WEINBERG Ch. 12.1, 12.2, 12.3

• Renormalization Beyond Renormalizability •

(or, computing loops in an Effective Field Theory)

We have not yet discussed Effective Field Theories (EFT) on general grounds, but we have already encountered the relevant example of the Chiral Theory, in both its 2 and 3-flavor versions. In these examples, we found a common and remarkable feature: the presence of couplings with negative energy dimension. Among the Relativistic Theories (the ones we will be mainly interested in) the presence of such a coupling is the defining property of the EFT. Even though the distinction is rather artificial, a theory is called Effective if it contains such couplings; if it is instead Renormalizable (i.e., only contains couplings of positive dimension) the theory is normally considered to be "Fundamental", or at least not an Effective one.

The presence of a coupling c of negative energy dimension (say, $[c] = \Delta < 0$) has two consequences. First, based on dimensional analysis, the interaction mediated by c will be

weak at low energy and strong at high energy. In a relativistic theory, in the limit in which E is much bigger than the masses of all the particles, c will always enter through the dimensionless combination

$$\mathcal{C} = c\, E^{-\Delta}$$

It is $\mathcal{C}$ that controls, for instance, the size of the corrections $\delta O/O$ (to any observable O) mediated by c. For $E \to \infty$ ($E \gg c^{1/\Delta}$) the corrections become big, so that the effects of the coupling cannot be computed using perturbation theory. This is very similar to what you already know to happen in QED: approaching the Landau Pole the coupling becomes large and perturbation theory loses its validity. Given that beyond perturbation theory we have no tools to make predictions, it is fair to say that when perturbation theory breaks down it is the entire theory that looses its validity: it will have to be replaced by some more fundamental description. The maximum energy of validity of the effective theory is called "the cutoff" Λ. More on this later

The second implication is more scary, so we deal with it right now: a theory with negative energy dimension coupling is <u>non-renormalizable</u> and you might have read on some old book that non-renormalizable means not predictive. If this was so, we would not be allowed to compute radiative corrections in the chiral theory of pions, and all what we could do is to use it at the tree-level, i.e. at the leading order in E/F_π, because it is only in the strict $E \to 0$ limit that its results coincide with the ones of the Current Algebra. Current Algebra was the original motivation to build the chiral theory, the framework of EFT will offer a different, more physical, motivation. In this framework, loop corrections will be calculable and will allow us to refine the predictions with E/F_π corrections and to go beyond Current Algebra.

Let us first of all summarize the essence of renormalization. Consider a generic theory, endowed with couplings $\{C_\kappa\}$ of energy dimension $[C_\kappa] = \Delta_\kappa$. For simplicity, consider the case in which all the physical states are scalars. Ignoring few subtleties, the discussion that follows applies to spin-one (massless or massive) particles as well, while some modifications are needed if willing to include fermions (see Weinberg chapter 12). Renormalization is best performed on 1-Particle Irreducible Feynman diagrams (1PI), in terms of which any amplitude or observable can be computed without encountering further UV divergences. The dimension in energy of a 1PI with "e" external lines is equal, for bosonic external states, to the one of the corresponding scattering amplitude (remember that the 1PI diagram is amputated). We therefore have:

$$[1PI(e)] = 4 - e$$

the structure of the 1PI is:

$$1PI = \int \frac{d^{4L}p}{(2\pi)^{4L}} \left[\prod_\kappa C_\kappa\right] \frac{N(p,q)}{D(p,q)}$$

where $\left[\prod_K c_K\right]$ indicates the product of the couplings that appear in the diagram, N and D are polinomials in the internal (p) and external (q) momenta. Consider now any sort of "momentum cutoff" regularization of the integrals (such as hard cutoff, Pauli-Villard, etc), with cutoff Λ. The "superficial degree of divergence" D is the highest power of Λ that (based on dimensional analysis) can appear in 1PI for $\Lambda \to \infty$. By definition, we have

$$1PI = \left[\prod_K c_K\right]\left[\Lambda^D + \Lambda^{D-1} E + \Lambda^{D-2} E^2 + \dots + \log\Lambda E^D + \right.$$
$$\left. + \text{"finite"}\right] \qquad (1)$$

where "E" collectively denotes the size of the external momenta q, which are all taken of the same order of magnitude. Remember also that in the above equation only the "UV part" of the 1PI integral is taken into account. This is the region of the d^4p integral in which all the components of p diverge simultaneously. The effects of "overlapping divergences" are ignored (see WEINBERG 12.2 if interested)

From the definition, and dimensional analysis, we immediately see that

$$D = 4 - e - \sum_K \Delta_K$$

Where $\sum_K \Delta_K$ is the total dimension of the vertices in the diagram (there might be repetitions, if one vertex appears several times).

