The symbol alphabet of one-loop Feynman integrals

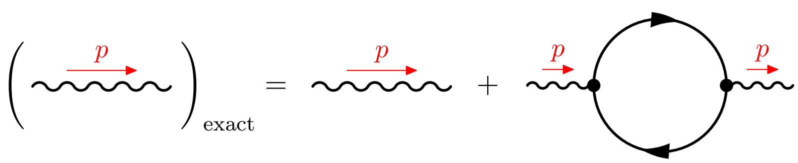
Alexander Farren† Eliza Somerville Mikey Whelan
supervised by Professor Ruth Britto
(2 August 2024)

Abstract

Improving our ability to compute Feynman integrals is critical as we continue in this precision era of particle physics. One-loop diagrams are promisingly amenable to be studied using the symbol of polylogarithms. Their alphabet is well understood for generic kinematics. Using the diagrammatic coaction of 1704.07931, we construct symbol recursions for various non-generic diagrams in uniform weight dimensional regularisation and hence identify their symbol alphabet and dictionaries. We extend the well-established connection between symbol letters and Gram determinants by clarifying the rules for taking limits of dictionaries towards non-generic diagrams.

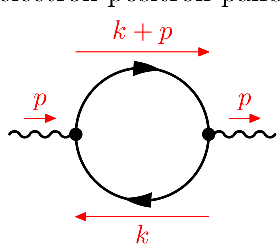
1 Introduction

Feynman diagrams were generously handed down to us humans by Richard Feynman in the 1940's [1] and have since helped theoretical physicists in calculating the probabilities involved in subatomic processes. For example, in quantum electrodynamics the exact contribution (to such probabilities) coming from a photon carrying four-momentum p between two points in spacetime is represented by the diagrams [2]



$$+\,\mathcal{O}(2\text{ loops}). \tag{1.1}$$

Here the squiggly lines are photons and the arrowed lines can be electrons or positrons. If the photon energy satisfies $E_{\mathbf{p}} \gtrsim 2m_e$, where m_e is the electron mass, an electron-positron pair may be created and subsequently annihilated back to a photon. This can happen arbitrarily many times. We say that the virtual electron carries momentum $k+p$ forward in time, while the positron carries momentum k backwards in time. But there is no constraint on the undetermined loop momentum k as long as the pair's four-momenta add to p to satisfy momentum conservation. We then must integrate over all values of k to account for all possible electron-positron pairs which could be created. This is a *Feynman integral*:



$$\sim \int d^4k\, \frac{\not{k}+m_e}{k^2-m_e^2}\,\frac{\not{k}+\not{p}+m_e}{(k+p)^2-m_e^2}. \tag{1.2}$$

Internal lines which join vertices in the diagram, and form the loop(s), are called *propagators*. The fractions in the integrand of (1.2) correspond to the respective propagators and follow from Feynman's rules for translating diagrams into probability amplitudes. For more information, consult for example [2]. We call diagrams with two propagators 'bubbles'.

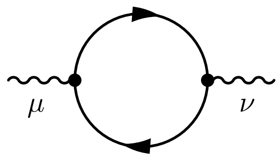
In general, Feynman integrals appear whenever there is a loop around which an undetermined momentum is flowing. For each loop in a diagram, there is an associated loop momentum and hence an integral.

†farrenal@tcd.ie

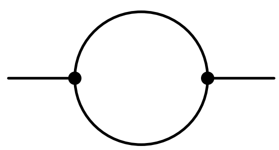
In (1.2) the two propagators correspond to an electron and a positron, which both have mass m_e. But massless particles such as photons and gluons can also appear as virtual propagators in physical processes. Different propagators in the same integral may also have different masses.

Furthermore, the physical theory we just discussed involves electrons, positrons and photons. All these particles have non-zero *spin* which means we need to specify (or in some cases sum over) the particles' spin states in the diagram translation. Fortunately, there is a non-trivial result which says such integrals involving particles with spin may be expressed in terms of integrals involving only particles without spin, which are called *scalar* particles. These simpler integrals are naturally called *scalar integrals*. Even more fortunately, any scalar integral can itself be expressed in terms of scalar integrals with at most four propagators (or equivalently four denominators in the integrand) [3–9].

To illustrate the difference in complexity between spin theories and scalar theories, let us compare the full expression for (1.2) with photon Lorentz indices μ and ν at the vertices given by (1.3) against the same diagram in a theory with one type of scalar particle (plain line) having mass m_e given by (1.4):

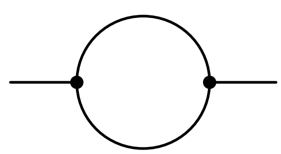


$$= (-\mathrm{i}e)^2 \int \frac{d^4k}{(2\pi)^4} \operatorname{tr}\left[\gamma^\mu \frac{\not{k} + m_e}{k^2 - m_e^2} \gamma^\nu \frac{\not{k} + \not{p} + m_e}{(k+p)^2 - m_e^2}\right], \tag{1.3}$$



$$= (-\mathrm{i}e)^2 \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 - m_e^2} \frac{1}{(k+p)^2 - m_e^2}. \tag{1.4}$$

(The constant e is the electric charge and just tells us how strongly the internal and external particles interact.) Even without a background in quantum field theory and its Dirac matrices γ^μ, the reader should be able to appreciate the benefit of working with scalar integrals. While (1.4) *is* simpler to compute, in 4 spacetime dimensions both (1.3) and (1.4) yield an infinity. This is obvious by looking at the integrand which has poles corresponding to the mass of each propagator. So, instead of working with infinities, we can opt for the unusual approach of treating the number of dimensions as a variable by setting $D = 4 - 2\epsilon$ where ϵ is a formal complex parameter. This is known as *dimensional regularisation*. With this prescription, it turns out (1.4) evaluates to a hypergeometric function, ${}_2F_1$, and has leading order contribution proportional to ϵ^{-1}. This amounts to isolating the divergence. But there are terms with higher powers of ϵ, such that in the most general form, the integral for the scalar bubble is a Laurent series



$$= \sum_{i=-1}^{\infty} c_i \epsilon^i. \tag{1.5}$$

Precise understanding of what coefficients c_i appear in (1.5) for any integral performed in dimensional regularisation would let us skip the hard part: doing the integral. Before we depart from physics entirely, let us motivate this understanding from a practical perspective.

In modern particle physics, collider experiments probe finer and finer physical processes which lead to more complicated diagrams, often with several loops [10]. Elsewhere in research, works exploring new and deeper theories which are currently not observed (such as supersymmetric theories) also demand faster and more practical tools for computing Feynman integrals [11]. Finally, one could of course argue for the study of Feynman integrals as mathematical objects, interesting in their own right. *Scattering amplitudes* is the broad area dedicated to understanding the structure behind these integrals, thereby facilitating the calculation of the outcomes of scattering processes. A prime – and not-so-distant – success story of this field is [12], in which a 17-page expression for a scattering amplitude was reduced to a few lines. We will soon see exactly what mathematical structure allowed for this epic simplification.

We hope the reader now believes Feynman integrals are of importance. But how do we actually study them? This work will focus on one-loop integrals I_n^D with $n \leq 4$ propagators, which we call n-point integrals in D dimensions. A nice feature of one-loop integrals is that their Laurent series coefficients are always linear combinations of *multiple polylogarithms* (MPLs) [13, 14]. These functions are a generalisation of the logarithm and have sum representation

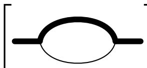
$$\mathrm{Li}_{n_1 \ldots n_k}(x_1, \ldots, x_k) = \sum_{m_1 > \ldots > m_k}^{\infty} \frac{x_1^{m_1}}{m_1^{n_1}} \cdots \frac{x_k^{m_k}}{m_k^{n_k}} \tag{1.6}$$

with a measure of transcendentality called the *weight* given by $n_1 + \ldots n_k$. It is a this point that we part ways with concrete physics. Earlier we introduced dimensional regularisation with $D = 4 - 2\epsilon$ since we live in four spacetime dimensions. However, it turns out that if we instead use round up to the nearest even number so $D_n = 2\lceil n/2 \rceil - 2\epsilon$, and assign ϵ a weight of -1, then every term in the Laurent series of $I_n^{D_n}$ has weight $\lceil n/2 \rceil$. This means the coefficients c_i are linear combinations of multiple polylogarithms *of the same weight* given by $\lceil n/2 \rceil + i$. We call this approach 'uniform transcendentality/weight' dimensional regularisation. The benefit of this prescription is in relation to a map $\mathcal{S}$ associated with MPLs called the *symbol* which only sees the highest weight in its argument. Schematically, the symbol acts as

$$\mathcal{S}(\text{MPL of weight } n) = \text{Linear combination of tensor products of } n \text{ MPLs of weight one.} \tag{1.7}$$

An MPL of weight one is just a logarithm. This immediately tells us a particular usefulness of the symbol: it allows for manipulation of MPLs of high weight through only using log rules. (This usefulness is what simplified the 17-pager.)

Since $\mathcal{S}$ outputs strings of tensor products with as many entries as the weight of the argument MPL, for us this means that every coefficient in the Laurent series of a Feynman integral gets mapped to what we call *words* of length $\lceil n/2 \rceil$ and having as many *letters*. For example, if we consider a scalar bubble with one massless propagator (shown thin), some words of length three are



$$\mathcal{S}\Big[\;\; \cdots \;\;\Big] \sim \big(\log m^2 \otimes \log m^2 \otimes \log p^2 - 2 \log m^2 \otimes \log\big(m^2 - p^2\big) \otimes \log p^2\big)\epsilon^2 + \ldots \,. \tag{1.8}$$

The letters of Feynman integrals depend on the external and internal kinematical variables, and inform on the analytic structure of the original integral. For a given diagram, we would like to know all of the possible words which appear on the right-hand side of (1.8) and then 'undo' the symbol as to obtain the integral[1]. Even assuming we *did* know all of the words, there is an issue because the symbol has a non-trivial kernel. However, it is often possible to reconstruct the original MPL algorithmically [12, 13, 17] or, as shown very recently, using machine learning [18]. This has a strong implication: if one can predict the symbol *alphabet* and *dictionary* of a Feynman integral, then reading off all of the eventual words at each order in ϵ should equate to doing the integral.

