On the Commutator Form of the Riemann Curvature Tensor

Aki Anandam

August 2026

1 Introduction

The Riemann curvature tensor prescribes how a vector changes as one parallel transports it around a parallelogram. In Euclidean space, the vector doesn't change. However, in more complex spaces, such as Spherical geometry, the vector can change, which is referred to as holonomy.

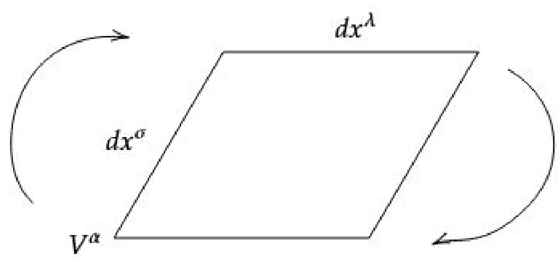

Figure 1: The vector V^α being parallel transported clockwise around a parallelogram in the $x^\sigma x^\lambda$ plane.

Figure 1 depicts the vector V^α being parallel transported around a parallelogram. It first moves positively in the dx^σ direction, then positively in the dx^λ direction, then negatively in the dx^σ direction, then negatively in the dx^λ. *

Call the change V^α undergoes δV^α. The Riemann tensor then relates δV^α to V^α:

$$\delta V^\alpha = R^\alpha_{\mu\lambda\sigma} dx^\lambda dx^\sigma V^\mu, \tag{1}$$

where

$$R^\alpha_{\mu\lambda\sigma} = \partial_\lambda \Gamma^\alpha_{\sigma\mu} - \partial_\sigma \Gamma^\alpha_{\lambda\mu} + \Gamma^\alpha_{\lambda\nu}\Gamma^\nu_{\sigma\mu} - \Gamma^\alpha_{\sigma\nu}\Gamma^\nu_{\lambda\mu}. \tag{2}$$

*This article uses the notation of https://web.mit.edu/sahughes/www/8.962/lec10.pdf and builds off the derivation presented.

Equation (2) is somewhat complicated, though this may not be so surprising as we are analyzing a complex process, and differential geometry is full of several complicated expressions for quantities with simple definitions.† However, what is surprising is that

$$R^{\alpha}_{\mu\lambda\sigma}V^{\mu} = [\nabla_{\lambda}, \nabla_{\sigma}]V^{\alpha}. \tag{3}$$

Many derivations of equation (2) do not make such a result so clear. For example, [1] derives equation (2) by analyzing how the components of V^{α} change as one goes around the parallelogram, adhering to the parallel transport rule. The covariant derivative, however, concerns the entire vector rather than just the components. In addition, if one simply expands the covariant derivative, they will find that the commutator on the right side of equation (3) only depends on the value of V^{α} at the designated point and not derivatives of V^{α}, indeed matching the Christoffel terms in equation (2). However, it is important to note the difference between what we mean by V^{α} in equation (3) and how we use V^{α} when transporting it around a parallelogram. Specifically, when transporting V^{α} around a parallelogram, we are moving a vector, and if we assign the value of that vector to each point along the journey, those values will not form a smooth vector field. This is observed by the fact that the value of V^{α} at the end of the journey will not match the value at the beginning of the journey, which creates a conflict because smooth vector fields can only assign one value to V^{α} at a particular point. By contrast, when we refer to V^{α} on the right side of equation (3), we are referring to a smooth vector field (it can actually be arbitrary so long as it has the desired value at the starting point, since derivatives of V^{α} ultimately will not matter in the curvature expression). Importantly, it is *not* the value of the vector that is parallel transported at different points in the manifold. If somehow that were the case, there would be an issue, since for the transported vector we require $\nabla_{\sigma}V^{\alpha} = 0$ along the first and third leg, which implies $\nabla_{\lambda}\nabla_{\sigma}V^{\alpha} = 0$ (the curvature contributions are zero since $\nabla_{\sigma}V^{\alpha}$ is zero). We can make the same argument for $\nabla_{\sigma}\nabla_{\lambda}V^{\alpha}$. This would suggest that the right side of equation (3) is always zero. However, this is not the case. Indeed, the notion of $\nabla_{\lambda}\nabla_{\sigma}V^{\alpha}$ is not well defined for the vector that is transported around the parallelogram, since it does not form a smooth vector field. Rather, in equation (3), we use an arbitrary smooth vector field that has the appropriate value at our starting point. A vector whose components are prescribed by this field would not experience the effect that going around the parallelogram results in a different final value than the initial value. However, as a consequence, we no longer can ensure that the vector is parallel transported along all four of the legs.