The "divergences" in the 1PI of equation (1), or better the dependence of 1PI on the cutoff Λ, is of a very specific form. The cutoff only appears in a Polynomial in the external momenta q (or E). Terms of this kind are called "local" or "contact terms", because they originate from tree-level insertions of local operators appearing in the Lagrangian. What eq.(1) tells us is that the "divergent" contributions (or, better, the cutoff-dependent contributions) are of the same form of a certain set of local operators, to be added to the Lagrangian. One normally says that, because they are of the same form, the divergences can be cancelled by these operators acting as counterterms.

More physically, one could say that the cutoff Λ (or, more generically, the regulator we use to define the integrals) is an unphysical modification of our original theory. Therefore, whatever depends on Λ is clearly something that cannot be predicted within our theory. What eq. (1) implies is that it is the coefficient of the local operators all what we cannot predict. Their values, or better the ones of the sum of the "bare" term in the Lagrangian and of the cutoff-dependent contributions, must be fixed by experiments. These are the usual renormalization conditions, imposed at a reference scale μ:

$$\left[\prod_k C_k\right]\Lambda^D + C_0 = C_0^\mu$$
$$\left[\prod_k C_k\right]\Lambda^{D-1} + C_1 = C_1^\mu$$
$$\vdots$$

where $C_{0,1\cdots D}$ are the (bare coefficients of the) operators that contribute to 1PI at the tree-level with, respectively, $0, 1, \dots, D$ derivatives:

$$[1PI]_{c.t} = C_0 + EC_1 + \cdots + E^D C_D$$

Independently of the interpretation, we are in any case obliged, to cancel the cutoff-dependence, to include in the theory all the local operators that receive contributions from the cutoff. A priori, these could be all the local operators one can write with the fields that are relevant for the theory under consideration. Actually, these operators are restricted by symmetries. Provided the symmetries are respected by the regulator, the 1PI will respect the corresponding identities. In particular, not all the (POTENTIALLY) divergent terms (ALLOWED BY DIMENSIONAL ANALYSIS) will be compatible with the symmetries, meaning they will not be generated. It is then enough to include only the local operators that are compatible with the assumed symmetries.

This provides the foundation of the modern approach to QFT:

$$\text{QFT} = \text{"Given Field Content and Set of Symmetries"} \Longrightarrow \text{Lagrangian} \Downarrow \text{Predictions}$$

The Lagrangian is a "derived" object, not the central one.

Normally, the Lagrangian contains all the infinite operators compatible with the symmetries. It is only in the special case of renormalizable theories that we can choose to have only a finite number of them. In the renormalizable case one puts all the operators with couplings of dimension $\Delta_K \geq 0$. If only these are present

$$D \leq 4 - e$$

and the divergences only arise in a finite number of 1PI(e); the ones with $e \leq 4$. Moreover, the divergences only contribute to local operators with coefficients c of energy dimension:

$$\Delta_0 = [c_0] = \sum_i \Delta_i + D = 4 - e \geq 0$$

$$\Delta_1 = [c_1] = 4 - e - 1$$

$$\vdots$$

$$\Delta_D = [c_D] = \sum_i \Delta_i \geq 0$$

All the couplings that get renormalized are again the ones of positive dimension, it is therefore self-consistent to include only these.

No one prevents us, of course, to treat all theories on the same footing. Even for "renormalizable theories", i.e. theories for which renormalizable interactions exist, we might always decide to include all the allowed operators. We will discuss in the following the implications of this different viewpoint on theories like QCD, QED, or the Standard Electro-Weak Theory.

But first of all we have to answer a crucial question: how do we deal with this infinite number of operators and of unknown coefficients? Do we really need an infinite number of couplings, to be fixed by an infinite number of measures, in order to make predictions? The answer is no: only a finite number of inputs is needed to perform predictions at a given order in the perturbative expansion. It is only if willing to compute infinite perturbative orders that you would really need infinite inputs. Provided perturbation theory holds, the non-renormalizable Effective Theory provides a well-definite algorithm to make predictions and compare with observations.