Significant progress was made in the last decade by exploring features which are connected to the symbol alphabet, such as canonical differential equations of Feynman integrals [19, 20], the cuts of said integrals [21, 22] or even the (tropical) geometry of polytopes associated to kinematic variables [23–25]. In particular, Gram and modified Cayley determinants (related to the volume of the latter polytopes) seem to play a key role in fixing the alphabet, or at the very least in constraining letters [4, 26, 27].

[1]The attentive reader may be worried about the number of dimensions being off, but there exist relations to shift things back to our spacetime [8, 15, 16].

In this work we will use these rich connections to answer interesting questions. To see a concrete example, let us return to (1.8). For brevity, only two words were shown in that example. The full list of three-letter words is the following:

1st letter	2nd letter	3rd letter
m^2	m^2	m^2
m^2	m^2	p^2
m^2	m^2	$(m^2 - p^2)$
m^2	p^2	p^2
m^2	p^2	$(m^2 - p^2)$
m^2	$(m^2 - p^2)$	p^2
m^2	$(m^2 - p^2)$	$(m^2 - p^2)$
$(m^2 - p^2)$	p^2	p^2
$(m^2 - p^2)$	p^2	$(m^2 - p^2)$
$(m^2 - p^2)$	$(m^2 - p^2)$	p^2
$(m^2 - p^2)$	$(m^2 - p^2)$	$(m^2 - p^2)$

Why is p^2 never a first letter? What about $(m^2 + p^2)$? What other patterns are there? For a systematic treatment of one-loop Feynman integrals, it would be best to instead ask:

– What letters can appear? In what entry?
– Given this alphabet, what words are possible?

We outline a way to answer such questions for this and other non-generic Feynman integrals using a recursive formula for the symbol [19]. In answering these questions, we encounter previously known constraints on symbol entries such as the (generalised) Steinmann relations [28] and the very recent genealogical constraints [29]. This recursion analysis will rely on the Gram and Cayley determinants previously mentioned, since these appear to be the ingredients which generate longer words through the recursion. These determinants enter through cut integrals, which can be viewed as taking the discontinuity of a Feynman integral across some branch cut.

Unlike generic cases where all propagator masses and external momenta are left unspecified (e.g. $p_1^2, p_2^2, \ldots, m_1^2, m_2^2, \ldots$), as soon as a propagator is made massless, or two energy scales are made equal, such determinants start to vanish. This is problematic because the symbol letters sometimes depend on the reciprocal of determinants, which could lead to indeterminacy in the recursion. We make first steps in reconciling this with the symbol calculated from the non-generic integral itself, rather than as a particular limit of a more generic kinematic configuration's symbol.

After having motivated this study and summarised our results, we will now introduce our convention for Feynman integrals, cut integrals and multiple polylogarithms which is that of [19, 21]. Then, the coproduct and symbol map of Feynman integrals will be defined such that we may start to understand alphabets.

2 Mathematical background

2.1 Feynman integrals

In the notation of [19], the scalar one-loop n-point Feynman integrals are defined as

$$I_n^D(\{p_i \cdot p_j\}; \{m_i^2\}; \epsilon) = e^{\gamma_E \epsilon} \int \frac{d^D k}{\mathrm{i}\pi^{D/2}} \prod_{j=1}^{n} \frac{1}{(k - q_j)^2 - m_j^2 + \mathrm{i}0}, \tag{2.1}$$

where $\gamma_E = \Gamma'(1)$ is the Euler-Mascheroni constant, and where we work in dimensional regularisation in $D = d - 2\epsilon$ dimensions, where d is an even positive integer and ϵ is a formal variable called the dimensional regulator. We denote the loop momentum by k, while the external momenta are labelled by p_i and satisfy the law of conservation of momentum, $\sum_{i=1}^{n} p_i = 0$. We define q_j to be a linear combination of the external momenta such that the momentum carried by the propagator labelled by j is $k - q_j$. Thus q_j can be obtained by imposing momentum conservation at each vertex of the diagram corresponding to the integral I_n^D:

$$q_j = \sum_{i=1}^{n} c_{ji} p_i, \quad c_{ji} \in \{-1, 0, 1\}. \tag{2.2}$$

For simplicity, we define the loop momentum k to be the momentum carried by the propagator labelled by 1, so that $q_1 = 0$. In [19], a convenient basis for all one-loop integrals was chosen to be

$$\widetilde{J}_n\big(\{p_i \cdot p_j\}; \{m_i^2\}; \epsilon\big) = I_n^{D_n}\big(\{p_i \cdot p_j\}; \{m_i^2\}; \epsilon\big), \tag{2.3}$$

where $D_n = 2\lceil n/2 \rceil - 2\epsilon$, or more explicitly

$$D_n = \begin{cases} n - 2\epsilon & \text{for } n \text{ even,} \\ n + 1 - 2\epsilon & \text{for } n \text{ odd.} \end{cases} \tag{2.4}$$

We note that the existence of dimensional shift identities [8, 15, 16] means that instead of choosing master integrals for a fixed dimension D, we may choose different basis integrals to be evaluated in different dimensions. The integrals $\widetilde{J}_n$ form a particularly convenient basis because they expected to be expressible in terms of multiple polylogarithms of uniform weight $\lceil n/2 \rceil + i$, up to an overall algebraic factor, at each order ϵ^i. This indicates that all one-loop Feynman integrals can be expressed in terms of multiple polylogarithms [19].

One parameterisation of (multiloop) scalar Feynman integrals is the *Feynman parameter integral*

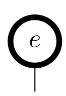
$$I = \frac{\Gamma\big(\nu - \frac{LD}{2}\big)}{\prod_{i=1}^{n} \Gamma(\nu_i)} \int_{\alpha_i \geq 0} d^n\alpha \ \delta\left(1 - \sum_{i}^{n} \alpha_i\right) \left(\prod_{i=1}^{n} \alpha_i^{\nu_i - 1}\right) \frac{\mathcal{U}^{\nu-(L+1)D/2}}{\mathcal{F}^{\nu - LD/2}} \tag{2.5}$$

where $\mathcal{U}$ and $\mathcal{F}$ are the two graph polynomials, the latter containing the kinematic data. For the one loop case, the delta function sets $\mathcal{U} = 1$ and reduces the integration region to a simplex over the parameters. For our purposes, the powers ν_i of the propagators are all 1, as is the loop number L.

Examples of Feynman Integrals

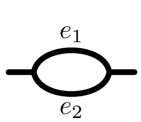
To better illustrate the form of the basis integrals, we now discuss some simple examples of one-loop Feynman diagrams and their corresponding integrals. The diagrams shown here follow the convention that massive propagators are represented by bold lines, while massless propagators are represented by normal lines.

The simplest example of a Feynman integral is the tadpole integral. This is a one-point integral, so taking $n = 1$ in (2.4) we see that using our conventions this diagram should be evaulated in $D = 2 - 2\epsilon$ dimensions. The integral is given by



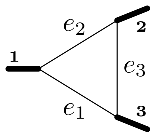
$$= \widetilde{J}_1(m^2) = e^{\gamma_E \epsilon} \int \frac{d^D k}{\mathrm{i}\pi^{D/2}} \frac{1}{k^2 - m^2 + \mathrm{i}0}. \tag{2.6}$$

The next simplest case is the bubble integral. This is a two-point integral, so as in the case of the tadpole we will evaluate it in $D = 2 - 2\epsilon$ dimensions. In the most general case where momentum p^2 flows through, and the two propagators have masses m_1^2 and m_2^2, the integral is given by



$$= \widetilde{J}_2(p^2; m_1^2, m_2^2) = e^{\gamma_E \epsilon} \int \frac{d^D k}{\mathrm{i}\pi^{D/2}} \frac{1}{(k^2 - m_1^2 + \mathrm{i}0)((k-p)^2 - m_2^2 + \mathrm{i}0)}. \tag{2.7}$$

A third important example is that of the triangle diagram with three external scales and three massless propagators. This is a three-point integral, so we will evaluate it in $D = 4 - 2\epsilon$ dimensions. The integral is given by

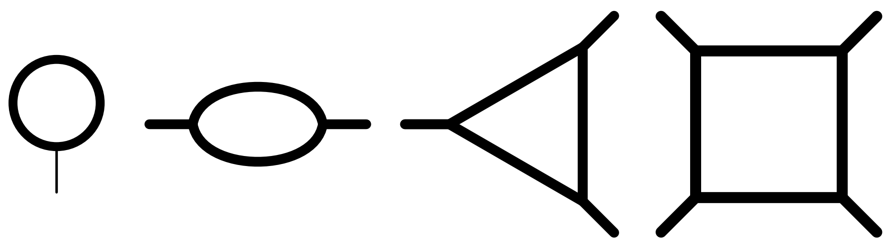


$$= \widetilde{J}_3(p_1^2, p_2^2, p_3^3) = e^{\gamma_E \epsilon} \int \frac{d^D k}{\mathrm{i}\pi^{D/2}} \frac{1}{(k^2 + \mathrm{i}0)((k-p_1)^2 + \mathrm{i}0)((k-p_1-p_2)^2 + \mathrm{i}0)}. \tag{2.8}$$

In this diagram, the number labelling each external edge indicates the index of the external momentum flowing through that edge.

