary 7.3** (At  $N = 3$  the transforming vessel is the middle one). [D] *Under  $H_{\text{clu}}$ , a vessel-localised pathway transforms the second of the three vessels, and the first and third form its flanking environment, exchanged by reversal.*

**Remark 7.4** (A consistency result, not a second derivation of  $N = 3$ ). Proposition 7.2 excludes  $N = 2$  and  $N = 4$  for a vessel-localised pathway, and  $N = 1$  admits no flanking environment at all, so  $N = 3$  is the least admissible length. That agrees with  $H_{\text{clu}}$ , which fixes  $N = 3$  on independent chemical grounds.

The agreement is a check and not a second derivation.

Vessel-localisation is itself a chemical premise — it is clause (6) of the pathway description of Remark 7.5 — so the argument replaces one chemical input by another rather than removing it. What it does establish is that the two inputs are compatible, and that an architecture transforming a terminal vessel would be incompatible with  $H_{\text{prod}}(0')$  and so would forfeit the capacity theorem.

**Remark 7.5** (The pathway, in outline). [E] The following is the chemical reading and enters no proof. It is given to the level the mathematical claims of this section require; the synthesis steps and their intermediates belong to the separate treatise named in the preamble to this section.

The cluster is the complete reaction vessel, comprising three complex molecular structures. Its middle unit carries the catalytic chrysene associated with the second nucleotide of the codon, and it is the one that undergoes transformation; the two outer tethered units act catalytically, maintain the integrity of the vessel, and shape the outcome. The intermediate is the complex structure described in the synthesis of US 12,195,423, a nucleotide tethered to a neighbouring chrysene catalyst — which is the realization of the intermediary of Remark 4.11, and is again not a free nucleobase.

The construction is internal. It requires no external substance beyond what is needed to proceed and to liberate the product, and intermediates are not released until the synthesis is complete, so that the outcome is a function of the completed cluster and not of any partial state. The backbone is formed in the reaction with nitric oxide at the phosphoethoxy arm; the residue is determined by the interaction of that arm with the chrysene core as shaped by the vessel.

Two further data enter the outcome: the orientation of the aromatic rings together with the substitution pattern, and the interaction of the vessel with the neighbouring nucleobases of the flanking units.

**Remark 7.6** (Locality is the formal content of internality). [A] Two clauses of Remark 7.5 are the chemical content of  $H_{\text{prod}}(0)$  and should be read as such: that the construction requires no external substance beyond the liberating step, and that no intermediate is released before completion. Together they assert that the product is a function of the completed cluster and of nothing outside it, which is locality.  $H_{\text{prod}}(0)$  is therefore not an idealisation imposed on the chemistry but a statement the chemistry makes; Remark 6.6 records that it nevertheless remains the load-bearing and least-tested obligation, since a determination under fixed conditions does not probe it.

**Remark 7.7** (The alternate substitution). [E] The account of Remark 7.5 is given for substitution at C6 or C12. The configuration substituted at C5 or C11 is an alternate of the same architecture; by A8 of [Sch26c] the locant pairs (12, 11) and (6, 5) are perimeter-adjacent primary–dual cipher pairs,  $\delta_{\text{dc}}(12) = 11$  and  $\delta_{\text{dc}}(6) = 5$ , so the two configurations are related by the cipher pairing. Nothing below distinguishes them and no statement of this paper depends on which is realized.

## 7.2 Which nucleobase may be read, and which may not

The pathway description refers to the neighbouring nucleobases. Taken literally, that is not an admissible argument for  $\Delta$ , and the following fixes the form in which it is.

**Proposition 7.8** (Amino/keto is frame-invariant; the base is not). [A] For a vessel with coordinates  $(b_1, b_2)$ , the registry  $b_1$  is independent of the frame and  $b_2$  is determined by it ([Sch26d, Prop. 4.4]). Under any letter assignment  $L$  carrying the registry partition to a partition of  $\{A, C, G, U\}$ , the class of a vessel is therefore a datum of the cluster while the identity of its base is not. Under the assignment of [Sch26f, Def. 2.16] the two registry classes are  $\{G, U\}$  and  $\{A, C\}$ , the keto and amino bases.

Sch26d, Prop. 4.4. for the two coordinates and [Sch26d, Prop. 3.16] for the fibres of  $b_1$  in marker coordinates. The identification of the fibres with  $\{G, U\}$  and  $\{A, C\}$  under that assignment is immediate from it, and by Proposition 4.12 nothing below depends on the assignment.  $\square$

**Corollary 7.9** (The admissible form of the reading). [D] A mechanism whose outcome depends on the amino or keto character of a flanking vessel is admissible: that character is  $b_1$ , hence an entry of the registry sequence, hence by Proposition 6.1 a permitted argument of  $\Delta$ . A mechanism whose outcome depends on the identity of a flanking base — adenine as against cytosine, or guanine as against uracil — is not admissible, those pairs being exchanged by flip, which fixes  $\Delta$  by Corollary 6.2.

**Remark 7.10** (What the corollary permits and forbids). [E] The chemical statement that an exocyclic amino group is available at a flanking position, and is not when the flanking base is keto, is of the admissible kind: adenine and cytosine both present the group and are the two members of one registry class, guanine and uracil present none and are the other. The statement survives translation into the formalism

unchanged.

The statement that adenine specifically, as against cytosine, fixes the outcome would not survive. It would make  $\Delta$  depend on  $b_2$ , hence on the projection reference, hence on a convention of the reading; by Theorem 5.4 the ceiling would rise from twenty to sixty-four and the paper would have no theorem. Any mechanism proposed for the remaining eighteen classes must be checked against this corollary before it is checked against anything else.

### 7.3 First case: the largest separation

**Example 7.11** ( $\overline{D} = 00000$  against  $\overline{D} = 00001$ ). [E] The two classes are

$$\begin{array}{llll} \overline{D} = 00000, & |D| = 0, & \text{size } 2, & \text{vessel registries } (0, 0, 0); \\ \overline{D} = 00001, & |D| = 1, & \text{size } 4, & \text{vessel registries } (0, 0, 1). \end{array}$$

The first is the wholly agreeing cluster: every adjacent pair of units presents the same sense, and all three vessels are keto. The second differs in one entry, at a terminal vessel. The pathway yields phenylalanine on the first and glycine on the second. In the terms of Remark 7.5, the internal synthesis at the tethered unit of the second nucleotide turns on the size relation between the adjacent nucleobases and the interleaving substitution arms: the wholly agreeing cluster presents the arms and the flanking bases in register, admitting the large aromatic residue, while a single disagreement at a terminal vessel removes the accommodation and leaves the residue unelaborated.

