In quantum field theory (QFT) we often need to deal with functionals, the most common of which is probably the generating functional, $Z[J]$. This being physics, however, we are also frequently interested in analyzing the symmetries of a system and extracting information from the demand that a theory is invariant under some transformation. Indeed, this is essentially the idea behind Noether's theorem.

But as we analyze symmetries, at the very least the language we use can become a bit muddled. Since symmetries are so important, it's worth taking some time to think carefully about what we are doing, and see if we can't perform the same analyses in more controlled environments in which we are able to compute everything we're interested in directly instead of by exploiting symmetry.

Contents

1 Relations Between Functions

Suppose we have some function $f : \mathbb{R} \to \mathbb{R}$ and transformation $T_\alpha : \mathbb{R} \to \mathbb{R}$ parametrized by α. We will choose the parametrization of T such that T_0 is the identity transformation and we will suppose that f satisfies $f(T_\alpha(x)) = f(x) + \Delta f(x, \alpha)$ for all x and all[1] α. Here $\Delta f(x, \alpha)$ is just some known function.

Since this relation holds $\forall \alpha, x$, we are free to take derivatives w.r.t. either x or α:

$$f'(T_\alpha(x))\frac{\partial T_\alpha}{\partial \alpha} = \frac{\partial}{\partial \alpha}\Delta f(x, \alpha), \qquad f'(T_\alpha(x))\frac{\partial T_\alpha}{\partial x} = f'(x) + \frac{\partial}{\partial x}\Delta f(x, \alpha). \tag{1.1}$$

Evaluating at $\alpha = 0$ so T is the identity, the first of these relations becomes

$$f'(x)\frac{\partial}{\partial \alpha}\bigg|_0 T_\alpha(x) = \frac{\partial}{\partial \alpha}\bigg|_0 \Delta f(x, \alpha). \tag{1.2}$$

This equation gives us a differential equation satisfied by any function which obeys the relation $f(T_a(x)) = f(x) + \Delta f(x, \alpha)$. In the case $\Delta f(x, \alpha) = 0$, we would say that this is a differential equation for functions invariant under the transformation T_α, at least for α sufficiently close to zero.

Consider for example $T_\alpha(x) = x + \alpha$ and $\Delta f(x, \alpha) = 2\alpha x + \alpha^2$. With these, (1.2) implies the differential equation

$$f'(x) = 2x \tag{1.3}$$

which we know is solved by $f(x) = x^2 + C$ for any constant C. We can also check that this function indeed satisfies $f(x+\alpha) = f(x) + (2\alpha x + \alpha^2)$. Though it should be noted that the α being sufficiently close to zero is actually very important here. For example, if we had chosen $\Delta f(x, \alpha) = 2\alpha x + 10\alpha^2$, (1.2) would have produced precisely the same differential equation but the original equation $f(T_\alpha(x)) = f(x) + \Delta f(x, \alpha)$ certainly wouldn't be satisfied.

This brings us to an interesting point. Though (1.2) is easy to write down and is certainly a necessary condition, it is not a sufficient one. Instead, we would need to go back to (1.1) and evaluate the first equation at $x = T_\alpha^{-1}(y)$:

$$f'(y)\frac{\partial T_\alpha}{\partial \alpha}\bigg|_{T_\alpha^{-1}(y)} = \frac{\partial}{\partial \alpha}\Delta f(T_\alpha^{-1}(y), \alpha) \tag{1.4}$$

where the α derivative on the RHS is understood to act only on the second argument of Δf. Writing this out explicitly for the example we had above, we find

$$f'(x) = 2x + 18\alpha \tag{1.5}$$

[1] It's often good enough that f satisfies this relation for only α in an open neighborhood of zero and similarly for x about some particular value.

which clearly has solution

$$f(x) = x^2 + 18\alpha x + C \tag{1.6}$$

for some constant C. However, there is an important auxiliary assumption we used without mentioning: $\frac{\partial}{\partial\alpha} f(x) = 0$. The function we are looking for cannot depend on the transformation we apply to it. The power of (1.4) is that we may now state with confidence that does not exist a function satisfying $f(x+\alpha) = f(x) + 2\alpha x + 10\alpha^2$.