Basis integrals

One can actually express any one-loop Feynman integral in terms of $\tilde{J}_n\big(\{p_i \cdot p_j\}; \{m_i^2\}; \epsilon\big)$ for $n \leq 4$ [3–9].



(2.9)

Tadpole Bubble Triangle Box (2.10)

Integrals with non-generic kinematics (e.g. $m_1^2 = 0, m_1^2 = m_2^2 \ldots$) are also part of this basis. Our goal will be to predict the symbol for such non-generic cases. The alphabet of generic integrals was derived in [19] but parts of the analysis fail when, for example, massless limits are taken. To understand this derivation and how to modify it to accommodate non-generic kinematics, we will need the notion of cut integrals.

2.2 Cut integrals

Cut Feynman integrals are closely related to the analytic structure of Feynman integrals. The concept of cut integrals originates from the cutting rules of Cutkosky [23]. The appearance of cut integrals in many different areas of study pertaining to the analytic structure of Feynman integrals stems from the fact that Feynman integrals are multi-valued functions, and cut integrals are related to the discontinuities of the original integral across its branch cuts [21]. In recent years, cut integrals have played a role in the study of integration-by-parts identities [30] and differential equations [31] satisfied by Feynman integrals.

A cut integral $\mathcal{C}_C \widetilde{J}_n$ is obtained by starting with a normal one-loop Feynman integral $\widetilde{J}_n$ and designating a subset C of propagators as *cut*. We call the remaining propagators *uncut*. Traditionally, the cut integral is computed by replacing the cut propagators by Dirac delta functions according to

$$\frac{1}{(k-q_j)^2 - m_j^2 + \mathrm{i}0} \to -2\pi\mathrm{i}\delta((k-q_j)^2 - m_j^2), \tag{2.11}$$

and then evaluating the integral under these constraints; this essentially corresponds to forcing the cut propagators on mass-shell.

However, the prescription (2.11) is not completely sufficient when studying the analytic structure of Feynman integrals, and a more precise definition of cuts is necessary. Such a definition was given for one-loop cut integrals in [21], where they were defined as residues integrated over a well-defined contour in dimensional regularisation. To fully state this definition, we must introduce two determinants which can be applied to a subset of propagators $C \subseteq [n] = \{1, \ldots, n\}$. The first is the Gram determinant,

$$\mathrm{Gram}_C = \det((q_i - q_*) \cdot (q_j - q_*))_{i,j \in C \setminus *}, \tag{2.12}$$

where $*$ denotes any particular element of C. The second is the modified Cayley determinant,

$$Y_C = \det\left(\frac{1}{2}(-(q_i - q_j)^2 + m_i^2 + m_j^2)\right)_{i,j\in C}. \tag{2.13}$$

These can be used to classify the singularities of Feynman integrals into two types. A singularity of the first type corresponds to a kinematic configuration where Gram_C vanishes for a subset C of propagators. To such a singularity we associate a cut integral $\mathcal{C}_C\widetilde{J}_n$, where the integration contour is deformed so as to encircle the poles of the propagators in C. When this integral is evaluated in terms of residues, one obtains

$$\mathcal{C}_C\widetilde{J}_n = \frac{(2\pi\mathrm{i})^{\lfloor n_C/2\rfloor} e^{\gamma_E\epsilon}}{(2\mathrm{i})^{n_C}\sqrt{Y_C}}\left(-\frac{Y_C}{\mathrm{Gram}_C}\right)^{(D_n-n_C)/2}\int\frac{d\Omega_{D_n-n_C}}{\mathrm{i}\pi^{D_n/2}}\left[\prod_{j\neq C}\frac{1}{(k-q_j)^2-m_j^2}\right]_C \mod \mathrm{i}\pi, \tag{2.14}$$

where $n_C = |C|$ is the number of cut propagators, $[\cdot]_C$ indicates that the function inside the sqaure brackets is evaluated on the zero locus of the inverse cut propagators, and we assume Minkowski kinematics.

A singularity of the second type corresponds to a configuration where Y_C vanishes for a subset C of propagators. To such a singularity we associate the cut integral $\mathcal{C}_{\infty C}\widetilde{J}_n$, where, as well as encircling the poles of propagators in C, the contour now also winds around the branch point at infinite loop momentum. It was shown in [21] that a cut integral associated to a singularity of the second type can be written as a linear combination of cut integrals associated to singularities of the first type. This means that the basis $\widetilde{J}_n$ of one-loop integrals can be lifted to a basis $\mathcal{C}_C\widetilde{J}_n$ of one-loop cut integrals, which can be chosen to contain only cut integrals associated with singularities of the first type.

It is often convenient to normalise each basis integral $\widetilde{J}_n$ to its maximal cut in integer dimensions j_n related to the leading singularity of the integral, which is defined by

$$j_n \equiv \lim_{\epsilon\to 0}\mathcal{C}_{[n]}\widetilde{J}_n = \begin{cases} 2^{1-n/2}\mathrm{i}^{n/2}Y_{[n]}^{-1/2}, & \text{for } n \text{ even},\\ 2^{(1-n)/2}\mathrm{i}^{(n-1)/2}\,\mathrm{Gram}_{[n]}^{-1/2}, & \text{for } n \text{ odd}.\end{cases} \tag{2.15}$$

Choosing this normalisation gives us the basis integrals

$$J_n = \widetilde{J}_n/j_n. \tag{2.16}$$

These are *pure* functions, meaning that the coefficients in their Laurent expansion in ϵ do not contain rational or algebraic functions of the external kinematic variables [32].

2.3 Multiple polylogarithms

Multiple polylogarithms (MPLs) are a class of functions that generalise the classical polylogarithms to several variables. Our interest in them here stems from the fact that they arise in the computation of a large class of Feynman integrals, and moreover that they may be endowed with a coproduct operation, which can in turn be used to define the symbol map. In this section, we define the integral representation of the multiple polylogarithms and discuss some of their properties, based on the treatment in [14, 32].

We begin by defining the iterated integral representation of the multiple polylogarithm $G(z_1,\ldots,z_k;y)$ for $y, z_i \in \mathbb{C}$ where all z_i are equal to zero by

$$G(\underbrace{0,\ldots,0}_{k\text{ times}};y) = \frac{1}{k!}\log^k(y). \tag{2.17}$$

If at least one z_i is nonzero, then we recursively define

$$G(z_1,z_2,\ldots,z_k;y) = \int_0^y \frac{dt}{t-z_1}G(z_2,\ldots,z_k;t). \tag{2.18}$$

We say that $G(z_1,\ldots,z_k;y)$ has a *trailing zero* if $z_k = 0$. For multiple polylogarithms without trailing zeros, the recursive definition gives

$$G(z_1,\ldots,z_k;y) = \int_0^y \frac{dt_1}{t_1 - z_1} \int_0^{t_1} \frac{dt_2}{t_2 - z_2} \cdots \int_0^{t_{k-1}} \frac{dt_k}{t_k - z_k}. \tag{2.19}$$

More generally, we can have

$$G_{n_1\ldots n_k}(z_1,\ldots,z_k;y) = G(\underbrace{0,\ldots,0,z_1}_{n_1},\ldots,z_{k-1},\underbrace{0,\ldots,0,z_k}_{n_k};y). \tag{2.20}$$

Here, the number $n_1 + \ldots + n_k$ is called the *weight* of the multiple polylogarithm and corresponds to how many integrals are needed to define the function. MPLs also have a sum representation

$$\mathrm{Li}_{n_1\ldots n_k}(x_1,\ldots,x_k) = \sum_{m_1>\ldots>m_k}^{\infty} \frac{x_1^{m_1}}{m_1^{n_1}} \cdots \frac{x_k^{m_k}}{m_k^{n_k}}, \tag{2.21}$$

which is related to the integral representation via

$$\mathrm{Li}_{n_1\ldots n_k}(x_1,\ldots,x_k) = (-1)^k G_{m_1\ldots m_k}\left(\frac{1}{x_1},\ldots,\frac{1}{x_1\cdots x_k};1\right).$$

For example, the classical polylogarithms are $\mathrm{Li}_n(x) = \sum_{m=1}^{\infty} x^m/m^n = -G_n(1/x;1)$.

Feynman integrals can be expressed as hypergeometric functions. When, for instance, the entries of a ${}_2F_1$ function are integers or half-intergers in the limit where ϵ goes to zero, the ϵ-expansion of the function can be expressed in terms of uniform-weight MPLs [33]. For example,

$${}_2F_1(-\epsilon, 1-\epsilon; 1+\epsilon; a) = [1]\epsilon^0 + [\log(1-a)]\epsilon^1 + \left[\frac{3}{2}\log^2(1-a) + 2\,\mathrm{Li}_2(a)\right]\epsilon^2 + \mathcal{O}(\epsilon^3). \tag{2.22}$$

All one-loop diagrams are expected to be expressed in this form [34], making symbol analysis appropriate.

2.4 Coproducts and the symbol

The primary focus of this work is a linear map called the *symbol* which can be applied to multiple polylogarithms. In order to define the symbol of an MPL, we must first briefly discuss the algebraic structure of the vector space of MPLs, and introduce the coproduct operation.