How this works is best illustrated with one example: consider the theory of one scalar field ϕ with the following symmetries:

$$\phi \to -\phi$$
$$\phi \to \phi + \varepsilon \quad (\varepsilon = \text{constant})$$

There are an infinite number of operators which are compatible with these symmetries, but all the allowed interactions are non-renormalizable. The Lagrangian takes the form

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + \sum_i \frac{g(n_i, d_i)}{M^{n_i + d_i - 4}} \mathcal{O}^{(n_i, d_i)}$$

where n_i and d_i are, respectively, the number of fields and of derivatives that appear in the interaction. The scale "M" has been factored out in order to make the g's dimensionless, and we will work under the assumption that (the renormalized values of) all these g's are of order unit. The idea is of course that the contributions from operators of high dimension will be suppressed by further powers of $E/M << 1$, so that at low energy it will be enough to make use of a limited number of operators.

Consider a generic 1PI with "e" external states and "V_i" vertices of type "i". The diagram might have any number "L" of loop. We just consider the finite part of such a diagram because the divergences are reabsorbed in the redefinition of the g's. The contribution of this generic diagram to the amplitude is:

$$SA \sim \prod_i \left[\frac{1}{M^{n_i+d_i-4}}\right]^{V_i} E^{\sum_i [n_i+d_i-4]V_i} E^{4-e}$$

$$\equiv \left(\frac{E}{M}\right)^{\nu} E^{4-e}$$

where

$$\nu = \sum_i V_i (d_i + n_i - 4)$$

we obtained the above equation by simply counting the powers of M that enter throught the vertices, and restored the powers of E by dimensional analysis. Remember now the identities:

$$I - L + 1 = \sum_i V_i = \text{"total number of vertices"}$$

which comes from the fact that each vertex leads to one momentum conservation, each propagator

to one $\int d^4p$, and the number of $\int d^4p$ which are left is $L-1$. The other identity is

$$\sum_i V_i n_i = 2I + e$$

which simply states that the lines outgoing from any vertex must end up somewhere. We can then rewrite ν as:

$$\nu = \sum_i V_i d_i + \sum_i V_i n_i - 4I + 4L - 4$$

$$= \sum_i V_i d_i + \sum_i V_i n_i + 2e - 2\sum_i V_i n_i + 4L - 4$$

$$= 2e - 4 + 4L + \sum_i V_i (d_i - n_i)$$

Clearly, the leading contribution to A comes (as long as $E/M \ll 1$) from the <u>smallest power of</u> ν. Because of the symmetry, $d_i \geq n_i$, which makes that all the terms in ν are <u>positive</u>. Therefore, at the leading order, <u>tree-level</u> diagrams ($L=0$) must be considered, and with vertices such that $d_i = n_i$. For a given process of "e" external states, <u>only a finite number of vertices</u> ($n_i \leq e$) have to be considered! If willing to include $(E/M)^2$ corrections, more loops

and more operators have to be considered, but always in a finite number.

Let us see this even more explicitely for the $e=4$ point function. The leading contribution is the tree-level one from the insertion of the operator

$$\frac{g(4,4)}{M^4}\,(\partial_\mu\phi\,\partial^\mu\phi)^2$$

$$\times \;\simeq\; g(4,4)\,\frac{E^4}{M^4}$$

The leading correction comes instead from

$$\frac{g(4,6)}{M^6}\,(\partial_\mu\partial_\nu\phi\,\partial^\mu\partial^\nu\phi)(\partial_\mu\phi\,\partial^\mu\phi)$$

$$\times \;\simeq\; g(4,6)\,\frac{E^6}{M^6}$$

which is subleading by E^2/M^2. At even higher orders, we have the contribution from other operators, plus the loop

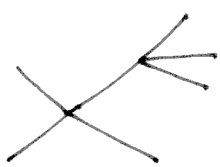
$$\sim \frac{[g(4,4)]^2}{M^8}\,E^8$$

FINITE

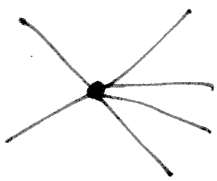
As long as $E/M < 1$, we see that we can compute each scattering amplitude, at each given order in E/M, by only computing a finite number of diagrams which involve a finite number of different interactions, whose coefficients must come from the observations. Our theory is therefore <u>perfectly predictive</u>

It is clear when this program fails: when $E \sim M$. The non-renormalizable theories have, as anticipated, a limited range of validity.