Everything we have said thus far generalizes immediately to higher dimensional examples: if $f : \mathbb{R}^n \to \mathbb{R}^m$ satisfies $f^a(T_\alpha(\boldsymbol{x})) = f^a(\boldsymbol{x}) + \Delta f^a(\alpha, \boldsymbol{x})$ for $a = 1, \ldots m$, then it follows that

$$\frac{\partial T^\mu_\alpha}{\partial\alpha} \frac{\partial}{\partial x^\mu}\bigg|_{T_\alpha(\boldsymbol{x})} f^a = \frac{\partial}{\partial\alpha} \Delta f^a(\boldsymbol{x}, \alpha). \tag{1.7}$$

Hence,

$$\frac{\partial T^\mu_\alpha}{\partial\alpha}\bigg|_{T^{-1}_\alpha(\boldsymbol{x})} \frac{\partial}{\partial x^\mu} f^a(\boldsymbol{x}) = \frac{\partial}{\partial\alpha} \Delta f^a(T^{-1}_\alpha(\boldsymbol{x}), \alpha). \tag{1.8}$$

Indeed, if we define

$$v_\alpha = \frac{\partial T^\mu_\alpha}{\partial\alpha}\bigg|_{T^{-1}_\alpha(\boldsymbol{x})} \frac{\partial}{\partial x^\mu}, \qquad Q^a_\alpha = \frac{\partial}{\partial\alpha} \Delta f^a(T^{-1}_\alpha(\boldsymbol{x}), \alpha) \tag{1.9}$$

then the variance condition may be written

$$v_\alpha(f^a) = Q^a_\alpha, \quad \text{subject to} \quad \frac{\partial}{\partial\alpha} f(x) = 0, \tag{1.10}$$

so we have, in fact, derived in the process the generator of the transformation T_α and in the case of invariance, so $\Delta f = 0$, we see that the generator of the transformation T_α must annihilate the invariant function.

An important point to note here is that (1.10) remains valid for all α and x, meaning we could continue to differentiate it with respect to either α or x to obtain additional identities satisfied by the function f.

Finally, we note that, though the second equation of (1.1) is satisfied by f, it is generally less useful because we are unable to isolate $f'(x)$ unless we take $\alpha = 0$, but when we do this the identity only tells us that $\frac{\partial}{\partial x} \Delta f(x, 0) = 0$ which is already implied by the identity $\Delta f(x, 0) = 0$. Since this equation does not rely on the continuity of α, however, there stands some hope that it might tell us something about discrete symmetries which are notoriously difficult to analyze.

2 Integral Identities

Here we will specialize to a particular kind of function f, namely those which are given by an integral whose bounds or integrand may depend on some parameters which we will continue to call x^μ (these are *not* necessarily valued in the same space we are integrating over). That is[2],

$$f(\boldsymbol{x}) = \int_\Sigma \mathcal{L}(\boldsymbol{z}, \boldsymbol{x}) \mathrm{d}z. \tag{2.1}$$

Our general analysis still holds, so if this function obeys $f(T_\alpha(\boldsymbol{x})) = f(\boldsymbol{x}) + \Delta f(\boldsymbol{x}, \alpha)$, then we may conclude (1.10). It is the common use of integral expressions as generating functions which makes considering them specifically worthwhile. In fact, this is the same reason they tend to appear in physics as well.

First, however, let's consider the example[3] $f(x) = \int_0^\infty e^{-xz} \mathrm{d}z$. We are able to integrate this function explicitly to find $f(x) = \frac{1}{x}$ and we can see that this function generates the integrals of polynomials times the exponential (up to a sign):

$$\begin{aligned} f(1) &= \int_0^\infty e^{-z} \mathrm{d}z, \qquad & f'(1) &= -\int_0^\infty z e^{-z} \mathrm{d}z, \qquad & f''(1) &= \int_0^\infty z^2 e^{-z} \mathrm{d}z, \\ f^{(3)}(1) &= -\int_0^\infty z^3 e^{-z} \mathrm{d}z & &\cdots & f^{(n)}(1) &= (-1)^n \int_0^\infty z^n e^{-z}. \end{aligned} \tag{2.2}$$

Now, since we already know $f(x) = \frac{1}{x}$, we may directly compute all of these integrals by taking sufficiently many derivatives. But generically, integrals are very hard to do and we may not know how to compute $f(x)$ directly. Instead, it may well be good enough that we find a set of identities relating these integrals to each other. This way, we might be able to say something interesting without actually needing to have the explicit expression for any of these integrals.