To this end, let $\mathcal{A}$ denote the $\mathbb{Q}$-vector space spanned by all multiple polylogarithms. This can be turned into an algebra using the fact that iterated integrals form a shuffle algebra, with a shuffle product given by

$$G(z_1, z_2, \ldots, z_k; y) \cdot G(z_{k+1}, \ldots, z_r; y) = \sum_{\text{shuffles } \sigma} G(z_{\sigma(1)}, z_{\sigma(2)} \ldots z_{\sigma(r)}; y), \tag{2.23}$$

where a permutation σ is said to be a *shuffle* of $(1,\ldots,k)$ and $(k+1,\ldots,r)$ if in

$$(\sigma(1), \sigma(2), \ldots, \sigma(r)) \tag{2.24}$$

the relative order of $1,2,\ldots,k$ and of $k+1,\ldots,r$ is preserved. The shuffle product preserves the weight, meaning that the shuffle product of two multiple polylogarithms of weights n_1 and n_2 is a linear combination of multiple polylogarithms of weight $n_1 + n_2$. Formally, we say that the algebra of multiple polylogarithms is *graded* by the weight,

$$\mathcal{A} = \bigoplus_{n=0}^{\infty} \mathcal{A}_n, \text{ with } \mathcal{A}_{n_1} \cdot \mathcal{A}_{n_2} \subseteq \mathcal{A}_{n_1+n_2}, \tag{2.25}$$

where $\mathcal{A}_n$ is the $\mathbb{Q}$-vector space spanned by all multiple polylogarithms of weight n, and we define $\mathcal{A}_0 = \mathbb{Q}$.

Moreover, the quotient space $\mathcal{H} = \mathcal{A}/(\mathrm{i}\pi\mathcal{A})$ is conjectured to form a Hopf algebra, and in particular can be equipped with a coproduct $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$, which is coassociative,

$$(\Delta \otimes \mathrm{id})\Delta = (\mathrm{id} \otimes \Delta)\Delta, \tag{2.26}$$

a homomorphism,

$$\Delta(a \cdot b) = \Delta(a) \cdot \Delta(b), \tag{2.27}$$

and respects the weight. For our purposes, the results of applying the coproduct Δ to the ordinary logarithm and the classical polylogarithms will be of most interest:

$$\Delta(\log z) = 1 \otimes \log z + \log z \otimes 1, \tag{2.28a}$$

$$\Delta(\mathrm{Li}_n(z)) = 1 \otimes \mathrm{Li}_n(z) + \sum_{k=0}^{n-1} \frac{1}{k!} \mathrm{Li}_{n-k}(z) \otimes \log^k(z). \tag{2.28b}$$

The coassociativity of the coproduct means that it can be uniquely iterated. For any partitition $(n_1, \dots, n_k)$ of n, we define the iterated coproduct

$$\Delta_{n_1,\dots,n_k} : \mathcal{H}_n \to \mathcal{H}_{n_1} \otimes \cdots \otimes \mathcal{H}_{n_k}. \tag{2.29}$$

The maximal iteration of the coproduct correpsonds to the partition $(1, \dots, 1)$, and gives the *symbol* $\mathcal{S}$ of a transcendental function F,

$$\mathcal{S}(F) \equiv \Delta_{1,\dots,1}(F) \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_1. \tag{2.30}$$

As every element of $\mathcal{H}_1$ is simply an ordinary logarithm, we usually omit the 'log' signs when discussing the symbol, and it is implicitly understood that each entry is actually the argument of a logarithm. This means that the symbol of a given multiple polylogarithm f has the form

$$\mathcal{S}(f) = \sum_{i_1,\dots,i_j} c_{i_1,\dots,i_j} f_{i_1} \otimes \cdots \otimes f_{i_j},$$

where the $c_{i_1,\dots,i_j}$ are coefficients and the values f_i are the symbol *letters*, which are the logarithm arguments [14]. In the case of Feynman integrals, the letters are functions of the external momenta and propagator masses. The letters which occur in the symbols of one-loop Feynman integrals have been found to exhibit some interesting patterns. For example, the 'first entry condition' states that the first entries of $\mathcal{S}(f)$ are logarithms of Mandelstam invariants [32].

The symbol alphabet for one-loop integrals in dimensional regularisation has been studied in detail for the generic case of finite integrals where all propagators are massive [19]. However, the symbol alphabet for non-generic one-loop integrals, such as those which are not finite or which involve massless propagators, is not yet fully understood. The study of the symbol dictionaries of these non-generic cases is the main focus of this work.

2.5 Symbol alphabet from discriminants

The principal A-determinant $\widetilde{E_A}$ is the product of the modified Cayley and Gram determinants and their non-zero minors. The factors of $\widetilde{E_A}$ are the rational letters in the symbol. In [26], square root letters are inferred by applying Jacobi determinant identities of the form

$$p \cdot q = f^2 - g = (f - \sqrt{g})(f + \sqrt{g}) \tag{2.31}$$

to give new letters

$$\frac{f - \sqrt{g}}{f + \sqrt{g}}. \tag{2.32}$$

Here p and q are both factors of the principal A-determinant, and the ratio of the factors on the right are taken to form the new letters. The number of letters of the general n-point graph with dimension $D = D_0 - 2\epsilon$ is

$$|W| = \begin{cases} 2^{n-3}\big(n^2+3n+8\big) - \frac{1}{6}\big(n^3+5n+6\big)\,, & D_0 \text{ even,} \\ 2^{n-3}\big(n^2+3n+8\big) - \frac{1}{2}\big(n^2+n+2\big)\,, & D_0 \text{ odd.} \end{cases} \tag{2.33}$$

Since we work in even dimension, the numbers of letters of interest are $|W| = 1, 5, 18, 57\ldots$.

The choice of alphabet is not unique and a change of variables leads to a different alphabet using properties of logs. When solving Feynman integrals and their cuts, there may be a more natural change of variables to express the solution in terms of known functions. These variables may not form the minimum alphabet. It may also be the case that a certain alphabet allows us to 'undo' the symbol more easily. The algorithmic procedure (see for example [13, 14, 17]) of reconstructing an MPL from a symbol depends on the alphabet chosen. As mentioned in the introduction, there is also a machine learning approach [18] for which the machine 'learns' to employ inversion, reflection, duplication and cyclic identities related to MPLs in order to simplify the symbol. This is obviously sensitive to a judicious choice of alphabet.

Starting with a generic diagram, it is possible to take a limit $\{m_i^2, p_j^2\} \to 0$ of $\widetilde{E_A}$ by simply removing any letters which vanish in the limit. While this limit is well-defined, it is not known if the letters resulting from the re-factorised $\widetilde{E_A}$ after the limit are subject to the same constraints as before. In other words, it is possible to take any limit of a symbol's alphabet (as shown in [26]), but in 4 we shall show this does not follow for the symbol's dictionary.

3 Using the recursion

3.1 Differential equations and symbol recursion

We have seen that the integrals J_n form a basis of one-loop Feynman diagrams. As argued in [19], in this basis of pure MPL functions the latter satisfy

$$dJ_{G_X} = \sum_{\emptyset \neq Y \subseteq X} \Omega_{X,Y} J_{G_Y}\,, \qquad \Omega_{X,Y} = \Omega^{(0)}_{X,Y} + \epsilon \Omega^{(1)}_{X,Y}. \tag{3.1}$$

Here (G, C) denotes a graph with edge set E_G of which C are cut and G_X is the induced graph from keeping only $X \subseteq E_G$. The differential equation (3.1) holds because the weight of pure MPLs is lowered by a derivative [24] where Ω is a matrix whose entries are logarithmic one-forms indexed by subsets X, Y of propagators. Using the properties of the diagrammatic coaction conjectured in [19], and the fact that cut integrals $\mathcal{C}_C J_G$ satisfy the same differential equation as J_G, one can ultimately derive from (3.1) the following recursion formulae for the symbol. For $n = |E_G|$ odd,

$$\begin{aligned} \mathcal{S}[J_G] = \epsilon \mathcal{S}[J_G] \otimes (\mathcal{C}_{E_G} J_G)^{(1)} + \sum_{e \in E_G} \mathcal{S}\big[J_{G\setminus e}\big] \otimes \big(\mathcal{C}_{E_G \setminus e} J_G\big)^{(0)} \\ + \sum_{\{e,f\} \in E_G} \mathcal{S}\big[J_{G\setminus\{e,f\}}\big] \otimes \left(\mathcal{C}_{E_G\setminus\{e,f\}} J_G + \frac{1}{2}\mathcal{C}_{E_G\setminus e} J_G + \frac{1}{2}\mathcal{C}_{E_G\setminus f} J_G\right)^{(0)}. \end{aligned} \tag{3.2}$$

For n even,

$$\begin{aligned} \mathcal{S}[J_G] = \epsilon \mathcal{S}[J_G] \otimes (\mathcal{C}_{E_G} J_G)^{(1)} + \sum_{e \in E_G} \epsilon \mathcal{S}\big[J_{G\setminus e}\big] \otimes \left(\mathcal{C}_{E_G\setminus e} J_G + \frac{1}{2}\mathcal{C}_{E_G} J_G\right)^{(1)} \\ + \sum_{\{e,f\} \in E_G} \mathcal{S}\big[J_{G\setminus\{e,f\}}\big] \otimes (\mathcal{C}_{E_G\setminus\{e,f\}} J_G)^{(0)}. \end{aligned} \tag{3.3}$$

The recursion is on the length of the words in the symbol. Recall that a symbol looks like

$$\mathcal{S}[J_n] = \sum_{i=-\lceil n/2 \rceil}^{\infty} (\text{words of length } \lceil n/2 \rceil + i)\epsilon^i \tag{3.4}$$

such that multiplying a symbol by ϵ amounts to shifting the length of the words down by 1. What feeds the recursions (3.2) and (3.3) is either the symbol with fewer letters of the diagram, or its once- or twice-pinched subgraphs, whereas the maximal $\mathcal{C}_{E_G} J_G$, next-to-maximal (Nmaximal) $\mathcal{C}_{E_G\setminus e} J_G$ and next-to-next-to-maximal cuts (NNmaximal) $\mathcal{C}_{E_G\setminus\{e,f\}} J_G$ generate the new letters. Besides the pinched symbols

$$\mathcal{S}\big[J_{G\setminus e}\big] \quad \text{and} \quad \mathcal{S}\big[J_{G\setminus\{e,f\}}\big], \tag{3.5}$$

the recursion involves the cuts

$$\begin{aligned} n \text{ odd:} \quad & \mathcal{C}^{(1)}_{E_G} J_G, \mathcal{C}^{(0)}_{E_G\setminus e} J_G, \mathcal{C}^{(0)}_{E_G\setminus\{e,f\}} J_G, \\ n \text{ even:} \quad & \mathcal{C}^{(1)}_{E_G} J_G, \mathcal{C}^{(1)}_{E_G\setminus e} J_G, \mathcal{C}^{(0)}_{E_G\setminus\{e,f\}} J_G. \end{aligned}$$