**Remark 7.12** (What the first case illustrates). It is the extreme case in two senses and is displayed for that reason. The separation of the two products is the largest available, and the separation of the two classes is the smallest: a single entry of the registry sequence, the minimum possible difference, carries the maximum difference in outcome. The class  $\overline{D} = 00000$  is also the most symmetric object in  $X$  — the unique class of weight zero, of size two, with palindromic registry — and is therefore the least representative of the twenty. A mechanism that succeeds here has been tested on the easiest instance.

### 7.4 Second case: a small separation

**Example 7.13** ( $\overline{D} = 00110$  against  $\overline{D} = 00111$ ). [E] The two classes are

$$\begin{array}{llll} \overline{D} = 00110, & |D| = 2, & \text{size } 4, & \text{vessel registries } (0, 1, 0); \\ \overline{D} = 00111, & |D| = 3, & \text{size } 4, & \text{vessel registries } (0, 1, 1). \end{array}$$

Both have the middle vessel unaligned, that is amino; they

differ in whether one terminal vessel is amino or keto. The pathway yields glutamate on the first and asparagine on the second.

In the terms of Corollary 7.9, the tethered substituted chrysene precursor acquires the exocyclic amino group when a terminal vessel presents an amino-class base, and does not when both terminal vessels are keto. The discrimination is therefore read off a single entry of the registry sequence, at a terminal position, and is of the admissible form.

**Remark 7.14** (Why the second case is the harder one). The two residues are separated by about thirteen cubic ångström in volume, carry identical Kyte–Doolittle hydrophathy, and differ chiefly in hydrogen-bonding pattern. No account turning on size or on hydrophobicity can separate them, and the mechanism is obliged to be specific: the discriminating datum is the presence or absence of a donor at a flanking position, which is exactly the amino/keto character that Proposition 7.8 certifies as frame-invariant.

This is the case on which the account should be judged. The first case could be passed by any mechanism sensitive to size; the second can be passed only by one reading the registry.

### 7.5 What the two cases establish

**Remark 7.15** (Two cases, and their limits). Three limits, stated together so that none is inferred from the absence of the others.

Two determinations are not a determination of  $\hat{\Delta}$ : eighteen classes are untreated, and Remark 6.9 records that this paper exhibits no rule.

Twenty determinations would not establish  $H_{\text{prod}}(2)$  as a matter of course, by Proposition 6.7, though they would test it; nor would they establish  $H_{\text{prod}}(0)$ , which no determination under fixed conditions probes, by Remark 7.6.

And the two cases are not independent evidence for the capacity theorem, which is proved without them. They bear on whether the ceiling is attained, that is on  $H_{\text{prod}}(1)$  and (2), and on nothing else.

**Remark 7.16** (An observation, used nowhere). [A] The four classes named in Examples 7.11 and 7.13 are assigned phenylalanine, glycine, glutamate and asparagine by the coarsening  $\varphi$  of [Sch26f, Thm. 5.10], which agrees with the outcomes recorded here. The agreement is recorded and used in no proof. It is the first evidence bearing on [Sch26f, Target 8.16], whether the realized map is that coarsening, and by Remark 3.4 it is not assumed anywhere in this paper; four agreements out of twenty are also well short of what

that target asks, the five classes fixed by the adjacency of that paper being free predictions and the remaining fifteen carrying the content.

**Remark 7.17** (The scope of the chemical claim). [E] Three things are asserted chemically in this section and no more: that the reactants are the complex intermediary products described in the synthesis of US 12,195,423, a nucleotide tethered to a neighbouring chrysene catalyst; that the process proceeds as outlined in Remark 7.5; and that on the four classes named it yields the four products recorded. The synthesis steps, the intermediates, the substitution patterns and the liberating step are not given here and are forthcoming as a separate chemical treatise.

The mathematical content of this section is independent of that development and is exhausted by three statements: Corollary 7.3, that a vessel-localised pathway must transform the central vessel; Corollary 7.9, that the amino or keto character of a flanking vessel is an admissible argument of  $\Delta$  and the identity of a flanking base is not; and Remark 7.6, that internality is the chemical content of  $H_{\text{prod}}(0)$ . Each is proved from the formalism and would stand if the two worked cases were withdrawn.

Accordingly nothing in this paper is offered for chemical review, and no statement of Sections 4 to 6 is conditional on the treatise. What the two cases show is that the discriminations the chemistry proposes are of a form the mathematics admits — which is a statement about the mathematics.

## 8 Non-Implications and Targets

Section 3 states the non-implications before the constructions they guard. This section restates them in one line each, so that the list may be checked against the completed text, and collects the Targets in one place. It adds no argument.

### 8.1 The non-implications restated

- (i) Twenty classes do not give twenty amino acids; the theorem is a ceiling and the count needs saturation (Proposition 3.1(i), Remark 5.5).
- (ii) The capacity theorem yields no assignment of amino acids to codons, and forbids one: parity excludes any table with a fibre of odd cardinality, the standard code included (Proposition 3.1(ii), Corollary 5.10).
- (iii) The partition result of the companion paper and the capacity theorem do not bear on one another; one depends on the letter assignment and on the standard code, the other on neither (Proposition 3.1(iii), Remark 4.13).

- (iv) Twenty at  $N = 3$  says nothing about other lengths; the counts are 2, 6, 20, 72, 272 (Proposition 3.1(iv), Remark 5.16).
- (v) The capacity theorem and the nucleobase alphabet are independent in both directions (Proposition 3.1(v), Proposition 4.12).
- (vi) Sufficiency of the codon supply constrains no cluster length;  $|X| < 4^N$  always, so the classical argument for the triplet cannot be run here (Proposition 3.1(vi), Remark 5.3).
- (vii) A bound on the products of a chrysene cluster is not a statement about the biological alphabet (Proposition 3.1(vii)).

Three further non-implications are stated at the point of use and are not repeated as clauses: that a six-unit cluster gives twenty classes only if both frame data are re-descriptions and not data of the cluster (Remark 3.3); that  $\Delta$  is not the coarsening of the companion paper (Remark 3.4); and that a family of per-class determinations does not entail injectivity (Proposition 6.7).

### 8.2 Inventory of Targets

Seven Targets are opened here.

**Target 8.1** (Locality). [T] Determine whether the product of a cluster varies with any datum outside it — temperature, concentration, the presence of neighbouring clusters, the order in which units assembled — and thereby discharge or refute  $H_{\text{prod}}(0)$ . Hypotheses required: a preparation in which the class of a cluster is known independently of its product, and in which one external variable can be moved while the class is held fixed. By Remark 7.6 the clauses of the pathway description assert locality; by Remark 6.6 no determination under fixed conditions tests it.