Let's see what a simple identity on f might be able to buy us. Consider a rescaling of the parameter x:

$$f(\lambda x) = \int_0^\infty e^{-\lambda x z} \mathrm{d}z = \frac{1}{\lambda} \int_0^\infty e^{-xz} \mathrm{d}z = \frac{1}{\lambda} f(x) = f(x) + \left(\frac{1}{\lambda} - 1 \right) f(x). \tag{2.3}$$

[2] For now we take the integration region Σ to be a constant with respect to the parameters x^μ, but this could easily be generalized via the generalization of Leibniz's rule.

[3] Though it is typical to call the parameter of this generating function α and the integration variable x, we abandon this notation to remain consistent with the notation used elsewhere in this document.

Writing[4] $\lambda = 1+\alpha$, we see that $T_\alpha(x) = (1+\alpha)x = x + \alpha x$. Furthermore, this tells us that

$$\Delta f(x,\alpha) = \left(\frac{1}{1+\alpha} - 1\right) f(x) = -\frac{\alpha}{1+\alpha} f(x). \tag{2.4}$$

Since our goal is to apply (1.10), we will first compute

$$\frac{\partial}{\partial \alpha}\Delta f(x,\alpha) = -\frac{1}{(1+\alpha)^2} f(x) \tag{2.5}$$

and so

$$Q_\alpha = \frac{\partial}{\partial \alpha}\Delta f(T_\alpha^{-1}(x),\alpha) = \frac{-1}{(1+\alpha)^2} f\left(\frac{1}{\lambda}x\right) = \frac{-\lambda}{(1+\alpha)^2} f(x) = \frac{-1}{1+\alpha} f(x) \tag{2.6}$$

where we have used again the identity satisfied by $f(x)$. Next, we compute

$$\frac{\partial T_\alpha}{\partial \alpha} = x \tag{2.7}$$

so

$$\left.\frac{\partial T_\alpha}{\partial \alpha}\right|_{T^{-1}(x)} = \frac{1}{\lambda}x = \frac{x}{1+\alpha}. \tag{2.8}$$

Now, our identity tells us that

$$\frac{x}{1+\alpha} f'(x) = \frac{-1}{1+\alpha} f(x). \tag{2.9}$$

The α dependence cancels nicely (so we don't need to worry about the $\frac{\partial}{\partial \alpha} f = 0$ condition) and we are left with the identity

$$x f'(x) = -f(x). \tag{2.10}$$

Unfortunately, there being no more α dependence means we cannot see how additional α derivatives would play with the identity, but we can start taking x derivatives. Applying an arbitrary number of derivatives[5], we would find the identity

$$n f^{(n)}(x) + x f^{(n+1)} = -f^{(n)}(x) \implies (n+1) f^{(n)}(1) + f^{(n+1)}(1) = 0 \tag{2.11}$$

[4] We shift the transformation parameter so the transformation can be the identity when the parameter is zero.

[5] At most one derivative can hit the x, the rest must hit the $f(x)$. Out of n derivatives applied, there are n ways to choose which of those derivatives should hit the x.

which holds for all integers $n > 0$. So, without needing to know anything about the values of the integrals (2.2), we have found a collection of relationships between them. In this particular case, the relationship happens to be a recurrence relation giving the $(n+1)^{\text{st}}$ of these integrals in terms of the n^{th} and the recurrence happens to be simple enough to solve by $f^{(n)}(1) = n! f(1)$ so if we could only compute $f(1)$, we could immediately know the values of all the other integrals. And, in fact, $f(1) = 1$, which if we recall that the Gamma function was given by $\Gamma[n+1] = \int_0^\infty z^n e^{-z} \mathrm{d}z = n!$ should be obvious in retrospect.

We could also consider an example in which the integral is actually invariant under the changes in the parameter as well so $\Delta f = 0$. A simple way to generate such examples in integrals would be to consider a change of variables that includes a parameter. For example, consider

$$f(x) = \int_{-\infty}^{\infty} e^{-z^2} \mathrm{d}z. \tag{2.12}$$

This function clearly does not depend on x since x does not actually appear in the integral. But if we apply a change of variables, so for example, $z' = z - x$, then

$$f(x) = \int_{-\infty}^{\infty} e^{-(z+x)^2} \mathrm{d}z. \tag{2.13}$$

This function is invariant under any transformation of x that we care to write down, but we can consider as an example $T_\alpha(x) = x + \alpha$. Under this transformation, we clearly have $f(T_\alpha(x)) = f(x)$ so $\Delta f = 0$.