But, in principle, if one could evaluate all these ingredients as well as the algebraic term of the symbol (at order $\epsilon^{-\lceil n/2 \rceil}$) then one could read off from the recursion what words are allowed; the symbol dictionary.

n **odd**	n **even**
$\otimes\mathcal{C}^{(1)}_{E_G} = \otimes\left(\frac{\text{Gram}_{E_G}}{Y_{E_G}}\right)$	$\otimes\mathcal{C}^{(1)}_{E_G} = \otimes\left(\frac{\text{Gram}_{E_G}}{4Y_{E_G}}\right)$
$\otimes\mathcal{C}^{(0)}_{E_G\setminus e} = \otimes\left(\frac{\sqrt{1-\eta}-1}{\sqrt{1-\eta}+1}\right)$	$\otimes\mathcal{C}^{(1)}_{E_G\setminus e} = -\frac{1}{2}\otimes\left(\frac{\text{Gram}_{E_G\setminus e}}{4Y_{E_G\setminus e}}\right) - \otimes\left(1+\sqrt{1-\frac{1}{\eta}}\right)$
$\otimes\mathcal{C}^{(0)}_{E_G\setminus\{e,f\}} = \otimes\left(\frac{a_1+a_2+a_3+a_4+a_5}{a_1+a_2+a_3+a_4-a_5}\right)$	$\otimes\mathcal{C}^{(0)}_{E_G\setminus\{e,f\}} = \frac{1}{2}\otimes\left(\frac{\sqrt{1-\rho}+1}{\sqrt{1-\rho}-1}\right)$

Table 1. Maximal, Nmaximal and NNmaximal cuts in terms of Gram and Cayley determinants for generic n-point one-loop integrals. These ultimately decide what letters appear in the symbol worlds.

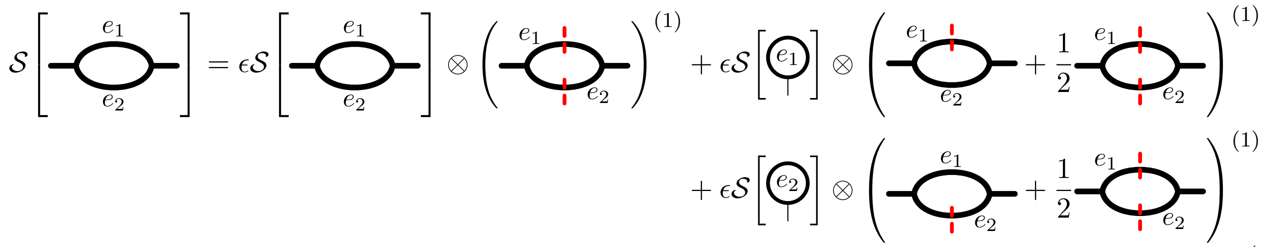
For generic diagrams, the necessary expressions (in shorthand) were found in [19] and are presented in Table 1, where a_i are complicated expressions in terms of determinants and can be found in [19, App. D]. We have also defined the useful ratios in accordance with [21]

$$\eta \equiv \frac{Y_{E_G}\,\text{Gram}_{E_G\setminus e}}{\text{Gram}_{E_G}\,Y_{E_G\setminus e}}, \qquad \rho \equiv \frac{Y_{E_G} Y_{E_G\setminus\{e,f\}}}{Y_{E_G\setminus e} Y_{E_G\setminus f}}. \tag{3.6}$$

The alphabet of generic one-loop Feynman integrals is known [26, 35]. Given the above recursion, one can in principle determine the rules governing the sequence and appearance of letters in the symbol of the generic bubble, triangle and box. However, our goal of obtaining the symbol dictionary for non-generic kinematics is not yet reached. Let us see this with some examples.

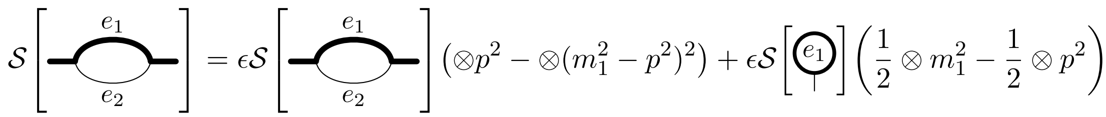
3.2 Example: bubble with one massive propagator

The following toy example displays all of the relevant themes at play when analysing symbols of non-generic one-loop integrals. The first step is to write the recursion for the generic version of the diagram. In this case the twice-pinched graphs do not contribute and we get, in diagram form,



(3.7)

Now suppose momentum p flows through the external legs, and the edges e_i have masses m_i. If we choose the edge e_2 to become massless, then the tadpole integral with mass m_2^2 will vanish, and so will its symbol. Also, cut integrals where a single massless propagator is cut identically vanish [21]. All that is left is



(3.8)

where $\otimes a$ is short for $\otimes \log a$ and we have substituted the result [19]

$$\otimes(\mathcal{C}_{e_1} J_2(p^2; m_1^2, 0))^{(1)} = \otimes(m_1^2 - p^2) + \frac{1}{2} \otimes m_1^2 - \otimes p^2, \tag{3.9}$$

$$\otimes(\mathcal{C}_{e_1 e_2} J_2(p^2; m_1^2, 0))^{(1)} = \otimes p^2 - 2 \otimes (m_1^2 - p^2). \tag{3.10}$$

One can perform the tadpole and bubble integrals manually and find the base of the recursion, which is the order $-\lceil n/2 \rceil = -1$ of the symbol: $\mathcal{S}\big[J_2(p^2; m_1^2, 0)\big]^{(-1)} = -\frac{1}{2}$, $\mathcal{S}\big[J_1(m_1^2)\big]^{(-1)} = -1$. In turn, we see a manifest cancellation of the letter p^2 at first non-trivial order. Since the symbol of $J_1(m_1^2)$ is just a string of $\otimes m_1^2$, there is no other opportunity for the letter p^2 to start a word: it has missed its chance at being a first letter. Furthermore, we see that m_1^2 only gets tagged on to words coming from the tadpole. Words of the type $m_1^2 \otimes \ldots \otimes m_1^2$ can of course reenter the recursion through the first term in (3.8), but at no point can p^2 or $(m_1^2 - p^2)$ precede m_1^2. Looking at the recursion in this way allows us to simply read off the alphabet[2]

$$\mathcal{A}\big(J_2(p^2; m_1^2, 0)\big) = \Big\{p^2, m_1^2, m_1^2 - p^2\Big\} \tag{3.11}$$

and the dictionary: (i) the letter p^2 cannot come first, (ii) no other letter can precede m_1^2. Finally we understand that table in the introduction and can pat ourselves on the back. The following could be asked:

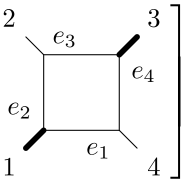
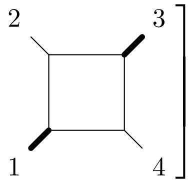
– Does this alphabet follow from the limit $m_2^2 \to 0$ of the generic bubble?
– Is this dictionary a special case of that of the generic bubble?

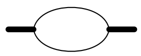
We postpone these issues until 4.

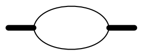
3.3 Example: two-mass easy box

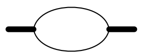
For diagrams with more than two propagators, the recursion is substantially more involved. Consider a box diagram where we set $p_2^2 = p_4^2 = 0$ with all massless propagators. This is commonly known as the two-mass easy box since it has a difficult sibling (where $p_2^2 \neq 0$ but $p_3^2 = 0$) and has recursion

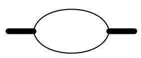
[2] Letters are defined up to an overall constant factor since the symbol works modulo $i\pi$ (taking care of minuses) and because the symbol can be defined as the $d\log$ structure of MPL iterated integrals (killing constants).

$\mathcal{S}\Big[$  $\Big] = \epsilon\mathcal{S}\Big[$  $\Big]\big(\otimes(s+t-p_1^2-p_3^2) - \otimes(st-p_1^2p_3^2)\big)$

$+\mathcal{S}\Big[\, p_3^2$  $\Big]\big(-\otimes p_3^2 + \otimes(p_3^2-s)(p_3^2-t) - \otimes(s+t-p_1^2-p_3^2)\big)$

$+\mathcal{S}\Big[\, t$  $\Big]\big(+\otimes t - \otimes(t-p_1^2)(t-p_3^2) + \otimes(s+t-p_1^2-p_3^2)\big)$

$+\mathcal{S}\Big[\, p_1^2$  $\Big]\big(-\otimes p_1^2 + \otimes(p_1^2-s)(p_1^2-t) - \otimes(s+t-p_1^2-p_3^2)\big)$

$+\mathcal{S}\Big[\, s$  $\Big]\big(+\otimes s - \otimes(s-p_1^2)(s-p_3^2) + \otimes(s+t-p_1^2-p_3^2)\big).$ (3.12)

The reason why there are no triangles despite the formula (3.3) is because of a useful relation between symbols of triangles with one or two massive external legs and bubbles, which was used here for a more compact result. To use the recursion, we start by looking at the base case. Only the massless bubbles contribute $\mathcal{S}\big[J_2(p^2)\big]^{(-1)} = -1$ towards the lowest order expression giving the expected $\mathcal{S}^{(-1)} = \otimes p_1^2 - \otimes p_3^2 - \otimes s - \otimes t$. Then at next order, the four bubbles exactly cancel this expression and so the long letter $(s+t-p_1^2-p_3^2)$ does not appear in two-letter words either! This explains why this letter only appears in the *third* entry onwards. The full alphabet is

$$\mathcal{A}\big(J_4(p_1^2,0,p_3^2,0)\big) = \Big\{p_1^2, p_3^2, s, t, \quad p_1^2-s,\ p_1^2-t,\ p_3^2-s,\ p_3^2-t, \quad st-p_1^2p_3^2,\ s+t-p_1^2-p_3^2\Big\} \tag{3.13}$$

and after some looking and some thinking, one can realise the dictionary is made of words satisfying:

- Only one of p_1^2, p_3^2, s, t appears in a word, and it appears at least once.
- A word consists first of a sequence of blue letters, then a single red letter may appear, then the rest of the letters are black.
- The letter $(s+t-p_1^2-p_3^2)$ can only appear in the third position onward.
- Each word that can appear does so uniquely and has a coefficient of ± 1.