In the example we just discussed, a different set of operators needs to be considered (at a given order) for each 1PI. For instance, the amplitude with "6" external legs receive contributions, already at the <u>leading order</u>, from both the $g(4,4)$ and the $g(6,6)$ couplings:



$$\sim \left(g(4,4)\right)^2 \frac{E^6}{M^8}$$



$$\sim [g(6,6)] \frac{E^6}{M^8}$$

We therefore see that both operators will have to be included if willing to compute the $2 \to 4$ scattering process. Similarly, more operators will be needed for $2 \to 6$ and so on.

It is interesting to observe that the situation is different in the chiral theory of pions. In the chiral theory, as we will now discuss, the leading contribution to any scattering amplitude is always mediated by a single operator (the only one we wrote until now). The single parameter associated to this "leading operator" is F_π, and any pion amplitude can be computed, at the leading order, in terms of this single parameter. The chiral theory is, in this sense, even more predictive than the example we just discussed.

Let us see how this works (WEINBERG vol 2 pg 197-199). The chiral theory is defined by the Goldstone triplet $\vec{\xi}$ which, as you have seen in an exercise, transforms in a

rather complicated way under the symmetry group $SU(2)_L \times SU(2)_R$. But this transformation property is such that

$$\vec{D}_\mu = \frac{\partial_\mu \vec{\pi}/F_\pi}{1 + \vec{\pi}^2/F_\pi^2}$$

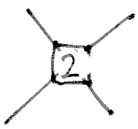
simply transforms as an $SU(2)$ triplet. In our "new" approach to the quantum field theory, we should include in the Lagrangian all the operators compatible with the symmetry. The only one we included until now was the $(\vec{D})^2$ one, which is easily seen to be the one with less derivatives (2 derivatives), some of the others are

$$\mathcal{L}_{eff} = + \frac{F_\pi^2}{2} \vec{D}_\mu \cdot \vec{D}^\mu - \frac{c_4}{4} \left(\vec{D}_\mu \cdot \vec{D}^\mu\right)^2 - \frac{c_4'}{4} \left(\vec{D}_\mu \cdot \vec{D}_\nu\right) \cdot \left(\vec{D}^\mu \cdot \vec{D}^\nu\right)$$

$$= \frac{F_\pi^2}{2} \vec{D}_\mu \cdot \vec{D}^\mu - \sum_i c^{(d_i)} \mathcal{O}_i^{(d_i)}$$

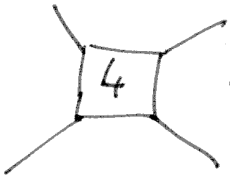
Each term already contains an infinite number of couplings among the pions. The power counting is different from before because now not each power of energy is accounted for by a power of $1/\Lambda$. There is often F_π. Consider first the $2 \to 2$

scattering. From the 2-der term we have, as we have also seen explicitely:



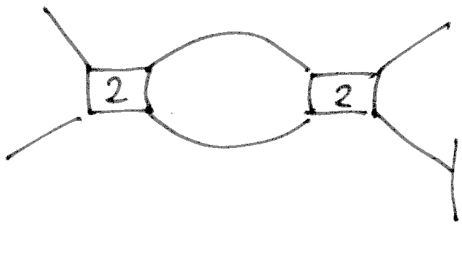
$$\simeq \frac{E^2}{F_\pi^2}$$

From the 4-der one, instead



$$\simeq \frac{C_4}{F_\pi^4} E^4 = \frac{E^2}{F_\pi^2}\left[\frac{C_4 E^2}{F_\pi^2}\right]$$

which provides a correction to the 2-der contribution as long as E is sufficientely small. We also have, at the same order, the 1-loop diagrams of the form



$$\simeq \frac{1}{16\pi^2}\frac{1}{F_\pi^4} E^4 = \frac{E^2}{F_\pi^2}\left[\frac{E}{4\pi F_\pi}\right]^2$$

which is also a small correction provided that

$$E \lesssim 4\pi F_\pi$$

$4\pi F_\pi \equiv \Lambda_{MAX}$ is called the "Maximum Cutoff" of the effective theory. Above Λ_{MAX}, the effective theory loses predictivity because we have that the one-loop correction becomes as large as

the tree-level term. The value of Λ_{MAX} is the "maximum" because the validity of the theory may stop much before. From the condition that the correction from the 4-der term must be small we find