**Target 8.2** (Initiator independence). [T] Determine whether the amino-acid pathway involves the initiator of [Sch26d, Def. 9.1], and thereby discharge or refute  $H_{\text{prod}}(0')$ . This decides between a ceiling of twenty and a ceiling of thirty-two, by Remark 3.3. A cluster prepared with the initiator at either end should yield the same product if the hypothesis holds.

**Target 8.3** (Saturation). [T] Determine the products of all twenty classes, and thereby discharge or refute  $H_{\text{prod}}(1)$  and  $H_{\text{prod}}(2)$ . By Remark 6.8 the determinations are made independently, so their joint injectivity is a prediction and not a construction; a collision refutes saturation and leaves the ceiling intact, which is the most informative negative result the programme admits.

**Target 8.4** (The operator). [T] Determine  $\hat{\Delta}$  as a function on the twenty registry classes. By Proposition 6.1 its arguments are the five agreements read modulo reversal and nothing else; by Proposition 6.4 the weight  $|D|$  has only six values and cannot by itself suffice, so a complete account must be sensitive to the pattern of agreements and not merely to their number. Section 7 supplies two values of twenty.

**Target 8.5** (Vessel-localisation). [T] Establish, rather than assume, that the pathway is vessel-localised in the sense of Definition 7.1. By Proposition 7.2 a localised pathway compatible with  $H_{\text{prod}}(0')$  must transform the central vessel and requires odd  $N$ ; Remark 7.4 records that this agrees with  $H_{\text{clu}}$  and that the agreement is a check and not a derivation. What is open is whether localisation holds at all, a delocalised pathway being consistent with everything proved here and giving the same ceiling.

**Target 8.6** (The ceiling at another length). [T] Prepare or identify a cluster of four catalyst–substrate interfaces and determine whether its product alphabet is bounded by seventy-two. By Theorem 5.4 with Theorem 5.1 the ceiling is the orbit count at the length, and the claim of this paper is a structural law with a free parameter rather than an assertion about the number twenty. This is the cleanest falsifiable corollary the capacity theorem admits, and a four-interface cluster whose products exceeded seventy-two would refute the theorem outright. Note that by Proposition 7.2 such a cluster admits no vessel-localised pathway, so a negative result would bear on Target 8.5 as well.

**Target 8.7** (Which classes are occupied). [T] Determine which of the twenty classes are realized by admissible clusters. [Sch26e, Prop. 3.11] establishes that every class contains an admissible marker assignment under the constraint set of [Sch26d, Def. 9.2], but that set is a construction and not an axiom, and the prior question — whether the marker is information-bearing at all — is [Sch26d, Prop. 3.8] and remains open. The capacity theorem is unaffected either way, being a bound; the count of clause (ii) is not, since an unoccupied class produces nothing and  $H_{\text{prod}}(1)$  fails for it.

### 8.3 Which are load-bearing

**Remark 8.8** (Three of the seven). Their number is not offered as evidence of anything. Target 8.1 is the deepest.  $H_{\text{prod}}(0)$  is what makes  $\Delta$  a function of the carrier at all, and without it Proposition 6.1 and Theorem 5.4 both fail. It is also the hypothesis a

synthesis programme is least likely to test, for the reason given at Remark 6.6.

Target 8.2 decides the number. Twenty and thirty-two are both theorems of this formalism; which one describes the chrysene system is a chemical question and the paper does not settle it.

Target 8.3 decides whether the paper's second clause stands. The ceiling is proved; that it is attained is not, and the twenty determinations are what would establish it.

The remaining four are not load-bearing. Target 8.4 asks for more than the paper claims; Target 8.5 concerns a premise of Section 7 and not of Section 5; Target 8.6 is a corollary test; and Target 8.7 bears on clause (ii) only.

### 8.4 Status of the predecessors' programme

**Remark 8.9** (No target is discharged). No deferred target of [Sch26c] is discharged here, none of the nine of [Sch26d], none of the eight of [Sch26e], and none of the nine of [Sch26f]. Three are touched and the manner is recorded so that it is not mistaken for a discharge.

[Sch26d, Prop. 3.8], that the marker is information-bearing, is carried unchanged into Target 8.7; nothing here bears on it.

[Sch26d, Target 10.4], the marker binary and emergent chirality, is not touched at all: no statement of this paper concerns configuration, and Remark 7.5 makes no claim about the handedness of the product.

[Sch26f, Target 8.16], whether the realized map is that paper's coarsening, acquires in Remark 7.16 four agreements out of twenty, of which none is among the five that paper fixes without a physical criterion. That is an observation and not a discharge, and Remark 3.4 records that the identification is assumed nowhere here.

**Remark 8.10** (What this paper adds to the inventory). Two of the Targets above are of a kind this corpus has carried only once before, in [Sch26f]: Targets 8.1, 8.2, 8.3 and 8.6 ask for measurements rather than for proofs or for data internal to the formalism. Target 8.6 is of a further kind again, being the first statement of the corpus that a single negative observation would refute outright.

## 9 Conclusion

The question this paper set out to answer was why the amino-acid alphabet has twenty members and not nineteen or twenty-one. The answer offered is a capacity argument from the structure of a physical carrier, and its distinguishing feature is negative: the standard genetic code is used in no step of it.

**A class is a cluster.** The classification of [Sch26e] counts orbits of marker strings under the frame group. Proposition 4.4 shows those orbits are not equivalence classes of descriptions but the objects described: the map from clusters to classes is a bijection, and the two or four marker assignments a class contains are one physical cluster read under the available frames. The statement appears in the predecessor only as a table caption, and a caption cannot carry a hypothesis register. Making it a proposition has a consequence that runs through the paper: frame-invariance of a product is not a hypothesis but a corollary (Corollary 4.6), and the assumptions move to where they can be attacked — that the product depends on the cluster's marker assignment alone, and that the pathway does not involve the initiator.

**The capacity theorem.** A product satisfying those two takes at most  $2^{2N-2} + 2^{N-1}$  values, which at  $N = 3$  is twenty (Theorem 5.4(i)). That is a ceiling. Exactly twenty requires in addition that every class produce an amino acid and that distinct classes produce distinct ones, which is the assertion that the system uses its capacity and which this paper does not establish (Theorem 5.4(ii), Remark 5.5). The witness at Proposition 3.1(i) — the constant map, which satisfies both standing hypotheses and has an image of one element — shows how much of the result the saturation clause carries.