This 'method' of deducing dictionaries requires data about contracted sub-integrals, the cuts of at certain orders in ϵ, and the base value of the symbol, i.e., the lowest non-zero weight words. However, the point being to *skip* the integral (and potentially difficult cuts), this is not a very satisfying explanation. Avoiding the circular issue of contracted graphs, the key difficulty is knowing exactly what letters appear through cuts in the right-hand side of recursion. This is also postponed to 4.

Still, we successfully determined alphabets and dictionaries for the simplest non-generic diagrams as tabulated in Table 2.

There is another important avenue to the integrals J_n through the relation with small cuts [19]:

$$\sum_{i\in[n]} \mathcal{C}_{\{i\}} J_n + \sum_{i<j\in[n]} \mathcal{C}_{\{i,j\}} J_n = -\epsilon J_n \mod \mathrm{i}\pi. \tag{3.14}$$

Clearly, if one knows the cuts on the left-hand side at any order in ϵ, one can obtain the integral itself. This amounts to (and agrees with) using the recursion for the bubble, where a maximal cut is the cut of two propagators. For non-generic $n \geq 3$ integrals, explicit formulae for the left of (3.14) are nearly as difficult as the integrals themselves and so the recursion which deals with maximal and Nmaximal cuts is friendlier. We are still working on the derivation of NNmaximal cuts in non-generic cases, which would yield (3.14) for triangles. This involves a parametrisation of the Feynman/cut integral in compactified space $\mathbb{CP}^{D+1}$ as shown in [21], but this time relaxing the assumption that determinants are non-zero.

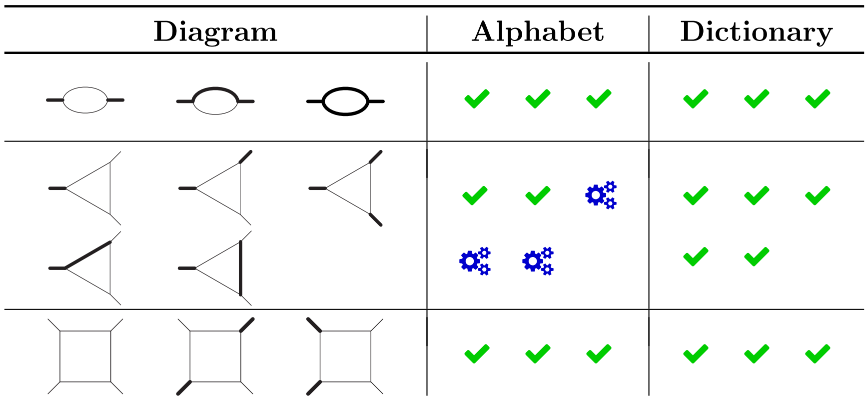


Table 2. Summary of results using the symbol recursions. A set of cogs means it is still unclear whether our alphabet is the minimal or 'proper' alphabet. See discussion in 4.

3.4 Comparison with literature

Many results for the symbol from the literature emerge from Landau analysis. The Landau variety L_C for some $C \subseteq [n]$ is defined by the set of kinematics variables such that $\mathrm{Gram}_C = 0$ or $Y_C = 0$, type-I and type-II singularities respectively. The discontinuity of a one-loop integral around L_C is defined as the difference before and after analytically continuing the external kinematics along a small positively-oriented circle around L_C. The discontinuity can be related to cuts by

$$\mathrm{Disc}_C J_G = -N_C \, \mathcal{C}_C J_G \mod \mathrm{i}\pi \,, \quad C \subseteq E_G \cup \{\infty\} \,, \tag{3.15}$$

where N_C is an integer [21]. Sequences of discontinuities are related to sequences of letters in the symbol. The Steinmann relations [28, 36] put constraints on sequential discontinuities in that scattering amplitudes do not have consecutive discontinuities in partially overlapping momentum channels.

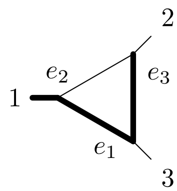
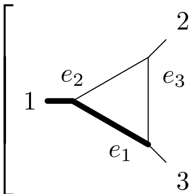
The recent *Euler characteristic test* [29] places negative constrains on which sequences of discontinuities can occur by taking discontinuities of $\mathcal{F}$ in a Feynman parameter integral. Plotting $\mathcal{F} = 0$ and setting a letter to zero, the integration region is altered such that some integration boundaries become obsolete. This means the associated singularities become no longer reachable. Taking the Euler characteristic of the space of Feynman parameters without the singular loci gives information about if the integral becomes singular, and can be used to compare if one letter can follow at any point after another. This method involves no integration itself and so is useful in deriving hierarchical constraints with little computational cost. Applied to the two-mass easy box above, it recovers the sequential constraints between the blue, red and black letters (3.13). However since this test does not give information about coefficients, it falls short of reporting that a word of the form

$$\text{blue} \otimes \cdots \otimes \text{blue} \otimes (s + t - p_1^2 - p_3^2) \otimes \cdots \tag{3.16}$$

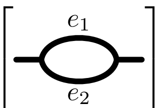
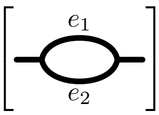
does not appear because it cancels in the recursion. In our approach, we use the cut integral data to compute the symbol in full, including coefficients. This means we can find this other type of constraint, that which derives from a hidden cancellation in the recursion. For instance, the triangle with one massive external leg and massive opposite propagator has a symbol with no one-letter words. This has consequences for the recursion, for example by prohibiting $(m^2 + p^2)$ from ever occurring in the first position. It is unclear how to predict such cancellations because the form of the recursion depends heavily on which kinematics are left generic or set to zero.

4 Taking non-generic limits

In general, it is not possible to take limits at the symbol level to recover the symbol of a less generic diagram. A simple counterexample involves the triangle integral $J_3(p_1^2, 0, 0; m_1^2, 0, m_3^2)$. The symbol at the first nontrivial order ϵ equals zero, with the contribution from all cuts in the recursion cancelling, while its limit $m_3^2 \to 0$ is non-zero:

$\mathcal{S}\Big[$  $\Big]^{(-1)} = 0,\qquad \mathcal{S}\Big[$  $\Big]^{(-1)} = \otimes(m_1^2 - p_1^2) - \otimes m_1^2 \qquad (4.1)$

However, there is a straightforward way to take the limit of letters of an alphabet. For example, the five-letter generic bubble alphabet $\{A_i\}$ (found using the method [26] described in 2.5) does not reduce to (3.11) in the limit $m_2^2 \to 0$: two letters go to zero or infinity. But if we ignore the divergent letters the recursion reduces properly. Concretely, for orders $N \geq 1$ the generic bubble recursion looks like

$S\Big[$  $\Big]^{(N)} = S\Big[$  $\Big]^{(N-1)} \otimes A_3 + \frac{1}{2}\left[(\otimes A_1)^N \otimes A_4 + (\otimes A_2)^N \otimes A_5\right] \qquad (4.2)$

where we recognise the tadpole symbols as strings of masses and the generic alphabet is [26]

$$A_1 = m_1^2, \qquad A_2 = m_2^2, \qquad A_3 = \frac{p^2}{\lambda(p^2, m_1^2, m_2^2)},$$
$$A_4 = \frac{-m_1^2 - m_2^2 - p^2 - \sqrt{\lambda(p^2, m_1^2, m_2^2)}}{-m_1^2 - m_2^2 - p^2 + \sqrt{\lambda(p^2, m_1^2, m_2^2)}}, \qquad A_5 = \frac{-m_1^2 - m_2^2 + p^2 - \sqrt{\lambda(p^2, m_1^2, m_2^2)}}{-m_1^2 - m_2^2 + p^2 + \sqrt{\lambda(p^2, m_1^2, m_2^2)}}, \tag{4.3}$$

in which we use the Källén function

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz. \tag{4.4}$$

In the limit $m_2^2 \to 0$ we have $\lambda(p^2, m_1^2, m_2^2) \to (m_1^2 - p^2)^2$ such that

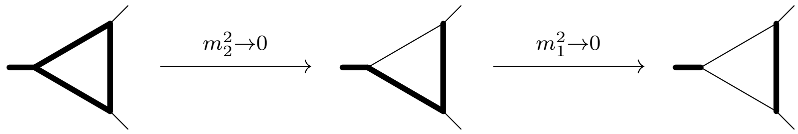
$$A_2 \to 0, \qquad A_3 \to \frac{p^2}{(m_1^2 - p^2)^2}, \qquad A_4 \to \frac{m_1^2}{p^2}, \qquad A_5 \to \infty. \tag{4.5}$$

The letter A_5 becomes indeterminate but, in this case, the correct alphabet (3.11) and recursion (3.8) are retrieved since the tadpole's symbol vanishing masks this indeterminacy. This agrees with the finding in [26] and [35] that taking non-generic kinematic limits is well-defined in the space of Landau solutions (letters) of one-loop Feynman integrals, but not for their differential equations. To retrieve the symbol of a non-generic one-loop Feynman integral, then, one might be tempted to conjecture the following prescription:

- First take the limit from the generic alphabet to this non-generic case[3].
- Then, considering the symbol recursion pertaining to the generic integral, identify letters which become 1, those which diverge (go to 0 or ∞) and those which degenerate to become identical.
- Simplify the recursion by omitting terms in which letters are 0, 1 or ∞, and by combining factors of newly identical letters.