For this example,

$$\frac{\partial T_\alpha}{\partial \alpha} = 1 \implies \left.\frac{\partial T_\alpha}{\partial \alpha}\right|_{T_\alpha^{-1}(x)} = 1. \tag{2.14}$$

Hence, we find that (1.10) tells us

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = 0 \tag{2.15}$$

which we probably could have guessed, but which we can see also follows nicely from the general framework we have built. As with the previous example, we note that this identity implies relations between many integrals. If we were to define the integrals

$$I_0 = \int_{-\infty}^{\infty} e^{-z^2} \mathrm{d}z, \quad I_1 = \int_{-\infty}^{\infty} z e^{-z^2} \mathrm{d}z, \quad \ldots \quad I_n = \int_{-\infty}^{\infty} z^n e^{-z^2} \mathrm{d}z, \tag{2.16}$$

then we may note that

$$f(0) = I_0, \quad f'(0) = -2I_1, \quad f''(0) = -2I_0 + 4I_2, \quad f^{(3)}(0) = 12I_1 - 8I_3, \quad \ldots \tag{2.17}$$

As with the previous example, the identities between these integrals allow us to deduce information about them without needing to know explicitly the values of any of them. For example, we can already see that I_1 and I_3 are zero, which we could have also noted from the asymmetry of their integrands.

A key take-away from these two examples is that it does not even matter whether the parameters $\boldsymbol{x}$ are real parameters in the integral or simply artifacts of a bad coordinate transformation. We can obtain identities in either case.

3 Generalization to Functionals

Here we will derive the generalization of (1.10) to functionals. That is, functions $F[\phi]$ which take as input some functions ϕ rather than just some set of parameters $\boldsymbol{x}$. Formally, very little changes, but it will be a useful exercise to derive the master identity generalizing (1.10) explicitly.

We consider here a transformation $T_\alpha[\phi]$ which maps functions ϕ into some other functions and we again assume that both $\alpha \in \mathbb{R}$ and $T_0 = id$. Suppose now that we have a functional[6] $F[\phi]$ which obeys $F[T_\alpha[\phi]] = F[\phi] + \Delta F[\phi; \alpha]$ for some known functional $\Delta F[\phi; \alpha]$.

We begin as we did before by differentiating with respect to α and applying the functional chain rule[7]:

$$\int \mathrm{d}x \frac{\partial T_\alpha(x)}{\partial \alpha}\bigg|_\phi \frac{\delta}{\delta\phi(x)}\bigg|_{T_\alpha[\phi]} F = \frac{\partial}{\partial \alpha}\Delta F[\phi; \alpha]. \tag{3.1}$$

Evaluating at $T_\alpha^{-1}[\phi]$, we now find

$$\int \mathrm{d}x \frac{\partial T_\alpha(x)}{\partial \alpha}\bigg|_{T_\alpha^{-1}[\phi]} \frac{\delta}{\delta\phi(x)} F[\phi] = \frac{\partial}{\partial \alpha}\Delta F[T_\alpha^{-1}[\phi]; \alpha]. \tag{3.2}$$

Defining

$$v_\alpha = \int \mathrm{d}x \frac{\partial T_\alpha(x)}{\partial \alpha}\bigg|_{T_\alpha^{-1}[\phi]} \frac{\delta}{\delta\phi(x)}, \qquad Q_\alpha = \frac{\partial}{\partial \alpha}\Delta F[T_\alpha^{-1}[\phi]; \alpha] \tag{3.3}$$

the above condition may be written in a form formally identical to (1.10):

$$v_\alpha(F) = Q_\alpha, \qquad \frac{\partial}{\partial \alpha} F = 0. \tag{3.4}$$

[6] The functional F could have some internal indices as well, but unless they also transform under the applied trasformation T, they will only be carried around in the background like the index a in (1.10). If the indices of F do transform under T, this will only equate to a contribution to ΔF.

[7] The variable $\boldsymbol{x}$ will now be returned to its status as a coordinate, which we may also think of as taking the place of the index μ in the mutlivariate case from earlier.

4 Noether's Theorem

To state and prove Noether's theorem properly in the language we have developed here, we should first set up exactly how we are thinking about mechanics. Once we do so, we will find the Noether's theorem follows quite easily from (3.4).