**Two questions, one input.** With the six-unit cluster fixed on chemical grounds independently of codons and of the alphabet size, the triplet codon follows from the carrier (Corollary 5.6) and so does the ceiling of twenty. Neither number is used to justify the other and neither is used to justify the cluster (Remark 2.9). The conventional derivation runs the other way — twenty amino acids need more than sixteen codons, so codons are triplets — and is in any case unavailable here:  $|X|/4^N = \frac{1}{4} + 2^{-(N+1)} < 1$  at every length, so the alphabet never outruns the codon supply and the sufficiency condition distinguishes no  $N$  (Proposition 5.2, Remark 5.3).

**What capacity forbids.** The property that explains twenty also forbids something, and the two cannot be separated. Every class has size two or four, so every fibre of an assembly-level map is even; the standard genetic code has two fibres of size one, methionine at AUG and tryptophan at UGG; hence no operator satisfying the two standing hypotheses produces the standard code (Proposition 5.9, Corollary 5.10). One line, and it needs no adjacency, no matching and no coarsening, which makes it strictly stronger than the fibre-profile argument of [Sch26e, Prop. 3.15(ii)].

The same argument allocates termination. The codons UAG and GCU lie in one class, so no function on classes assigns termination to the first and an amino acid to the second: termination is supplied by the reading and not by the cluster (Corollary 5.12). The formalism therefore requires termination to be recognised

by the translating apparatus rather than carried as a cognate chemistry, which is where it is observed. That conclusion is derived here and not assumed.

**The operator, and what it may see.**  $\Delta$  is a function of the registry class: five binary entries recording whether each adjacent pair of units agrees in marker, read modulo end-to-end reversal, and nothing else (Proposition 6.1). No pathway variable referring to the absolute orientation of a unit, or distinguishing one end of the cluster from the other, can determine it (Corollary 6.2). Two consequences bear directly on the chemistry. A vessel-localised pathway must transform the central vessel and requires odd  $N$ , since reversal fixes no other vessel (Proposition 7.2); and a mechanism may read the amino or keto character of a flanking vessel, that character being the frame-independent coordinate, but may not read the identity of a flanking base (Proposition 7.8, Corollary 7.9).

Two discriminations are worked in those terms. The wholly agreeing cluster against one differing in a single terminal entry, where the minimum difference in the registry carries the maximum difference in product; and two classes differing only in whether a terminal vessel is amino or keto, where the residues are separated by neither size nor hydrophathy and the discrimination is read off one entry of the registry. Both are illustrations. Eighteen classes are untreated, and by Proposition 6.7 twenty determinations would test injectivity rather than establish it.

**What is not done.** No axiom is introduced (Remark 2.2), and A3, A4, A5, A7, A9, A10 and A11 are used nowhere (Proposition 2.5), so nothing here concerns sequence dependence. No deferred target of any predecessor is discharged (Remark 8.9). No criterion determining  $\Delta$  is exhibited (Remark 6.9), and  $\Delta$  is not the coarsening of the companion paper (Remark 3.4). Seven non-implications are recorded (Proposition 3.1) and seven Targets opened, of which three are load-bearing.

Two limits deserve restating rather than leaving to the list. The count is twenty only for the full frame group; it is thirty-two if the initiator is a datum of the cluster rather than a reader of it, and which of the two describes the chrysene system is a chemical question this paper poses and does not settle (Remark 3.3, Target 8.2). And the theorem bounds the products of a chrysene cluster. That the biological alphabet is that set is an identification of the kind [Sch26b, §9.2] records as the corpus's largest unproved premise, and the paper declines it (Proposition 3.1(vii)). The title asks a question the paper answers for a carrier, not for biology.

**What remains.** A carrier of six units admits exactly twenty distinguishable states, and that count is a property of the carrier and of the group of frame changes, not of any codon table. Its output is therefore a base triplet, and any product attributable to it is bounded by twenty. Attaining the bound is a hypothesis;

producing the standard code is impossible at this level; termination belongs to the reading. What a pathway may depend on is fixed to five relative orientations read from either end, and two of the twenty values are worked. The rest is measurement.

## A The Twenty Classes

Table 1 is a computation and not a construction. It is included so that every statement of Sections 4 to 7 may be checked against a finite object rather than against a formula, and it is regenerated by the supplementary script from the definitions of Section 2, importing nothing from this appendix.

**Reading the table.** The representative is the lexicographically least marker assignment of the class in the convention  $F < B$ .  $\overline{D}$  is the registry sequence modulo reversal, of length five, and  $|D|$  its Hamming weight, the number of disagreeing adjacent pairs. The size is the cardinality of the Fr-orbit, which by Proposition 4.4 is the number of frames under which the cluster presents a distinct description. The stabiliser type is  $r$  for  $\{\text{id}, \text{rev}\}$ ,  $\text{fr}$  for  $\{\text{id}, \text{fliprev}\}$  and 1 for the trivial stabiliser, in the sense of [Sch26e, Prop. 3.6(i)].  $\overline{D}_3$  is the central entry, which by [Sch26e, Prop. 3.6(iv)] separates the two families of size-two classes. The last column gives the three *vessel registries* ( $D_1, D_3, D_5$ ), the odd-indexed entries being the registries of the three vessels; by Proposition 7.8 a 0 marks a keto-class vessel and a 1 an amino-class one, under the letter assignment of [Sch26f, Def. 2.16] and under no other statement of this paper.

