Intuitive Guide to Principles of Communications
www.complextoreal.com

Lattice and Trellis Coded Modulation (TCM)

Concept of Lattice

The concept of lattice is important in coding and particularly relevant in Trellis Coded Modulation (TCM). A lattice is a simple concept to visualize. It is just a set of points that have some “regularness”.

First let’s look at a few of the important lattices used in TCM and their properties.

The lattice below is a general 2 dimensional lattice and it has a specific name called $\mathbf{Z}^2$. Z stands for a field consisting of integers and 2 for the order of the dimension. A $\mathbf{Z}^N$ lattice would the same thing in N dimensions.

Figure shows a 1-dimensional lattice and this would be written as just plain **Z**.

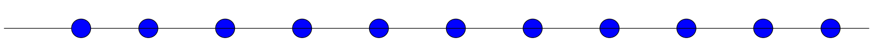

Fig. 1 – A two dimensional lattice, **Z**

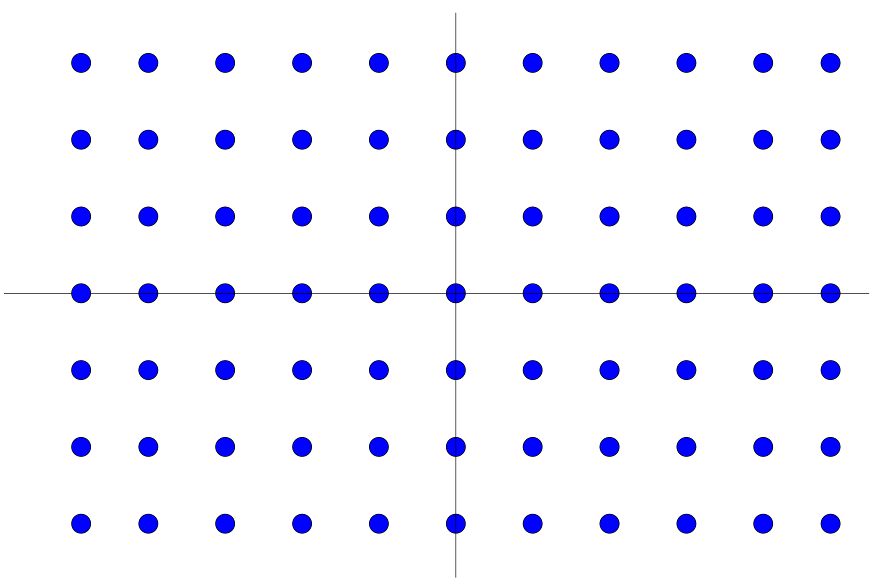

Fig. 2 – A two dimensional lattice, $\mathbf{Z}^2$

The basis vectors for the 2 dimensional lattices are

$$x_1 = [1,0]$$
$$x_2 = [0,1]$$

and the generator matrix is given by

$$G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

What is a generator matrix? It is the smallest unit of points that can be rotated or translated to produce all other points in the lattice.

A rotated lattice $\mathbf{RZ}^2$

If we rotate the above 2-dimensional lattice, $\mathbf{Z}^2$ by 45 degrees we get the following lattice. It is designated by adding **R,** a rotation parameter in front of the $\mathbf{Z}^2$. (in TCM we like to keep our points at integer units so we multiply it by $\sqrt{2}$.) The rotated lattice looks like this.