Equation (3) then unites two different ways of looking at curvature. First, if we fix the covariant derivatives by requiring that the vector be parallel transported along all four legs, then curvature requires that the end components V^{α} will not line up with the initial components. However, if we instead fix the vector field by ensuring that the components line up and there is only one

†See https://akianandam.com/articles/christoffel.pdf for another example.

vector at each point on the manifold, then the covariant derivatives must be adjusted and may no longer satisfy the parallel transport condition for all four legs. Specifically, if we want two things:

1. $\nabla_\mu V^\alpha = 0$ for all four legs.
2. The components V^α form a smooth vector field around the parallelogram, such that a vector moving around the parallelogram, whose components are given by the vector field, will have the same final components as initial components (i.e. $\delta V^a = 0$).

Curvature requires that we can pick one of the two. Whichever one we choose to forgo, the discrepancy of not satisfying the condition (either $\delta V^\alpha \neq 0$ in the first case or $[\nabla_\lambda, \nabla_\sigma] \neq 0$ in the second) will give us the value of the curvature. The question then arises: why do both these different discrepancies give the value of curvature?

This will be explained in the following section.

2 Why the two discrepancies are getting at the same thing

Fundamentally, the reason why the inability of the components to line up is related to the inability of the covariant derivatives to line up comes down to the equation for the covariant derivative:

$$\nabla_\mu V^\alpha = \partial_\mu V^\alpha + \Gamma^\mu_{\alpha\gamma} V^\gamma. \tag{4}$$

The Christoffel symbols are the same regardless of whether we choose option 1 or option 2. Deep down, the Christoffel symbols are what encode curvature. This is made manifest by equation (2) which only depends on the Christoffel symbols. If we fix the Christoffel symbols in place, we see in equation (4) that the covariant derivative and the change in the components of V^α are related one-to-one. We will now show clearly the equivalence we seek. Specifically, that

$$R^\alpha_{\mu\lambda\sigma} dx^\lambda dx^\sigma V^\mu = \delta V^\alpha = dx^\lambda dx^\sigma [\nabla_\lambda, \nabla_\sigma] V^\alpha. \tag{5}$$

To start, we first establish that to leading order, when calculating δV^α, it makes no difference whether we measure the change in components at the bottom left or top right corner. Originally, we measure the change in components at the bottom left corner, adding up the changes the vector undergoes from the top left to top right to bottom right to bottom left corners. However, we have another option. We could take one vector, say V_1^α, and transport it to the top left corner and then to the top right corner. We then take another vector, V_2^α, and transport it to the bottom right corner and then to the top right corner.

Figure 2 depicts the accumulations each vector undergoes, with blue illustrating the legs V_1^α traverses and red indicating the legs V_2^α traverses. We then

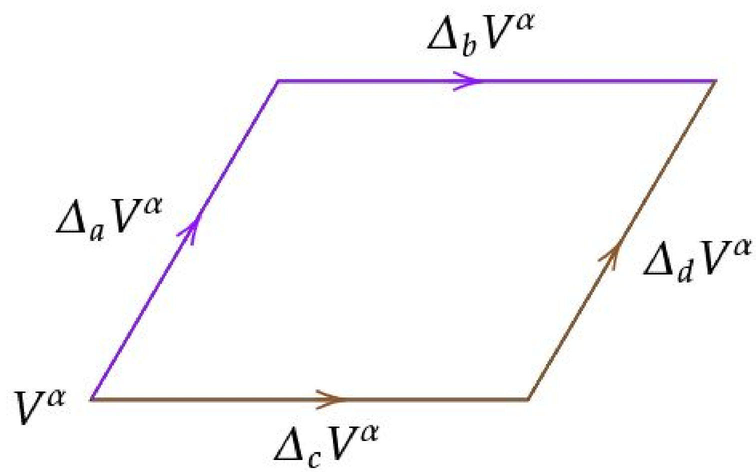

Figure 2: Comparing two different routes to get from the bottom left to the top right corner.

take the difference, $V_1^\alpha - V_2^\alpha$. Along each journey, V_1^α accumulates two changes, $\Delta_a V^\alpha$ and $\Delta_b V^\alpha$. Meanwhile, V_2^α accumulates the changes $\Delta_c V^\alpha$ and $\Delta_d V^\alpha$. Thus,

$$\delta V^\alpha \approx V_1^\alpha - V_2^\alpha = \Delta_a V^\alpha + \Delta_b V^\alpha - \Delta_c V^\alpha - \Delta_d V^\alpha. \tag{6}$$