| #                                                                                                           | Representative | $\overline{D}$ | $ D $ | Size | Stab        | $\overline{D}_3$ | $(D_1, D_3, D_5)$ |
|-------------------------------------------------------------------------------------------------------------|----------------|----------------|-------|------|-------------|------------------|-------------------|
| <i>Size two — palindromic registry; four of type <math>r</math> and four of type <math>\text{fr}</math></i> |                |                |       |      |             |                  |                   |
| 1                                                                                                           | BBBBBB         | 00000          | 0     | 2    | $r$         | 0                | 000               |
| 2                                                                                                           | BBBFFF         | 00100          | 1     | 2    | $\text{fr}$ | 1                | 010               |
| 3                                                                                                           | BBFFBB         | 01010          | 2     | 2    | $r$         | 0                | 000               |
| 4                                                                                                           | BBFBFF         | 01110          | 3     | 2    | $\text{fr}$ | 1                | 010               |
| 5                                                                                                           | BFFFFB         | 10001          | 2     | 2    | $r$         | 0                | 101               |
| 6                                                                                                           | BFFBBF         | 10101          | 3     | 2    | $\text{fr}$ | 1                | 111               |
| 7                                                                                                           | BFBFBF         | 11011          | 4     | 2    | $r$         | 0                | 101               |
| 8                                                                                                           | BFBFBF         | 11111          | 5     | 2    | $\text{fr}$ | 1                | 111               |
| <i>Size four — non-palindromic registry, trivial stabiliser</i>                                             |                |                |       |      |             |                  |                   |
| 9                                                                                                           | BBBBBF         | 00001          | 1     | 4    | 1           | 0                | 001               |
| 10                                                                                                          | BBBBFF         | 00010          | 1     | 4    | 1           | 0                | 000               |
| 11                                                                                                          | BBBBFB         | 00011          | 2     | 4    | 1           | 0                | 001               |
| 12                                                                                                          | BBBFFB         | 00101          | 2     | 4    | 1           | 1                | 011               |
| 13                                                                                                          | BBBFBB         | 00110          | 2     | 4    | 1           | 1                | 010               |
| 14                                                                                                          | BBBFBF         | 00111          | 3     | 4    | 1           | 1                | 011               |
| 15                                                                                                          | BBFFFB         | 01001          | 2     | 4    | 1           | 0                | 001               |
| 16                                                                                                          | BBFFBF         | 01011          | 3     | 4    | 1           | 0                | 001               |
| 17                                                                                                          | BBFBBF         | 01101          | 3     | 4    | 1           | 1                | 011               |
| 18                                                                                                          | BBFBFB         | 01111          | 4     | 4    | 1           | 1                | 011               |
| 19                                                                                                          | BFBBBF         | 10011          | 3     | 4    | 1           | 0                | 101               |
| 20                                                                                                          | BFBFFB         | 10111          | 4     | 4    | 1           | 1                | 111               |

**Table 1: The twenty classes of a six-unit cluster.** By Proposition 4.4 each row is a cluster and not an equivalence class of descriptions. The weight  $|D|$  takes six values with fibre sizes 1, 3, 6, 6, 3, 1, so by Proposition 6.4 it cannot by itself determine  $\Delta$ ;  $\overline{D}$  is the complete invariant. Rows 1 and 9 are the classes of Example 7.11 and rows 13 and 14 those of Example 7.13; the last column is the datum those two discriminations are read from. The numbering is that of [Sch26f, Tab. 1], which orders the size-two classes first and then the size-four, each by  $\overline{D}$ .

## B Concordance, Ledger, and Verification

### B.1 Scope and method

This appendix records, for every prior statement of the corpus that this paper touches, whether it is *preserved*, *extended*, *restated*, *corrected*, *not applicable* or *noted*, against the criterion  $C$  of [Sch26b, Def. 2]. One entry requires comment before the tables.

| $N$ | $2N$ units | $ X  = 2^{2N-2} + 2^{N-1}$ | $4^N$ | $ X /4^N$ |
|-----|------------|----------------------------|-------|-----------|
| 1   | 2          | 2                          | 4     | 0.5000    |
| 2   | 4          | 6                          | 16    | 0.3750    |
| 3   | 6          | <b>20</b>                  | 64    | 0.3125    |
| 4   | 8          | 72                         | 256   | 0.2812    |
| 5   | 10         | 272                        | 1024  | 0.2656    |
| 6   | 12         | 1056                       | 4096  | 0.2578    |
| 7   | 14         | 4160                       | 16384 | 0.2539    |
| 8   | 16         | 16512                      | 65536 | 0.2520    |

**Table 2: The ceiling at other cluster lengths.** The capacity of Theorem 5.4 is the orbit count at the length, so the claim of this paper is a structural law with a free parameter that the chemistry sets and not an assertion about the number twenty;  $N = 3$  is fixed by  $H_{\text{clu}}$  and by nothing else. The last column is  $|X|/4^N = \frac{1}{4} + 2^{-(N+1)}$ , strictly below one at every length: the alphabet never outruns the codon supply, so the classical sufficiency argument for the triplet codon distinguishes no  $N$  and cannot be run here (Proposition 5.2, Remark 5.3). A four-interface cluster carries a ceiling of seventy-two and is Target 8.6.

[Sch26e, Tab. 1] asserts in its caption that the members of a class are the same physical stack read under different conventions. This paper makes that assertion load-bearing — the capacity theorem rests on it — and a caption cannot carry a hypothesis register. It is therefore *restated* as Proposition 4.4 and proved from [Sch26d, Props. 3.4, 3.19, 6.5, 6.7]. Nothing in the predecessor is contradicted; what was recorded there informally is given hypotheses here.

## B.2 Concordance with the classification papers

| Prior statement                                                 | Disposition                                                                                                                                                                 |
|-----------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [Sch26e, Lem. 3.1], the orbit count                             | Preserved; imported at Theorem 5.1 and not reproved                                                                                                                         |
| [Sch26e, Thm. 3.3(i)], the orbit map is $\overline{D}$          | Preserved; it is what makes $\Delta$ a function of five agreements (Proposition 6.1)                                                                                        |
| [Sch26e, Thm. 3.3(iii)], assembly quantities factor through $X$ | Preserved and <i>extended</i> : that theorem presupposes the attribution; Proposition 4.4 supplies it                                                                       |
| [Sch26e, Tab. 1] caption, a class is one stack                  | <i>Restated</i> as Proposition 4.4; see B.1                                                                                                                                 |
| [Sch26e, Prop. 3.6], the stabiliser refinement                  | Preserved; used at Proposition 6.4(ii)–(iii)                                                                                                                                |
| [Sch26e, Cor. 3.7], eight and twelve                            | Preserved                                                                                                                                                                   |
| [Sch26e, Prop. 3.11], realization of every class                | Preserved; carried to Target 8.7                                                                                                                                            |
| [Sch26e, Prop. 3.15(ii)], the fibre profiles differ             | Preserved; noted that Proposition 5.9 is strictly stronger, parity alone sufficing                                                                                          |
| [Sch26e, Prop. 3.15(i)], the multiset coincidence               | Not applicable; used nowhere here                                                                                                                                           |
| [Sch26e, §6], the variational criterion                         | Not applicable; no selection is made                                                                                                                                        |
| [Sch26f, Def. 2.14], <i>cluster</i>                             | Preserved; the naming discipline is inherited                                                                                                                               |
| [Sch26f, Def. 2.16], the letter assignment $L$                  | Noted; used in no statement, by Proposition 4.12                                                                                                                            |
| [Sch26f, Thm. 5.2], the partition criterion                     | Noted; the non-implication is Proposition 3.1(iii)                                                                                                                          |
| [Sch26f, Prop. 5.8], the five fixed classes                     | Noted; it is the source of the candidate pairs displayed in Section 7, ledger item 15                                                                                       |
| [Sch26f, Thm. 5.10], the coarsening $\varphi$                   | Noted; $\Delta \neq \varphi$ is Remark 3.4, and the four agreements of Remark 7.16 are used in no proof                                                                     |
| [Sch26f, Hyp. 2.30], $H_{\text{dual}}(\text{i}), (\text{ii})$   | <i>Restated</i> as $H_{\text{prod}}(0')$ and (2); clause (i) becomes Corollary 4.6 and the assumption moves to $H_{\text{prod}}(0)$ . No clause is weakened or strengthened |
| [Sch26f, Target 8.16]                                           | Not discharged; see Remark 8.9                                                                                                                                              |