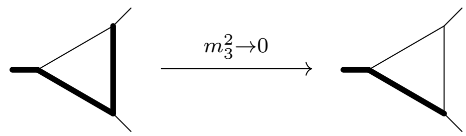
[3] As we will see with triangles, there seems to be no issue with the *order* in which we take limits.

This is wrong. For example, the following sequence of limits



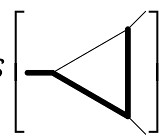
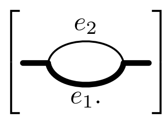
(4.6)

is one such sequence where it *is* possible to follow the above prescription (in fact the limit is well-defined at symbol level), while for the limit



(4.7)

this prescription fails. The reason for this is that the recursion for $J_3(p_1^2, 0, 0; m_1^2, 0, m_3^2)$ has a term of the form

$\mathcal{S}$  $\sim S$ 

$\otimes \left(\frac{p_1^2 + m_3^2 - m_1^2}{m_3^2}\right)$ (4.8)

whereas the 1-mass opposite and 1-mass adjacent have respectively for this cut

$$\mathcal{C}_{E_G \backslash e_3} J_3(p_1^2, 0, 0; 0, 0, m_3^2) = \log\left(\frac{p_1^2 + m_3^2}{m_3^2}\right), \tag{4.9}$$

$$\mathcal{C}_{E_G \backslash e_3} J_3(p_1^2, 0, 0; m_1^2, 0, 0) = \log\left(\frac{(p_1^2 - m_1^2)^2}{p_1^2}\right). \tag{4.10}$$

Clearly it is possible to take the $m_1^2 \to 0$ limit of (4.8) to recover the 1-mass opposite term, yet it is not possible to do the same for $m_3^2 \to 0$ to recover the 1-mass adjacent diagram. The full expression

$$\mathcal{C}_{[n-1]} J_n = -2\frac{\Gamma^2(1-\epsilon)}{\Gamma(2-2\epsilon)} \frac{\eta^\epsilon}{\sqrt{1-\eta}+1} \, {}_2F_1\left(1, 1-\epsilon; 2-2\epsilon; \frac{2}{\sqrt{1-\eta}+1}\right) \tag{4.11}$$

for the Nmaximal cut expanded to write (4.8) needs to be re-expanded with a different limiting procedure due to a vanishing of the Cayley determinant which was assumed nonzero in the generic case. In the generic case of $0 < \eta < 1$ for η defined in (3.6), the 1-mass opposite cut takes the form

$$\mathcal{C}_{[n-1]} J_n = \log\left(\frac{\sqrt{Y_{[n]}\,\mathrm{Gram}_{[n-1]} - \mathrm{Gram}_{[n]}\,Y_{[n-1]}} - \sqrt{-\,\mathrm{Gram}_{[n]}\,Y_{[n-1]}}}{\sqrt{Y_{[n]}\,\mathrm{Gram}_{[n-1]} - \mathrm{Gram}_{[n]}\,Y_{[n-1]}} + \sqrt{-\,\mathrm{Gram}_{[n]}\,Y_{[n-1]}}}\right) + \mathcal{O}(\epsilon), \tag{4.12}$$

while the 1-mass adjacent requires the special case $\eta = 0$

$$\mathcal{C}_{[n-1]} J_n = \frac{1}{\epsilon} + \log\left(\frac{\mathrm{Gram}_{[n-1]}}{4Y_{[n-1]}}\right) + \mathcal{O}(\epsilon). \tag{4.13}$$

At the point where one of these expressions was chosen for the cuts, it made the other expression inaccessible via limit-taking.

Even when the limit of the symbol can be taken from one diagram to another, negative constraints on words may become invalid as letters degenerate. For instance, a constraint on the two-mass easy box is that $s \otimes s \otimes (s+t-p_1^2-p_3^2)$ would be forbidden. In the massless box, this forbidden word becomes $s \otimes s \otimes (s+t)$ in the limit. However, a different word $s^2 \otimes (p_1^2 - s) \otimes (s+t-p_1^2-p^2)$ reduces down to same forbidden word, invalidating the naive rule. Additionally, after the limits the coefficients of words may be altered as letters degenerate, possibly introducing new cancellations and thus new constraints. For instance, $s \otimes s \otimes t$ does not

Limit	$J_2^{(0)}(p^2; m_1^2, m_2^2) \mod i\pi$	**Full determinants**
–	$\log\left[\sqrt{\frac{r_0}{r_1}}\left(1+\sqrt{1-\frac{r_1}{r_0}}\right)\sqrt{\frac{r_0}{r_2}}\left(1+\sqrt{1-\frac{r_2}{r_0}}\right)\right]$	$Y_{[2]} = -\frac{1}{4}\left((p^2)^2 - 2p^2(m_1^2 - m_2^2) + (m_1^2 + m_2^2)^2\right)$ $\text{Gram}_{[2]} = p^2$
$m_2^2 \to 0$	$\log\left[\sqrt{\frac{r_0}{r_1}}\left(1+\sqrt{1-\frac{r_1}{r_0}}\right)2\sqrt{r_0}\right]$	$Y_{[2]} = -\frac{1}{4}\left(p^2 - m_1^2\right)^2$ $\text{Gram}_{[2]} = p^2$
$m_{1,2}^2 \to 0$	$\log[2\sqrt{r_0}2\sqrt{r_0}] = \log[4r_0]$	$Y_{[2]} = -\frac{1}{4}(p^2)^2$ $\text{Gram}_{[2]} = p^2$

Table 3. The prescription $1/r_i \to 1$ when $r_i = Y_{\{i\}} = m_i^2 \to 0$ reproduces the correct expression for the bubble with two, one or zero massive propagors.

occur. This could only be predicted if the exact coefficients of all words were known before such that new coefficients can be determined. On the other hand, the two-mass easy rule that $s+t-p_1^2-p_3^2$ can only occur in the 3rd position onward still holds in the massless limit where $s+t$ only occurs in the 3rd position onward. This is because the new letter $s+t$ can only be reached one way. From the counterexamples above we may conclude that is not possible to take the limit of a generic dictionary as some words may cancel in the symbol.

That sums up the discussion concerning the symbol recursion. Let us investigate what happens when we take non-generic limits in the other avenue to the symbol provided by (3.14). For the bubble, adding the maximal and Nmaximal cuts at order ϵ^1 to access $J_2^{(0)}$ yields the results in Table 3 where the ratios are

$$r_0 \equiv \frac{Y_{[2]}}{\text{Gram}_{[2]}}, \qquad r_i \equiv \frac{Y_{\{i\}}}{\text{Gram}_{\{i\}}} = m_i^2. \tag{4.14}$$

It seems that the prescription of setting $1/r_i = 1$ whenever the limit $m_i^2 = r_i \to 0$ is taken faithfully yields the right expression. This hints at a possible way of correctly taking limits of the kinematic variables when dealing with cuts in terms of determinants. The next easiest example would be triangle diagrams, for which there are many more non-generic cases to consider. As mentioned in 3.3, the above analysis would require NNmaximal cuts of triangles for which an integral parametrisation is presented in [21] with the assumptions $Y_{\mathcal{C}} \neq 0$, $m_i^2 \neq 0$.

However, a level of scepticism is warranted regarding this 'hint' from Table 3. The bubble is particularly simple in that $\text{Gram}_{\{i\}} = 1$, and the NNmaximal cut formulae never enter consideration. This is related to the polytope picture of Feynman integral kinematics whereby the external momentum variables $q_{j\in C}$ are directed from a chosen origin, say q_1, and whose vectors then point to the vertices of a $(c-1)$-simplex in the case of propagators $1,\ldots,c$ being cut. (See the discussion in [21, p. 3.4].) For bubble, the $(c-1)$-simplex of Nmaximal cuts has no volume $\text{Gram}_{\{i\}} = 1$ since it is just a line. This is hardly enlightening.

5 Conclusion

By using the recursive formula for symbols as a sandbox, we were able to clarify the status of symbol alphabets and dictionaries of non-generic one-loop Feynman integrals in uniform weight dimensional regularisation. This is summarised in Figure 1 below.

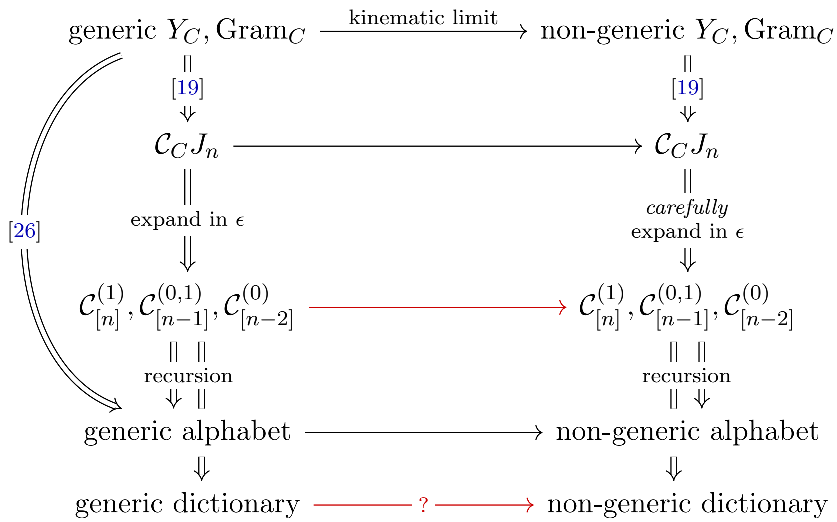


Figure 1. A schema following the flow of information from a diagram's determinants all the way until the symbol alphabet and dictionary. A red line indicates the kinematic limit is not guaranteed to yield the correct expression.