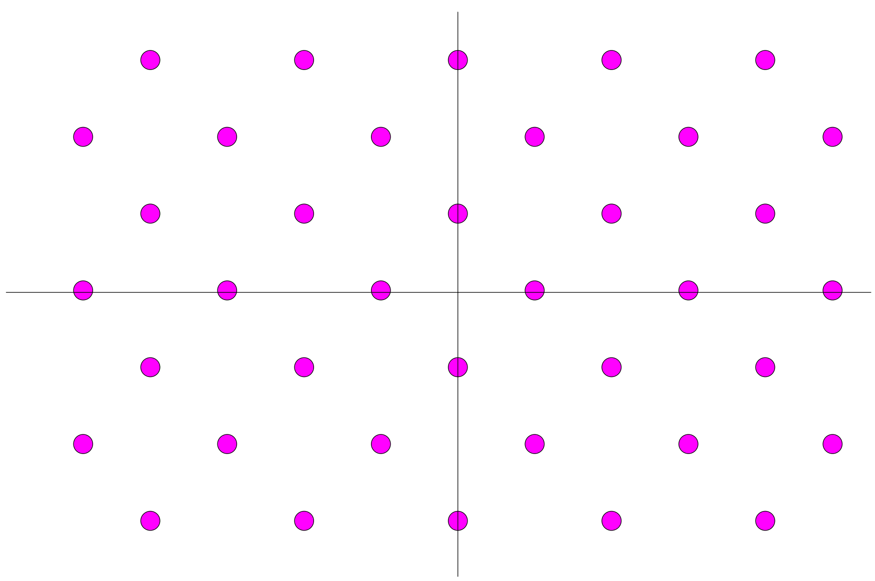

Fig. 3 – A rotated lattice $\mathbf{RZ}^2$

The basis for this lattice is

$$x_1 = [1, \quad 1]$$
$$x_2 = [1, -1]$$

and its generator matrix G is given by

$$G = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

If you look carefully, this rotated lattice can be obtained from the regular $\mathbf{Z}^2$ lattice by getting rid of alternate points in $\mathbf{Z}^{2.}$. So we can get two $\mathbf{RZ}^2$ lattice from one $\mathbf{Z}^2$ lattice as shown below in alternating colors for points belonging to two $\mathbf{RZ}^2$ lattice. (I will use the term lattice for plural as well.)

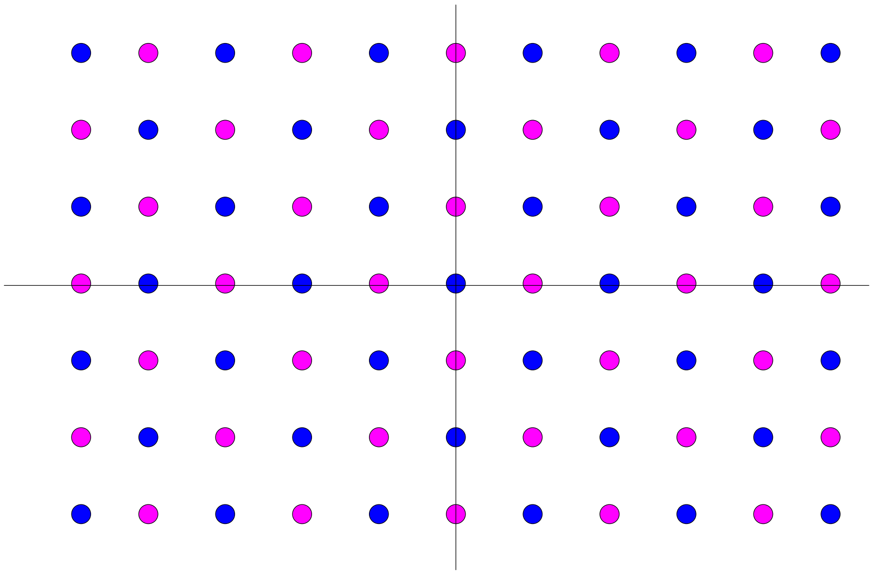

Fig. 4 – The original lattice which contains two $\mathbf{RZ}^2$ lattice.

The blue points belong to one $\mathbf{RZ}^2$ and the pink points belong to an $\mathbf{RZ}^2$. When you add the two you get the original back $\mathbf{Z}^{2.}$ So we say that that $\mathbf{RZ}^2$ is a coset $\mathbf{Z}^2$.

Scaled integer lattice 2Z²

A scaled $\mathbf{2Z}^2$ is shown below. It is just a scaled version of $\mathbf{Z}^2$. The term 2 in front of $\mathbf{Z}^2$ means that is has been scaled by factor of 2 in one of the directions. The distance horizontally has doubled to 2 units.