### B.3 Concordance with the encoded-stack and axiom papers

| Prior statement                                                       | Disposition                                                                                                                                                                   |
|-----------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [Sch26d, Defs. 3.1–3.3], stack, placement, marker                     | Preserved                                                                                                                                                                     |
| [Sch26d, Prop. 3.4], non-collapse                                     | Preserved; injectivity clause of Proposition 4.4(i)                                                                                                                           |
| [Sch26d, Prop. 3.14], the affine bijection                            | Preserved; Proposition 4.8                                                                                                                                                    |
| [Sch26d, Prop. 3.16], registry in marker coordinates                  | Preserved; Proposition 7.8                                                                                                                                                    |
| [Sch26d, Prop. 3.19] with $H_{\text{ind}}$                            | Preserved; surjectivity clause of Proposition 4.4(i)                                                                                                                          |
| [Sch26d, Prop. 4.4], what depends on the frame                        | Preserved and made load-bearing: it is why amino/keto is admissible and the base identity is not (Corollary 7.9)                                                              |
| [Sch26d, Prop. 6.5], $\Phi$ is a torsor                               | Preserved; Proposition 4.4(ii)                                                                                                                                                |
| [Sch26d, Prop. 6.7], decomposition of the frame action                | Preserved and made load-bearing: every occurrence of twenty depends on it, through Proposition 4.4(ii) and Proposition 7.2. Restated in the open rather than cited in passing |
| [Sch26d, Rem. 6.6], the two half-pointed torsors                      | Preserved; Remark 4.14                                                                                                                                                        |
| [Sch26d, §2.2], the localized cipher involution $\Delta_{\mathbf{k}}$ | Noted, and <i>kept apart</i> : the production operator of this paper carries the same letter unsubscripted. The involution is used in no statement here; see Remark 2.22      |
| [Sch26d, Defs. 6.1, 9.1], the initiator                               | Preserved; excluded from the cluster by Definition 4.1 under $H_{\text{prod}}(0')$ , with the cost stated at Remark 4.2                                                       |
| [Sch26d, Prop. A.7], four words from one stack                        | Preserved; Remark 4.5                                                                                                                                                         |
| [Sch26d, Prop. 3.8], the marker is information-bearing                | Preserved as [T]; Target 8.7                                                                                                                                                  |
| [Sch26d, Rem. 3.7], the marker is not a chirality                     | Not applicable; no statement here concerns configuration                                                                                                                      |
| [Sch26d, Thm. 4.17], sequence dependence                              | Not applicable; A7 unused (Proposition 2.5)                                                                                                                                   |
| A1, A2, A6, A8, A12                                                   | Preserved; no axiom added (Remark 2.2)                                                                                                                                        |
| A8, the cipher pairing $\delta_{\text{dc}}$                           | Preserved; used to relate the two substitution configurations at Remark 7.7, and in no proof                                                                                  |
| [Sch26c, Thm. 61], [Sch26c, Prop. 39]                                 | Preserved, not re-established                                                                                                                                                 |
| [Sch26c, Thm. 47], emergent chirality                                 | Not touched                                                                                                                                                                   |
| [Sch26b, §3.1], the governing restraint                               | Preserved; Section 3 is written in its form                                                                                                                                   |
| [Sch26b, §9.2], the identification premise                            | Preserved and <i>declined</i> : Proposition 3.1(vii) states that the theorem bounds the products of a cluster and asserts nothing about the biological alphabet               |

### B.4 Chemical scope

**Remark B.1** (What is placed before the domain review, and what is not). [E] This paper is not offered for chemical review and no statement of Sections 4 to 6 is conditional on a chemical development. Remark 7.17 fixes the scope: the reactants are identified as the complex intermediary products of the synthesis of US 12,195,423; the process is outlined at Remark 7.5 to the level the mathematical claims require; and the synthesis steps, intermediates, substitution patterns and liberating step are forthcoming as a separate chemical treatise.

Three items are nevertheless placed before the domain pass because they are chemical statements that a proof uses.  $H_{\text{clu}}$ , that the reacting unit is six alternating units — on which the triplet and the ceiling both depend, and which must be established without reference to codons or to the alphabet size (Remark 2.9).  $H_{\text{int}}$ , that each interface produces one nucleobase. And  $H_{\text{prod}}(0')$ , that the amino-acid pathway does not involve the initiator — which decides between a ceiling of twenty and one of thirty-two and is the paper's central empirical commitment (Remark 3.3, Target 8.2).

The species and locant assignment of [Sch26d, Rem. B.3] is inherited unchanged, and the alternate configuration at C5 or C11 is recorded at Remark 7.7 and used nowhere.

## B.5 Self-identified defect ledger

Sixteen items, in the crit/rev/ver form of [Sch26b, Def. 3]. Items 1–14 were raised at the pre-drafting pass against the concept outline and are recorded as such, not retrospectively presented as design decisions; items 15 and 16 arose during drafting. All sixteen are disposed fully accepted.