$$E^2 \leq F_\pi^2 / c_4$$

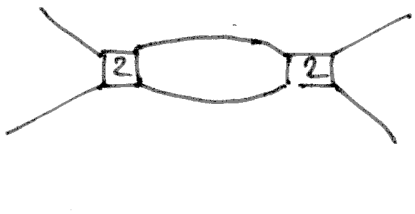
which of course can become a very low upper bound if c_4 happens to be very large! Remember that c_4 is fixed by experiments, we have little control on it. However, what we know is that c_4 must be small enough if the effective theory has to hold until Λ_{MAX}. For this being the case:

$$F_\pi^2 / c_4 \geq \Lambda_{MAX}^2 = 16\pi^2 F_\pi^2$$

$$\Rightarrow \quad \boxed{c_4 \leq \frac{1}{16\pi^2}}$$

There is another argument, that allows us to set the size of c_4. Remember that c_4 is the renormalized coupling, the sum of the bare and of the cutoff-dependent

"divergent" contribution. This is given by



$$= \frac{1}{16\pi^2}\frac{E^2}{F_\pi^2}\frac{\Lambda^2}{F_\pi^2} + \frac{1}{16\pi^2}\frac{E^4}{F_\pi^4}\log\Lambda$$

the first term renormalizes F_π, while the second gives a contribution to C_4 of the form:

$$\delta C_4 \simeq \frac{1}{16\pi^2}\log\Lambda \sim \frac{1}{16\pi^2}$$

$$C_4 = C_4^0 + \delta C_4$$

the "naturalness criterion" states that, in the absence of cancellations that take place in the underlying fundamental theory (QCD, in the present case), we must have that

$$C_4 \gtrsim \delta C_4 = \frac{1}{16\pi^2}$$

The two bounds together give, obviously,

$$C_4 \simeq \frac{1}{16\pi^2}$$

This is the so-called Naive Dimensional Analysis (NDA), an estimate that works pretty well in the QCD theory.

The NDA estimate presented above is not "portable" (with suitable modifications) to a generic effective theory, because the cutoff does not typically coincide with the maximum one. The theories for which this instead happens are, like QCD, the ones in which there is strong coupling. The Naturalness Criterion we have employed to derive not the NDA estimate, but the lower bound on the coupling, is instead quite robust. It can be stated by thinking to some "UV theory" (QCD, in the present case) in which c_4 is perfectly calculable, and finite. In this theory, c_4 will have the form

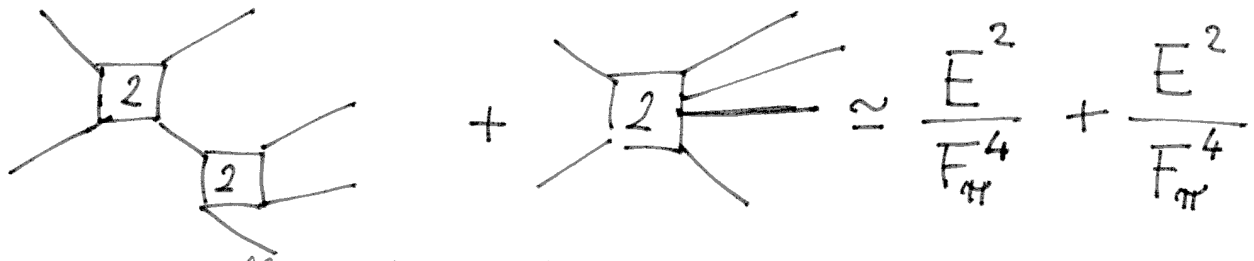
$$c_4 = \int dE \; F(E)$$

in terms of some form factor F that describes the effect of quanta exchanges at energy $E \in [0, \infty]$. If the pion theory correctly describes the evolutions of the relevant (pion) quanta for $E < \Lambda$, the form-factor F of the "true theory" must coincide with the one computed in the effective one. We therefore write:

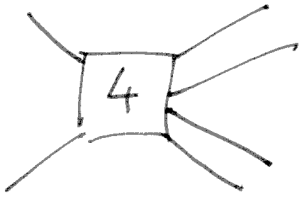
$$c_4 = c_4^{IR} + c_4^{UV} = \int^{\Lambda} dE F(E) + \int_{\Lambda} dE F(E)$$

The naturalness criterion states that c_{UV} and c_{IR} are unlikely to be of almost equal size and of opposite size. It is "unnatural" that c^{UV} and c^{IR} nearly cancel each other to give a very small total result for c_4. For many people, this criterion should always hold, This is related with the famous Hierarchy problem we will discuss in the next lesson.