In equation (6), we note that the difference we calculated is approximately equal to δV^α in equation (1). Let T_α^β denote the operator which takes vectors at the top right corner and parallel transports them to the bottom left corner. Then, we see that

$$\delta V^\beta = T_\alpha^\beta (V_1^\alpha - V_2^\alpha). \tag{7}$$

This is because $T_\alpha^\beta V_1^\alpha$ represents the vector which is transported along all four legs of the parallelogram. V_1^α already covered the first two legs, and then T_α^β carries it the last two legs. Meanwhile, $T_\alpha^\beta V_2^\alpha$ represents the initial vector, since V_2^α first went out along the last two legs but then gets brought back in exact cancellation. Thus, $T_\alpha^\beta (V_1^\alpha - V_2^\alpha)$ represents the difference between the final and initial vectors encapsulated in δV^α. Now, to leading order, T_α^β is just the identity, since the additional terms will pick up factors of dx as one integrates derivatives along the infinitesimal legs. The quantity $V_1^\alpha - V_2^\alpha$ already contains all the high order terms we seek (terms of order $dxdx$), and we are discarding all higher order terms (three or more terms of dx). Thus, to leading order, we have shown equation (6).

Now, we seek to show that $\delta V^\alpha = [\nabla_\lambda, \nabla_\sigma]V^\alpha$ for a smooth vector field V^α. Even though we can't ensure $\nabla_\mu V^\alpha = 0$ on all four legs, we are allow to make it zero on two legs. The curvature will then be encoded in the value of the covariant derivative on the remaining two legs. Specifically, we choose the left and bottom legs (a and c) to be zero.

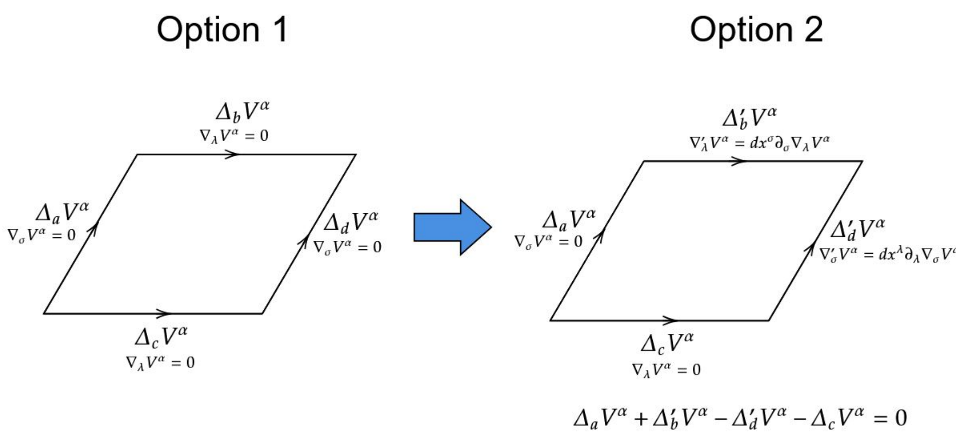

Figure 3: Transitioning from option 1 to option 2

Figure 3 depicts both option 1 and option 2. Specifically, it denotes the case of option 1 on the left which enforces the parallel transport condition of all four legs. The changes V^α undergoes as it moves along each leg, in the direction of the arrow, is given by $\Delta_x V^\alpha$ for leg x. Meanwhile, on the right, we have option 2. We are able to enforce the parallel transport condition on legs a and c. Consequently, the change V^α undergoes on either leg, $\Delta_a V^\alpha$ and $\Delta_c V^\alpha$, is unchanged. However, this is not doable on legs b and d. Instead, the covariant derivative on leg b is given by $dx^\sigma \partial_\sigma \nabla_\lambda V^\alpha$. Likewise, on leg d, the covariant derivative is given by $dx^\lambda \partial_\lambda \nabla_\sigma V^\alpha$. Throughout the article, we are going to speak of the process of going from option 1 to option 2. What this means is that we imagine taking every value of V^α at each point in the manifold in option 1 and constructing a vector field assigning that value to that point. The process of going from option 1 to option 2 is then modifying that vector field along legs b and d, by changing $\Delta_b V^\alpha$ and $\Delta_d V^\alpha$, in order to ensure the smooth condition.

Because the covariant derivative on legs a and c are still zero in option 2, we have

$$\nabla_\sigma \nabla_\lambda V^\alpha = \partial_\sigma \nabla_\lambda V^\alpha, \tag{8}$$

since the curvature terms attached to $\nabla_\lambda V^\alpha$ vanish.