| #  | crit                                                                                                                                           | rev / ver                                                                                                                                                         |
|----|------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | “Exactly twenty” presented as the theorem                                                                                                      | Split into ceiling and saturation, both in the statement. Explicit: Theorem 5.4                                                                                   |
| 2  | $N = 3$ justified from the codon being a triplet                                                                                               | Superseded by item 12                                                                                                                                             |
| 3  | The frame group proved for the letter, not the product                                                                                         | Superseded by item 11                                                                                                                                             |
| 4  | “ $\Delta$ produces the twenty and the stop sites”                                                                                             | Parity forbids it; phrasing removed. Explicit: Corollary 5.12                                                                                                     |
| 5  | The partition result presented as intrinsic                                                                                                    | Disclosed as $L$ -dependent and <i>a posteriori</i> . Explicit: Proposition 3.1(iii)                                                                              |
| 6  | Entropy conflated with the count                                                                                                               | Remark cut entirely, item 14                                                                                                                                      |
| 7  | One worked case offered as a method                                                                                                            | Section 7 labelled illustration. Explicit: Remark 7.15                                                                                                            |
| 8  | Naming: $\Delta$ against $\varphi$ , $H_{\text{prod}}$ against $H_{\text{dual}}$                                                               | Renamed; both substitutions recorded. Explicit: Remark 2.22                                                                                                       |
| 9  | The intermediary read as a nucleobase                                                                                                          | Remark added and the capacity shown independent of $L$ . Explicit: Remark 4.11, Proposition 4.12                                                                  |
| 10 | Frame-invariance registered as a hypothesis; it is a consequence, and the real assumption was missing                                          | Four-clause register adopted; $H_{\text{prod}}(0)$ introduced. Explicit: Corollary 4.6, Remark 4.7                                                                |
| 11 | The initiator breaks the reversal symmetry, giving thirty-two                                                                                  | Resolved by $H_{\text{prod}}(0')$ ; the sensitivity table rewritten as a conditional. Explicit: Remark 4.2, Remark 5.15                                           |
| 12 | $N = 3$ still justified from the code                                                                                                          | $H_{\text{clu}}$ registered as a chemical statement prior to codons; the triplet becomes a consequence. Explicit: Remark 2.9, Corollary 5.6                       |
| 13 | A load-bearing claim cited to a table caption                                                                                                  | Proved as a proposition. Explicit: Proposition 4.4                                                                                                                |
| 14 | The entropy remark needs its measure named                                                                                                     | Cut; the thesis is a count                                                                                                                                        |
| 15 | Raised at drafting: the two worked cases take their candidate pairs from [Sch26f, Prop. 5.8], which is computed from the standard genetic code | The distinction is exposition against proof: the choice of what to display has that provenance, what Section 7 asserts does not. Disclosed. Explicit: Remark 6.10 |
| 16 | Raised at drafting: an [E] claim used inside the proof of the triplet corollary                                                                | The chemistry is registered as $H_{\text{clu}}$ and $H_{\text{int}}$ , with the molecular reading confined to a remark. Explicit: Section 2.4, Remark 5.7         |

## B.6 Verification record

Seven checks. Each is an exhaustive enumeration or a closed computation over a finite set, rebuilt from the definitions by a supplementary script that imports nothing from the tables of this paper.

**V1** Orbits on the sixty-four marker assignments under every subgroup of Fr: 64, 32, 36, 36, 20 (Proposition 5.14).

**V2** The class count under all twenty-four letter assignments: twenty in every case (Proposition 4.12).

**V3**  $|X|$  at  $N = 1, \dots, 8$  against  $4^N$ , and the ratio  $\frac{1}{4} + 2^{-(N+1)} < 1$  throughout; the closed formula checked against direct enumeration for  $N \leq 4$  (Theorem 5.1, Proposition 5.2).

**V4** Every class has size two or four, eight and twelve respectively; the standard codon table has exactly two singleton fibres, Met and Trp (Proposition 5.9, Corollary 5.10). *This is the only check that reads the codon table.*

- V5** The twenty classes with  $\overline{D}$ ,  $|D|$ , size, stabiliser type and central entry; the weight fibres 1, 3, 6, 6, 3, 1; size two if and only if  $\overline{D}$  is palindromic; the central entry splitting the eight four and four (Proposition 6.4, Table 1).
- V6** Reversal fixes a vessel if and only if  $N$  is odd, and at  $N = 3$  the fixed vessel is the middle one (Proposition 7.2).
- V7** The four classes of Section 7 with their weights, sizes and vessel registries (Examples 7.11 and 7.13).

## B.7 Meta-review

**Remark B.2** (Disposition). Sixteen items, all disposed fully accepted; none partially resolved, unresolved or no longer applicable. All sixteen carry *ver explicit*, there being no withdrawal whose verification would be an absence. Seven checks run, seven passing.

**Remark B.3** (Three observations). First, the two items with the largest consequence, 10 and 11, were raised before drafting and are not defects of proof. Item 10 found a hypothesis that was not one and a hypothesis that was missing; item 11 found that the classification's own symmetry is broken by an object the predecessor places at the end of the stack. Both changed what the paper assumes rather than whether a step follows, which is the same asymmetry [Sch26b, Rem. 1] reports and [Sch26f, Rem. B.11] records on a second manuscript.

Second, the pass improved the result rather than narrowing it, which is unusual in this corpus. Item 13 turned a caption into a proposition and item 10 turned a hypothesis into a corollary; item 12 turned an assumption into Corollary 5.6, so that the paper answers two questions where the outline answered one. Three of the sixteen made the manuscript claim *more*, not less, and each did so by making an assumption explicit.

Third, the paper's dependence is now concentrated rather than diffuse. Every occurrence of the number twenty passes through [Sch26d, Prop. 6.7] and  $H_{\text{prod}}(0')$ , and Remark 3.3 states the alternative value it takes if either fails. That is a better position than a result resting on many small assumptions, and it is why Target 8.2 is named load-bearing.

**Remark B.4** (What the review does not certify). Targeted verification re-runs the mathematical stage and does not employ an independent reviewing system; by [Sch26b, §9.4] independence of pass is not independence of reviewing system, and every entry marked *ver explicit* inherits that limitation. The seven numerical verifications are not subject to it, being reproducible from the definitions.

Seven Targets are open, three load-bearing (Remark 8.8). Four hypotheses —  $H_{\text{prod}}(0)$ ,  $(0')$ ,  $(1)$  and  $(2)$  — are registered and undischarged, and the capacity theorem's second clause stands on the last two.  $H_{\text{clu}}$  and  $H_{\text{int}}$  are chemical and are placed before the domain pass. No synthesis is reported, and the chemical development on which Section 7's two cases rest is forthcoming and not part of this paper.

## C Established Literature

As in the literature appendices of [Sch26c, Sch26b, Sch26d, Sch26e, Sch26f], none of the subjects below is claimed to be novel in isolation. The purpose is to identify the established machinery each part of this paper draws upon and to keep it visibly distinct from the organisation proposed here.

The list is shorter than the companion paper's, and the reason is the paper's principal claim rather than an economy: no matching theory, no permanent, and no assignment problem occurs here, because no coarsening is selected and no adjacency is built. What remains is orbit counting, torsors, and two external tables used once each.

### C.1 Group actions and orbit counting

The orbit-counting lemma underlying Theorem 5.1 and Proposition 5.14 is classical and is stated in [Bur11, §119]; a modern treatment, with the orbit–stabiliser theorem used in Proposition 6.4, is [Rot95, Ch. 3]. The attribution to Burnside is conventional and inaccurate, the result being due to Cauchy and Frobenius; the name is retained because the corpus will be searched under it.