In general, we assume that we have some action functional $S[\phi]$. We will assume that the fields ϕ are valued in some configuration space Γ where Γ is the space of sections over some bundle whose base space is space-time. However, there may be reasons why we would like to restrict from the space of all sections down to a smaller collection $C \subseteq \Gamma$. For example, C may be the space of sections which satisfy some boundary conditions. We will suppose that there exists a good differential operator δ on Γ and that $\mathfrak{X}(\Gamma)$ is the space of vector fields over Γ with $\mathfrak{X}(C)$ similarly defined[8].

With this notion of a differential, we are free to compute the differential of the action, $\delta S[\phi]$. Next, let $\eta \in \mathfrak{X}(C)$ be arbitrary. There exists a unique decomposition $\delta S[\phi] = \delta S_\eta[\phi] + \delta S_K[\phi]$ such that $\delta S_K[\phi]$ is in the kernel of ι_η for all η and ϕ, that is $\iota_\eta \delta S_K[\phi] = 0$ $\forall \eta \in \mathfrak{X}(C), \phi \in \Gamma$, and if $\iota_\eta \delta S_\eta[\phi_s] = 0$ $\forall \eta \in \mathfrak{X}(C)$ and fixed $\phi_s \in C$, then $\delta S_\eta[\phi_s] = 0$.

In this language, the equations of motion take the form $\iota_\eta \delta S[\phi] = 0$ for all $\eta \in \mathfrak{X}(C)$. Solutions ϕ_s of this functional equation clearly imply based on the above discussion that $\iota_\eta \delta S_\eta[\phi_s] = 0$ $\forall \eta$. Now, if T_α is some transformation on the configuration space and the action satisfies $S[T_\alpha[\phi]] = S[\phi] + \Delta S[\phi; \alpha]$, then based on our general discussion in the previous section, S must satisfy (3.4). But if we rewrite the action of v_α on S as a contraction of v_α with δS, then we find on the equations of motion

$$Q_\alpha[\phi_s] = v_\alpha(S)[\phi_s] = \iota_v \delta S[\phi_s] = \iota_v \delta S_K[\phi_s]. \tag{4.1}$$

This is the content of Noether's theorem.

5 Special Cases of Noether's Theorem

Let's consider first the special case where the action is a local functional depending on only the value of the fields and their first derivatives, $S[\phi] = \int \mathrm{d}x \mathcal{L}(\phi, \partial\phi)$. We will take as

[8] Strictly speaking, I think we really need this to be the space of vector fields over Γ whose flows do not leave C, but I imagine this probably follows from a set of boundary conditions being well-posed.

definition that T_α represents a symmetry if $\Delta S[\phi;\alpha]$ is the integral of a total derivative. And, if $\Delta S[\phi;\alpha]$ is the integral of a total derivative, then so is Q_0. Say, $Q_0 = \int \mathrm{d}x \partial_\mu K^\mu$ for some K^μ.

Next, note that the variation of this action is given by

$$\delta S[\phi] = \int \mathrm{d}x \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right] \delta\phi + \int \mathrm{d}x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi\right). \tag{5.1}$$

Assuming we take as boundary conditions that the fields ϕ fall off "sufficiently fast" at infinity, thus specifying C, we see that the two terms we have written down already factor the differential into the sum δS_η and δS_K, the total derivative giving us δS_K.

If we now take ϕ_s to be a solution to the equations of motion, then Noether's theorem tells us

$$\iota_v \int \mathrm{d}x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi\right) = \int \mathrm{d}x \partial_\mu K^\mu. \tag{5.2}$$

Writing $T_\alpha[\phi](y) = \phi(y) + \alpha t[\phi](y)$, we see that the vector field v_0 becomes

$$v_0 = \int \mathrm{d}y t[\phi](y) \frac{\delta}{\delta\phi(y)} \tag{5.3}$$

and so

$$\int \mathrm{d}x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} t[\phi]\right) = \int \mathrm{d}x \partial_\mu K^\mu. \tag{5.4}$$

But since no properties of the integration domain where used, we must also have equality of the integrands. Thus,

$$\partial_\mu J^\mu \equiv \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} t[\phi] - K^\mu\right) = 0 \tag{5.5}$$

which is the standard statement of Noether's theorem.