Similarly,

$$\nabla_\lambda \nabla_\sigma V^\alpha = \partial_\lambda \nabla_\sigma V^\alpha. \tag{9}$$

In addition to different covariant derivatives along legs b and d, we also have different changes in the components V^α on these two legs. We denote these changes by $\Delta'_b V^\alpha$ and $\Delta'_d V^\alpha$. Now, the smooth vector field condition of option 2 requires that the total change V^α undergoes as it goes around the parallelogram must sum to zero (or, equivalently, that going from the bottom left corner to the top right corner should yield the same result via the blue or red line in Figure 2). Therefore,

$$\Delta_a V^\alpha + \Delta'_b V^\alpha - \Delta'_d V^\alpha - \Delta_c V^\alpha = 0.$$

This implies

$$\Delta_\alpha V^\alpha - \Delta_c V^\alpha = \Delta'_d V^\alpha - \Delta'_b V^\alpha.$$

Inputting this into equation (6), swapping an approximate equality for a plain equality given that we ignore higher order terms,

$$\delta V^\alpha = \Delta'_d V^\alpha - \Delta_d V^\alpha - (\Delta'_b V^\alpha - \Delta_b V^\alpha). \tag{10}$$

Now,

$$\nabla_\lambda V^\alpha = \partial_\lambda V^\alpha + \Gamma^\alpha_{\lambda\mu} V^\mu. \tag{11}$$

For leg b, in option 1, it is simply zero, and thus

$$dx^\lambda \nabla_\lambda V^\alpha = \Delta_b V^\alpha + dx^\lambda \Gamma^\alpha_{\lambda\mu} V^\mu = 0. \tag{12}$$

Thus for leg b,

$$dx^\lambda \Gamma^\alpha_{\lambda\mu} V^\mu = -\Delta_b V^\alpha, \tag{13}$$

and similarly for leg d

$$dx^\sigma \Gamma^\alpha_{\sigma\mu} V^\mu = -\Delta_d V^\alpha. \tag{14}$$

Note that in equations (13) and (14) we take V^μ along legs b and d respectively.

Now, when we switch to option 2, the covariant derivatives along legs b and d take on nonzero values, $\nabla'_\lambda V^\alpha$ and $\nabla'_\sigma V^\alpha$ respectively.

Now,

$$dx^\lambda \nabla'_\lambda V^\alpha = \Delta'_b V^\alpha + dx^\lambda \Gamma^\alpha_{\lambda\mu} V^\mu. \tag{15}$$

We are able to substitute equation (13) into equation (15). Importantly, the reason we are able to do this is because to leading order, V^μ is the same in leg b for both options 1 and 2. This is a result of the fact that the accumulated change in V^μ along leg a is the same in options 1 and 2, since in both we preserve the parallel transport condition. Thus,

$$dx^\lambda \nabla'_\lambda V^\alpha = dx^\lambda dx^\sigma \partial_\sigma \nabla_\lambda V^\alpha = \Delta'_b V^\alpha - \Delta_b V^\alpha.^\ddagger \tag{16}$$

Likewise,

$$dx^\lambda dx^\sigma \partial_\lambda \nabla_\sigma V^\alpha = \Delta'_d V^\alpha - \Delta_d V^\alpha. \tag{17}$$

Thus, inputting into equation (10) and factoring in equations (8) and (9), we find

$$\delta V^\alpha = dx^\lambda dx^\sigma (\nabla_\lambda \nabla_\sigma V^\alpha - \nabla_\sigma \nabla_\lambda V^\alpha). \tag{18}$$

Matching to equation (1), we find

$$R^\alpha_{\mu\lambda\sigma} V^\mu = [\nabla_\lambda, \nabla_\sigma] V^\alpha, \tag{19}$$

which is exactly what we set out to prove. The algebraic steps reveal precisely why the discrepancies in option 1 and option 2 result in the same quantity. However, it is still illuminating to take a step back and see exactly why the two quantities were expected to be the same. We do so now.

3 A retrospect

It all comes down to δV^α. Analyzing how the components of V^α change as one goes around the parallelogram is how [1] derives the Riemann tensor and is typically how most students conceptually remember it.