Two uses are worth separating. The count of Theorem 5.1 is the lemma applied to a Klein four-group acting on  $\{0, 1\}^{2N}$ , and is imported from [Sch26e]. The count of Proposition 6.1 — binary sequences of length five modulo reversal,  $(2^5 + 2^3)/2$  — is

the same lemma for a two-element group, and is the form in which the capacity bound is a statement about the arguments of the production operator rather than about strings.

## C.2 Torsors

The notion of a torsor under a finite abelian group is standard; see [Bou70, Ch. I, §5] and, for the finite case, [Ser79, Ch. I]. It enters Proposition 4.4 through [Sch26d, Prop. 6.5], the frame-setters forming a torsor under  $\mathbb{Z}/2 \times \mathbb{Z}/2$ , and the property used is that a torsor has no canonical base point, so that a statement about its elements must be invariant under the acting group or must name the choice that breaks the symmetry. Remark 4.14 applies that discipline to the three four-element objects of this paper.

## C.3 Elementary enumeration

The counts of Proposition 5.2 and Table 2 are elementary; the standard reference for the surrounding combinatorics is [Sta12, Ch. 1]. That the classes are a quotient of the words, and hence that  $|X| < 4^N$  without further argument, is immediate from Proposition 4.4 and is recorded only because Remark 5.3 turns on it.

## C.4 The two external tables

The standard genetic code is cited to [Cri68] and is used in exactly one place: Corollary 5.10 and Corollary 5.12, where two of its properties — that methionine and tryptophan occupy one codon each, and that UAG and GCU are assigned differently — exhibit what an assembly-level operator cannot do. The properties were recomputed from the codon table at the domain pass rather than read from the citation, in the manner of [Sch26e, Rem. 3.16], and the recomputation is check V4 of Appendix B. No other statement of this paper reads a codon table, which is the content of Remark 2.4.

The residue volumes and hydropathies of Remark 7.14 are [Zam72] and [KD82]. They are used to say that two candidate products are close, and enter no proof. The IUPAC one-letter ambiguity classes, under which the two registry classes of a vessel are the amino and keto bases, are [NCIUB85]; that identification belongs to the companion paper's letter assignment and by Proposition 4.12 no statement of Sections 4 to 6 depends on it.

## C.5 What is drawn from the Chrysene Formalism itself

Axioms A1–A12, the status vocabulary, the non-vacuity theorem and the acyclicity of the axiom ladder are [Sch26c]. The marker and its non-collapse, the vessel and its two binaries, the frame and its decomposition, the initiator, and the witness stack are [Sch26d]; the load-bearing imports are its Propositions 3.4, 3.19, 6.5 and 6.7, and the last of these carries every occurrence of the number twenty. The classification into frame classes, the orbit count, the stabiliser refinement and the realization proposition are [Sch26e]. The cluster naming, the letter assignment, the partition criterion and the coarsening against which  $\Delta$  is distinguished are [Sch26f]. The review architecture and the non-implication conventions are [Sch26b], and the survival programme, to which nothing here bears, is [Sch26a]. The molecular architecture is claimed in [Sch25a] and [Sch25b].

None of these is external literature in the sense of this appendix; they are listed so that the boundary between established machinery and the formalism's own claims is visible in one place.

## References

- [Bou70] Nicolas Bourbaki. *Algebra I: Chapters 1–3*. Hermann / Addison-Wesley, 1970.
- [Bur11] William Burnside. *Theory of Groups of Finite Order*. Cambridge University Press, 2nd edition, 1911.
- [Cri68] Francis H. C. Crick. The origin of the genetic code. *Journal of Molecular Biology*, 38(3):367–379, 1968.
- [KD82] Jack Kyte and Russell F. Doolittle. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1):105–132, 1982.
- [NCIUB85] Nomenclature Committee of the International Union of Biochemistry. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *European Journal of Biochemistry*, 150(1):1–5, 1985.

- [Rot95] Joseph J. Rotman. *An Introduction to the Theory of Groups*, volume 148 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 4th edition, 1995.
- [Sch25a] Charles D. Schaper. Encoded design integrated synthesis of nucleic acids, and phospholipids, and related pharmaceutical products, 2025. US Patent 12,195,423.
- [Sch25b] Charles D. Schaper. Encoded design integrated synthesis of nucleic acids, and phospholipids, and related pharmaceutical products, 2025. US Patent 12,473,247.
- [Sch26a] Charles D. Schaper. Stable Transfer Geometry of the Substituted Chrysene Network:  $H^\infty$  Transformation Metrics, Dissipative Well-Posedness, and Survival Under Persistent Disturbance. *Annals of the Chrysene Formalism*, 1(2):603–645, 2026.
- [Sch26b] Charles D. Schaper. The Chrysene Formalism as Candidate Mathematics and Candidate Biogenesis: A Quantitative and Categorical Survey of Volume 1, Number 1. *Annals of the Chrysene Formalism*, 1(2):476–488, 2026.
- [Sch26c] Charles D. Schaper. The Chrysene Formalism Revisited: A Twelve-Axiom Foundation for Molecular Characterization, Tensorial Noncommutative Geometry, and Operator Response. *Annals of the Chrysene Formalism*, 1(2): 489–560, 2026.
- [Sch26d] Charles D. Schaper. The Encoded Stack: A Frame-Indexed Base-Four Structure on the  $\mathbb{Z}^3$ -Fibrated Chrysene Network, its Sequence Algebra, and the Set-Up for Survival. *Annals of the Chrysene Formalism*, 1(2):646–675, 2026.
- [Sch26e] Charles D. Schaper. The Twenty Registry Classes of a Chrysene Stack: Frame Orbits, Descent of a Letter Transformation, and a Variational Criterion. *Annals of the Chrysene Formalism*, 1(2):676–704, 2026.
- [Sch26f] Charles D. Schaper. Coarsenings of the Genetic Code by the Twenty Registry Classes of a Chrysene Stack: A Partition Criterion, a Filter Selection, and the Double-Sided Encoding Hypothesis. *Annals of the Chrysene Formalism*, 1(2): 730–767, 2026.
- [Ser79] Jean-Pierre Serre. *Local Fields*, volume 67 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979.
- [Sta12] Richard P. Stanley. *Enumerative Combinatorics, Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2nd edition, 2012.
- [Zam72] Alexander A. Zamyatnin. Protein volume in solution. *Progress in Biophysics and Molecular Biology*, 24:107–123, 1972.