In the special case in which all of the fields ϕ are scalar-valued and the transformation T is a result of a coordinate transformation generated by some vector field ξ, then $t[\phi] = \mathcal{L}_\xi \phi = -\xi^\mu \partial_\mu \phi$ where, due to an unfortunate confluence of notation, $\mathcal{L}_\xi$ is the Lie derivative w.r.t. ξ. Then

$$\begin{aligned}
\Delta S[\phi;\alpha] &= -\alpha \int \mathrm{d}x \left(\frac{\partial \mathcal{L}}{\partial \phi} \xi^\mu \partial_\mu \phi + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \partial_\nu (\xi^\mu \partial_\mu \phi)\right) \\
&= -\alpha \int \mathrm{d}x \left(\xi^\mu \partial_\mu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \partial_\nu \xi^\mu \partial_\mu \phi\right) \\
&= -\alpha \int \mathrm{d}x \left[\partial_\mu (\xi^\mu \mathcal{L}) + \left(\partial_\nu \xi^\mu \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \partial_\mu \phi - \partial_\mu \xi^\mu \mathcal{L}\right)\right]
\end{aligned} \tag{5.6}$$

where we have assumed that $\mathcal{L}$ does not depend directly on the coordinates. If we assume that ξ^μ is a constant, then we find that

$$K^\mu = -\xi^\mu \mathcal{L} \tag{5.7}$$

and so Noether's theorem implies

$$J^\mu = -\xi^\nu T^\mu_\nu = -\xi^\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \delta^\mu_\nu \mathcal{L} \right) \tag{5.8}$$

is conserved.

If we did not assume ξ^μ to be constant, then we would not have found ΔS to have been given by a total derivative. However, by sufficient application of the product rule, (4.1) would imply $\xi^\mu \partial_\nu T^\nu_\mu = 0$ just the same.

6 Why Should Any of This Work?

Let's take a moment to think carefully about why Noether's theorem works. The key uncertainty in all of Noether's theorem is how we actually calculate this function K^μ. The most obvious method, which is also the standard method, is to take $S[T_\alpha[\phi]]$ and simply expand to first order in α. But then $\int \mathrm{d}x \partial_\mu K^\mu$ is nothing more than the first α derivative of $S[T_\alpha[\phi]]$ evaluated at $\alpha = 0$. The on-shell variation of the action contracted on the particular vector field corresponding to this transformation, namely the surface term of (5.1), is also precisely the derivative of the action with respect to α after evaluating on-shell.

So, we might ask ourselves why K^μ is in any way distinct from $\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi$ when taken on-shell. They represent the same quantity after all. And if they are the same, how is it possible for Noeher's theorem to tell us anything new? As a matter of fact, we could even take $\partial_\mu T^\mu_\nu$, assume the form $T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \delta^\mu_\nu \mathcal{L}$, and show that $\partial_\mu T^\mu_\nu = 0$ using only the Euler-Lagrange equations. No need to check variations. So, the statement $\partial_\mu T^\mu_\nu = 0$ is rendered trivially true on-shell.

The trick of the answer is actually to note that Noether's theorem only holds on-shell, but is also trivial on-shell, and that this is okay. The real key of Noether's theorem is that, despite only holding on-shell, it's utility is entirely off-shell. Or, perhaps better say, its utility is to tell us something about the shell.

We should think about things as follows. There is some configuration space for the fields ϕ, for example the space C of Section 4. There is some subspace $\tilde{C} \subseteq C$ of configurations for which the equations of motion are satisfied. Again related to previous notation, we would say that $\forall \phi_s \in \tilde{C}$, $\iota_\eta \delta S[\phi_s] = 0$ $\forall \eta \in \mathfrak{X}(C)$. In any case, we know that $\forall \phi \in \tilde{C}$, $\partial_\mu J^\mu \big|_\phi = 0$

trivially. The power then comes when we don't already know anything about the properties[9] of $\tilde{C}$. We are always free to consider the value of $\partial_\mu J^\mu$ for any configuration in C. And, in fact, we might define the subset $C_J = \{\phi \in C | \partial_\mu J^\mu\big|_\phi = 0\}$. Noether's theorem guarantees that $\tilde{C} \subseteq C_J$. Hence, a Noetherian conservation law tells us something about the on-shell configuration space without relying directly on the equations of motion.

This is why application of Noether's theorem do not always result in the conservation laws being trivially true: we consider the value of $\partial_\mu J^\mu$ *off-shell* and demand that it be zero as a *constraint* on what configurations could possibly be in $\tilde{C}$.

[9] We of course know that all elements of $\tilde{C}$ satisfy the equations of motion by definition, and of course all properties of $\tilde{C}$ are therefore implied by the equations of motion. However, this does not mean all the properties of $\tilde{C}$ are determinable by human means from the equations of motion alone.