

## 6.7: Abstract Index Notation

### Learning Objectives

- Developing some mathematical notation related to waves

This chapter has centered on the physics of waves, but along the way we've found it helpful to build up some mathematical ideas such as covectors, which have applications in a much broader physical context. In this section we'll develop some related notation.

Expressions in birdtracks notation such as

can be awkward to type on a computer, which is why we've already been occasionally resorting to more linear notations such as  $(\nabla C) \rightarrow s$ . For more complicated birdtracks, the diagrams sometimes look like complicated electrical schematics, and the problem of generating them on a keyboard get more acute. There is in fact a systematic way of representing any such expression using only ordinary subscripts and superscripts. This is called abstract index notation, and was introduced by Roger Penrose at around the same time he invented birdtracks. For practical reasons, it was the abstract index notation that caught on.

The idea is as follows. Suppose we wanted to describe a complicated birdtrack verbally, so that someone else could draw it. The diagram would be made up of various smaller parts, a typical one looking something like the scalar product  $u \rightarrow v$ . The verbal instructions might be: *"We have an object  $u$  with an arrow coming out of it. For reference, let's label this arrow as  $a$ . Now remember that other object  $v$  I had you draw before? There was an arrow coming into that one, which we also labeled  $a$ . Now connect up the two arrows labeled  $a$ ."*

Shortening this lengthy description to its bare minimum, Penrose renders it like this:  $u_a v^a$ . Subscripts depict arrows coming out of a symbol (think of water flowing from a tank out through a pipe below). Superscripts indicate arrows going in. When the same letter is used as both a superscript and a subscript, the two arrows are to be piped together.

Abstract index notation evolved out of an earlier one called the Einstein summation convention, in which superscripts and subscripts referred to specific coordinates. For example, we might take 0 to be the time coordinate, 1 to be  $x$ , and so on. A symbol like  $u_\lambda$  would then indicate a component of the dual vector  $u$ , which could be its  $x$  component if  $\lambda$  took on the value 1. Repeated indices were summed over.

The advantage of the birdtrack and abstract index notations is that they are coordinate-independent, so that an equation written in them is valid regardless of the choice of coordinates. The Einstein and abstract-index notations look very similar, so for example if we want to take a general result expressed in abstract-index notation and apply it in a specific coordinate system, there is essentially no translation required. In fact, the two notations look so similar that we need an explicit way to tell which is which, so that we can tell whether or not a particular result is coordinate-independent. We therefore use the convention that Latin indices represent abstract indices, whereas Greek ones imply a specific coordinate system and can take on numerical values, e.g.,  $\lambda = 1$ .

The following are some examples of equivalent equations written side by side in birdtracks and abstract index notations.

Observer  $o$ 's displacement in spacetime is a vector:

$$\rightarrow o \quad o^a \quad (6.7.1)$$

In Einstein notation, it's awkward to express a vector as a whole, because in a notation like  $o^\lambda$ ,  $\lambda$  is supposed to take on a particular value. If we used  $o^\lambda$  to mean the whole vector, it would be an abuse of notation. In abstract index notation, however, the  $a$  is simply a name we gave to a pipe coming into vector  $o$ ; the fact that we didn't need to refer to the name in order to connect it to some other pipe is irrelevant.

A wave's frequency is a covector:

$$\omega \rightarrow \quad \omega_a \quad (6.7.2)$$

An observer experiences proper time  $\tau$ :

$$o \rightarrow o = \tau^2 \quad o_a o^a = \tau^2 \quad (6.7.3)$$

There are no external arrows in the birdtracks version, and in the abstract-index version all lower indices (pipes coming out) have been paired with upper indices (pipes coming in); this indicates that the proper time is a scalar, and therefore independent of any choice of coordinate system. In Einstein notation, this becomes  $o_\lambda o^\lambda$ , with an implied sum over the repeated index,  $\sum_\lambda o_\lambda o^\lambda$ . The

$\lambda$  refers to a particular coordinate system, so in the Einstein notation it is no longer obvious that the equation holds regardless of our choice of coordinates.

A world-line along which a wave propagates lies along a vector that is orthogonal to the wave's frequency covector:

$$\omega \rightarrow u = 0 \quad \omega_a u^a = 0 \quad (6.7.4)$$

The frequency covector is the gradient of the phase:

$$\omega \rightarrow = \bigcirc \phi \rightarrow \quad \omega_a = \nabla_a \phi$$

The following grammatical rules apply to both abstract-index and Einstein notation:

1. Repeated indices occur in pairs, with one up and one down and the two factors multiplying each other.
2. Disregarding indices that are paired as in rule 1, all other indices must appear uniformly in all terms and on both sides of an equation. "Appear uniformly" means that an index can't be missing and can't be a superscript in some places but a subscript in others.
3. For reasons to be explained in section 7.4, a partial derivative with respect to a coordinate, such as  $\partial/\partial x^k$ , is treated as if the index were a subscript, and conversely  $\partial/\partial x_k$  is considered to have a superscripted  $k$ .

In abstract-index notation, rule 1 follows because the indices are simply labels describing how, in birdtracks notation, the pipes should be hooked up. Violating rule 1, as in an expression like  $v^a v^a$ , produces a quantity that does not actually behave as a scalar. An example of a violation of rule 2 is  $v^a = \omega_a$ . This doesn't make sense, for the same reason that it doesn't make sense to equate a row vector to a column vector in linear algebra. Even if an equation like this did hold in one frame of reference, it would fail in another, since the left-hand and right-hand sides transform differently under a boost.

In section 6.4 we discussed the notion of finding the covector that was dual to a given vector, and the vector dual to a given covector. Because the distinction between vectors and covectors is represented in index notation by placing the index on the top or on the bottom, relativists refer to this kind of thing as raising and lowering indices. In general, this type of manipulation is called "index gymnastics." Here's what raising and lowering indices looks like.

Converting a vector to its covector form:

$$u_a = g_{ab} u^b \quad (6.7.5)$$

Changing a covector to the corresponding vector:

$$u^a = g^{ab} u_b \quad (6.7.6)$$

The symbol  $g^{ab}$  refers to the inverse of the matrix  $g_{ab}$ .

---

This page titled [6.7: Abstract Index Notation](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Benjamin Crowell](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.