When we go from option 1 to option 2, what are we doing? We are modifying legs b and d, altering the covariant derivative such that the net accumulated change around the parallelogram goes to zero. In this alteration, we are changing the way the components vary along legs b and d, swapping $\Delta_b V^\alpha$ for $\Delta'_b V^\alpha$ and $\Delta_d V^\alpha$ for $\Delta'_d V^\alpha$. We've already established that

$$dx^\lambda dx^\sigma \nabla_\sigma \nabla_\lambda V^\alpha = \Delta'_b V^\alpha - \Delta_b V^\alpha.$$

In going from option 1 to option 2, we turned $\Delta_b V^\alpha$ into $\Delta'_b V^\alpha$ - we'll call the difference

$$\delta_b V^\alpha = \Delta'_b V^\alpha - \Delta_b V^\alpha = dx^\lambda dx^\sigma \nabla_\sigma \nabla_\lambda V^\alpha.$$

Likewise,

$$\delta_d V^\alpha = \Delta'_d V^\alpha - \Delta_d V^\alpha = dx^\lambda dx^\sigma \nabla_\lambda \nabla_\sigma V^\alpha.$$

It makes sense that the change in the covariant derivative is the same as the change in the accumulations in V^α, since in equation (11) we see that if

‡A note on notation: the primes in $\nabla'_\lambda V^\alpha$, $\Delta'_b V^\alpha$ and so forth are intended to distinguish option 1 and option 2, calling attention to the fact that terms without primes are equivalent in both cases. However, in option two, we can still call the covariant derivative along the b leg $\nabla_\lambda V^\alpha$ which, in this case, equals $dx^\sigma \partial_\sigma \nabla_\lambda V^\alpha$ and similarly for the d leg.

the Christoffel symbols remain fixed, which they do since in going from option 1 to option 2 we don't change the underlying curvature of the manifold, then there is a direct equality between these changes. In option 1, we had our total curvature δV^α which we sought. In going to option 2, we made changes by adding $\delta_b V^\alpha$ and subtracting $\delta_d V^\alpha$ (it's a subtraction because going around the parallelogram clockwise entails walking backward along legs c and d). These changes canceled δV^α:

$$\delta V^\alpha + \delta_b V^\alpha - \delta_d V^\alpha = 0. \tag{20}$$

$$\delta V^\alpha = \delta_d V^\alpha - \delta_b V^\alpha = dx^\lambda dx^\sigma (\nabla_\lambda \nabla_\sigma V^\alpha - \nabla_\sigma \nabla_\lambda V^\alpha) \tag{21}$$

Figure 4 briefly summarizes this process. We start in option 1, with covariant derivatives zero on all legs and δV^α nonzero. Then, to go to option 2, we make necessary changes to legs b and d to bring δV^α to zero. These necessary changes amount to adding additional terms to $\Delta_b V^\alpha$ on leg b and $\Delta_d V^\alpha$ on leg d. Due to the Christoffel symbols remaining fixed as we switch from option 1 to 2, the additional terms are precisely the changes in the covariant derivative of V^α. These additional terms serve to cancel δV^α (after taking into account the sign reversal for leg d). If adding these terms to δV^α brings it to zero, adding the reverse to zero yields δV^α, and thus δV^α equals the reverse of these terms. Thus, we're able to finally conclude with equation (19).

This picture clears up a conceptual issue with the signage of equation (3) that some students may have. Since, in calculating δV^α, we go positive along leg a but negative along leg d, it seems like perhaps the relevant quantity is $-\nabla_\lambda \nabla_\sigma V^\alpha$ instead of $\nabla_\lambda \nabla_\sigma V^\alpha$, since we presumably would want to be subtracting the term at greater x^λ. However, we see now that $\nabla_\lambda \nabla_\sigma V^\alpha$ doesn't reflect a part of δV^α but rather what it takes to bring δV^α to zero, and thus the actual contributing term is the reverse, $-\nabla_\lambda \nabla_\sigma V^\alpha$.

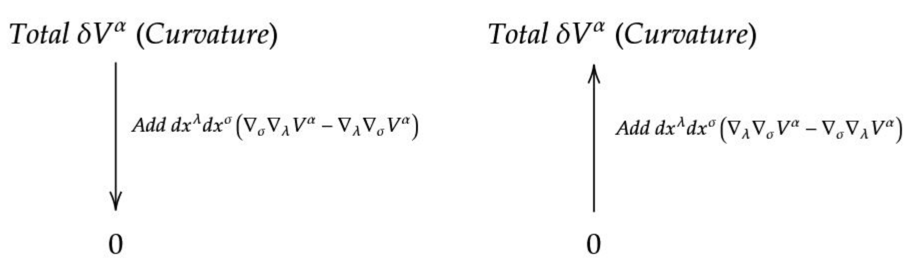

Figure 4: Summary of how to transition from option 1 to 2 and vice versa.

References

[1] Scott A. Hughes General Relativity Massachusetts Institute of Technology, web.mit.edu/sahughes/www/8.962/lec10.pdf.