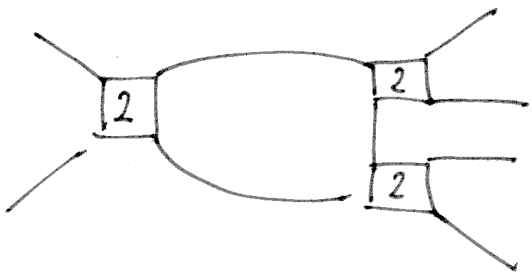
Let us go back to the issue of calculability, and consider now the 6-point function. The tree-level contribution from dim 2 es:



where 2 different contributions of the same order come from inserting 2 quartic vertices or a one with 6 legs from expanding the dim-2 operator. We also have dim-4



and loops such as, say



$$\simeq \frac{1}{16\pi^2}\frac{1}{F_\pi^6}E^2$$

we find that, as anticipated, tree-level diagrams with the dim-2 coupling provide the leading contribution to $2\to 4$ as well.

This is a general fact: consider an amplitude with V_d vertices, L loops and I external states. It scales with the energy as

$$E^\nu \quad ; \quad \nu = \sum_d V_d d_d + 4L - 2I$$

where I am simply counting the powers of E that come from loops and vertices and propagators. But since

$$L = 1 + I - \sum_d V_d$$

$$\nu = +2 \quad -2L - 2\sum_d V_d + 4L + \sum_d V_d d_d$$

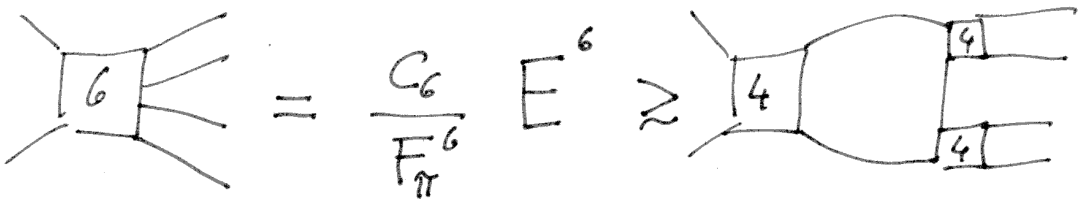
$$= 2 + 2L + \sum_d V_d (d_d - 2)$$

All the terms in the sum are positive. Zero loop diagrams with $d_d = 2$ vertices are the dominant ones for any process.

Corrections of order E^2/Λ^2_{MAX} to any scattering amplitude (4, 6, etc..) originate from one loop diagrams with the 2-der couplings and from tree-level insertions of the dim-4 operators. These effects have been computed in the chiral theory, and successfully compared with observations, providing a validation of the approach we are following. Smaller effects, of order E^4/Λ^4_{MAX}, are difficult to compute and to measure, but it is nevertheless important to know that they come from:

1) $L=2$ diagrams with 2-der couplings
2) $L=1$ diagrams with one insertion of the 4-der couplings
3) $L=0$ insertion of one 6-der operator

The NDA size of the 6-der operator is easily estimated, similarly to what we did for the 4-der one:



$$= \frac{C_6}{F_\pi^6} E^6 \gtrsim \Big|_{div} = \frac{1}{16\pi^2}\left[\frac{1}{16\pi^2}\right]^3 \frac{1}{F_\pi^{12}} E^6 \Lambda^4_{MAX} + \dots$$

$$\| \; C_4^3$$

where the inequality becomes equality in a strongly-coupled theory like QCD:

$$c_6 = \left[\frac{1}{16\pi^2}\right]^4 \frac{\Lambda_{MAX}^4}{F_\pi^6} = \frac{1}{16\pi^2}\,\frac{1}{\Lambda_{MAX}^2}$$

Similarly, the NDA estimate of all the terms is:

$$\mathcal{L}_{eff} = \frac{\Lambda_{MAX}^4}{16\pi^2}\left[\frac{1}{2}\,\frac{\vec{D}_\mu \cdot \vec{D}^\mu}{\Lambda_{MAX}^2} - \sum_i \frac{c_i}{\Lambda_{MAX}^{d_i}}\,\mathcal{O}_i^{(d_i)}\right]$$

where $\Lambda_{MAX} = 4\pi F_\pi$ and c_i = "$\mathcal{O}(1)$ coefficients"