We began by presenting two examples of how one can determine the dictionary for any non-generic integral. This is a case-by-case 'method' and does not provide much insight into *why* certain letters are constrained to appearing in the third entry onward, for example. Unlike the Steinmann relations or recent hierarchical constraints on sequences of letters, which are weaker but more general statements having clear ties to discontinuities in physical channels, using the symbol recursion as a lens is limited by the data available for the given (non-generic) integral. However, the recursion dissects the symbol's evolution in the length of letters by isolating the sources of the letters: cut integrals. This made manifest a problematic feature of cut formulae in terms of hypergeometric functions, which is that expanding in the dimensional regulator ϵ *before* taking kinematic limits leads to divergences and does not yield the correct result depending on the chosen form of said function. This has the important consequence that it is, in general, *not possible* to take non-generic limits of dictionaries.

While it is true that the method of A-determinants in [26] for obtaining a generic alphabet correctly reduces to any non-generic alphabet, this does not inform on how these letters make up the words of a symbol. One must infer these rules, which characterise the dictionary, from the recursion. Starting from a generic dictionary, this means one must go all the way back up the flow of information in Figure 1 to the cuts $\mathcal{C}_C J_n$, and only *then* take the kinematic limit. While not terribly convoluted for bubbles, simple triangles and simple boxes, this becomes impossible for most boxes and even some triangles.

Through the recursion, or the relation between one-loop Feynman integrals and their one- and two-propagator cuts, we hope to continue this work such that we may identify a more tractable and encompassing method for determining (non-generic) dictionaries directly from determinants. This would correspond to

$$Y_C, \mathrm{Gram}_C \overset{?}{\Longrightarrow} \mathcal{S} \text{ dictionary} \tag{5.1}$$

and would promise much in the way of computing scattering amplitudes.

Acknowledgements

We would like to thank Ruth for her attentive guidance this summer, as well as for her captivating teaching these past few years. Thank you to Marvin Anas Hahn for organising the internship this year, and lastly our thanks go to the Hamilton Trust for sponsoring this work.

References

[1] R. P. Feynman. 'Space-time approach to quantum electrodynamics'. *Physical Review* 76 (1949).

[2] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley, 1995.

[3] L. M. Brown and R. P. Feynman. 'Radiative Corrections to Compton Scattering'. *Physical Review* 85 (2 1952).

[4] D. B. Melrose. 'Reduction of Feynman diagrams'. *Nuovo Cimento* 40 (1965).

[5] G. 't Hooft and M. J. G. Veltman. 'Scalar One Loop Integrals'. *Nuclear Physics B* 153 (1979).

[6] G. Passarino and M. J. G. Veltman. 'One Loop Corrections for e+ e- Annihilation Into mu+ mu- in the Weinberg Model'. *Nuclear Physics B* 160 (1979).

[7] W. L. van Neerven and J. A. M. Vermaseren. 'Large loop integrals'. *Physics Letters B* 137 (1984).

[8] Z. Bern, L. Dixon and D. A. Kosower. 'Dimensionally regulated one-loop integrals'. *Physics Letters B* 302.2–3 (1993). arXiv: `9212308 [hep-th]`.

[9] Z. Bern, L. Dixon and D. A. Kosower. 'Dimensionally-regulated pentagon integrals'. *Nuclear Physics B* 412.3 (1994). arXiv: `9306240 [hep-th]`.

[10] C. Duhr. 'Hopf algebras, coproducts and symbols: an application to Higgs boson amplitudes'. *Journal of High Energy Physics* 2012.8 (2012). arXiv: `1203.0454 [hep-ph]`.

[11] N. Arkani-Hamed, T.-C. Huang and Y.-t. Huang. *Scattering Amplitudes For All Masses and Spins*. 2021. arXiv: `1709.04891 [hep-th]`.

[12] A. B. Goncharov, M. Spradlin, C. Vergu and A. Volovich. 'Classical Polylogarithms for Amplitudes and Wilson Loops'. *Physical Review Letters* 105.15 (2010). arXiv: `1006.5703 [hep-th]`.

[13] C. Duhr, H. Gangl and J. R. Rhodes. 'From polygons and symbols to polylogarithmic functions'. *Journal of High Energy Physics* 2012.10 (2012). arXiv: `1110.0458 [math-ph]`.

[14] S. Weinzierl. *Feynman Integrals*. Springer, 2022. arXiv: `2201.03593 [hep-th]`.

[15] O. V. Tarasov. 'Connection between Feynman integrals having different values of the space-time dimension'. *Physical Review D* 54.10 (1996). arXiv: `9606018 [hep-th]`.

[16] R. Lee. 'Space–time dimensionality D as complex variable: Calculating loop integrals using dimensional recurrence relation and analytical properties with respect to D'. *Nuclear Physics B* 830.3 (2010). arXiv: `0911.0252 [hep-ph]`.

[17] J. Broedel, C. Duhr, F. Dulat, B. Penante and L. Tancredi. 'Elliptic symbol calculus: from elliptic polylogarithms to iterated integrals of Eisenstein series'. *Journal of High Energy Physics* 2018.8 (2018). arXiv: `1803.10256 [hep-th]`.

[18] A. Dersy, M. D. Schwartz and X. Zhang. *Simplifying Polylogarithms with Machine Learning*. 2022. arXiv: `2206.04115 [cs.LG]`.

[19] S. Abreu, R. Britto, C. Duhr and E. Gardi. 'Diagrammatic Hopf algebra of cut Feynman integrals: the one-loop case'. *Journal of High Energy Physics* 2017.12 (2017). arXiv: `1704.07931v2 [hep-th]`.

[20] N. Arkani-Hamed and E. Y. Yuan. *One-Loop Integrals from Spherical Projections of Planes and Quadrics*. 2017. arXiv: `1712.09991 [hep-th]`.

[21] S. Abreu, R. Britto, C. Duhr and E. Gardi. 'Cuts from residues: the one-loop case'. *Journal of High Energy Physics* 2017.6 (2017). arXiv: 1702.03163 [hep-th].

[22] S. Abreu, R. Britto and H. Grönqvist. 'Cuts and coproducts of massive triangle diagrams'. *Journal of High Energy Physics* 07 (2015). arXiv: 1504.00206 [hep-th].

[23] R. E. Cutkosky. 'Singularities and Discontinuities of Feynman Amplitudes'. *Journal of Mathematical Physics* 1.5 (1960).

[24] N. Arkani-Hamed, J. Bourjaily, F. Cachazo, A. Hodges and J. Trnka. 'A note on polytopes for scattering amplitudes'. *Journal of High Energy Physics* 2012.4 (2012). arXiv: 1012.6030 [hep-th].

[25] N. Arkani-Hamed, A. Hillman and S. Mizera. 'Feynman polytopes and the tropical geometry of UV and IR divergences'. *Physical Review D* 105.12 (2022). arXiv: 2202.12296 [hep-th].

[26] C. Dlapa, M. Helmer, G. Papathanasiou and F. Tellander. 'Symbol alphabets from the Landau singular locus'. *Journal of High Energy Physics* 2023.10 (2023). arXiv: 2304.02629 [hep-th].

[27] X. Jiang, J. Liu, X. Xu and L. L. Yang. 'Symbol letters of Feynman integrals from Gram determinants'. (2024). arXiv: 2401.07632 [hep-ph].

[28] K. E. Cahill and H. P. Stapp. 'Optical theorems and Steinmann relations'. *Annals of Physics* 90.2 (1975).

[29] H. S. Hannesdottir, L. Lippstreu, A. J. McLeod and M. Polackova. 'Minimal Cuts and Genealogical Constraints on Feynman Integrals'. (2024). arXiv: 2406.05943 [hep-th].

[30] K. J. Larsen and Y. Zhang. 'Integration-by-parts reductions from unitarity cuts and algebraic geometry'. *Physical Review D* 93.4 (2016). arXiv: 1511.01071 [hep-th].

[31] A. Primo and L. Tancredi. 'On the maximal cut of Feynman integrals and the solution of their differential equations'. *Nuclear Physics B* 916 (2017). arXiv: 1610.08397 [hep-th].

[32] S. Abreu, R. Britto, C. Duhr and E. Gardi. 'From multiple unitarity cuts to the coproduct of Feynman integrals'. *Journal of High Energy Physics* 2014.10 (2014). arXiv: 1401.3546 [hep-th].

[33] M. Y. Kalmykov, B. F. Ward and S. Yost. 'All order ϵ-expansion of Gauss hypergeometric functions with integer and half/integer values of parameters'. *Journal of High Energy Physics* 2007.4 (2007). arXiv: 0612240 [hep-th].

[34] J. M. Henn. 'Multiloop Integrals in Dimensional Regularization Made Simple'. *Physical Review Letters* 110.25 (2013). arXiv: 1304.1806 [hep-th].

[35] J. Chen, C. Ma and L. L. Yang. 'Alphabet of one-loop Feynman integrals *'. *Chinese Physics C* 46.9 (2022). arXiv: 2201.12998 [hep-th].

[36] J. L. Bourjaily, H. Hannesdottir, A. J. McLeod, M. D. Schwartz and C. Vergu. 'Sequential discontinuities of Feynman integrals and the monodromy group'. *Journal of High Energy Physics* 2021.1 (2021). arXiv: 2007.13747 [hep-th].