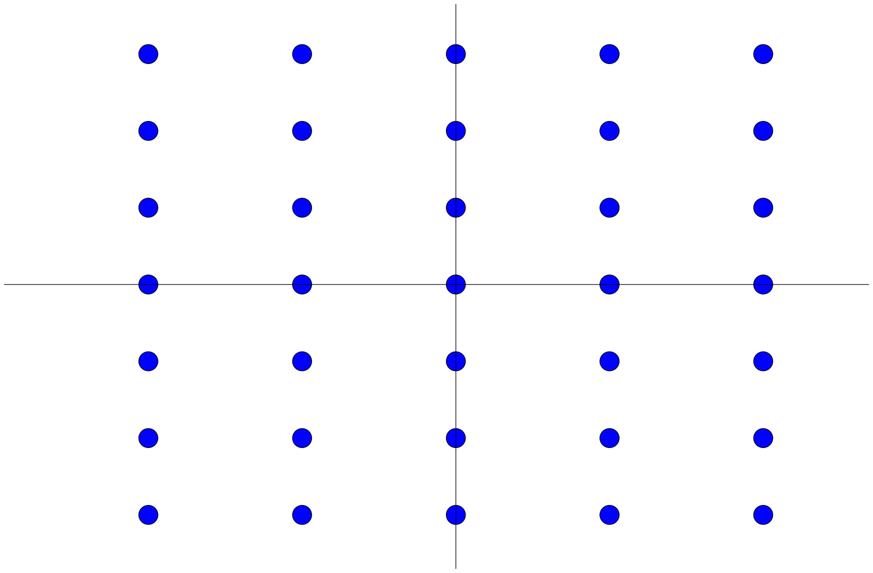

Fig 5 – Scaled lattice, $2\mathbf{Z}^2$

In current state of communications, the modulation signal design is based on these lattice patterns. The lattice is the general form behind constellations such as QPSK, QAM and M-PSK 32-QAM etc. which can all be considered a subset of these general lattices. Our communication constellations inherit the properties of these lattice.

The rectangular lattice actually are not the most efficient. The most efficient lattice is a hexagonal lattice shown below. Where each point has six neighbors.

This hexagonal lattice is called $\mathbf{A}^2$

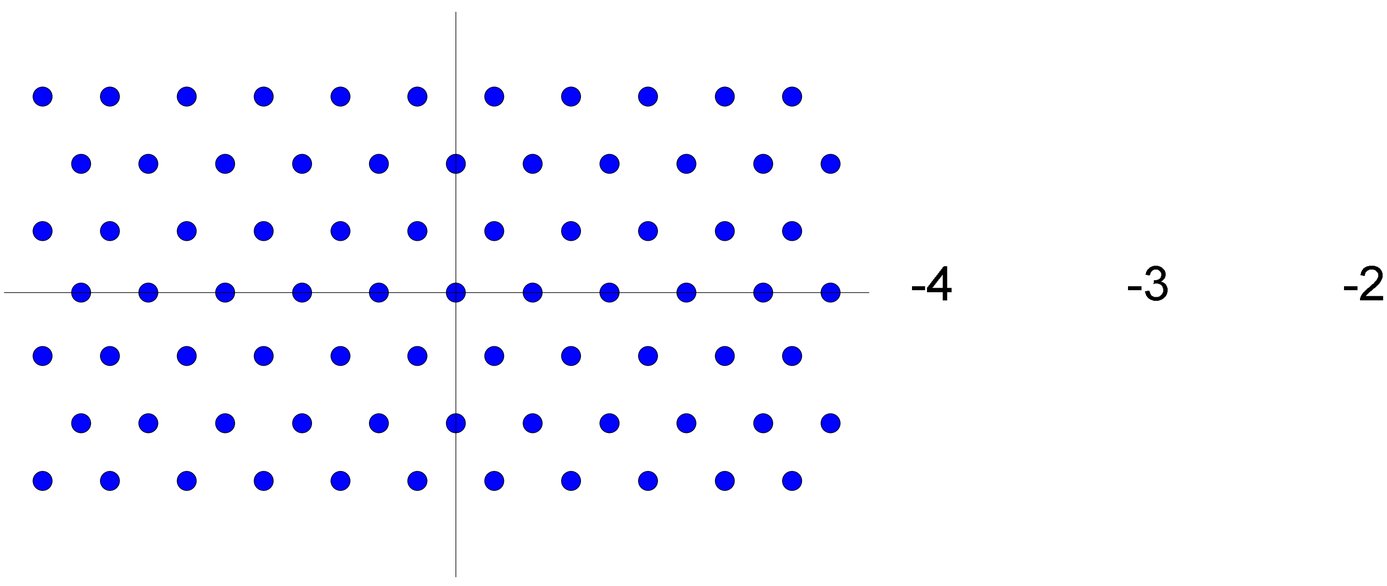

Fig 6 – hexagonal lattice, $\mathbf{A}^2$

And its generator matrix is given by

$$G = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.5\sqrt{3} \end{bmatrix}$$

the coordinates of this transformation are of course no longer integers. $\mathbf{A}^2$

In modulation, this constellation would be better to use, because we can pack in more symbols per bandwidth but due to limitations of current state of the art in signal processing, we use rectangular constellations even though there are other more optimal options.

Four dimensional lattice, $\mathbf{D}^4$

This is a four dimensional Lattice, which I can not draw for you. You can think of these as sequences that are orthogonal to each other. The sum of the components of each vector in four dimensions for a lattice is even.

The generator function is given by

$$G = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Higher dimension lattice have proven to be very efficient and are now used in TCM as we shall see later.

Sublattice, lattice partitions and cosets

Cosets

A translation of a sub-lattice (or a subset) by an element of the original lattice is called a coset. This confusing definition is made easy by the following figure.

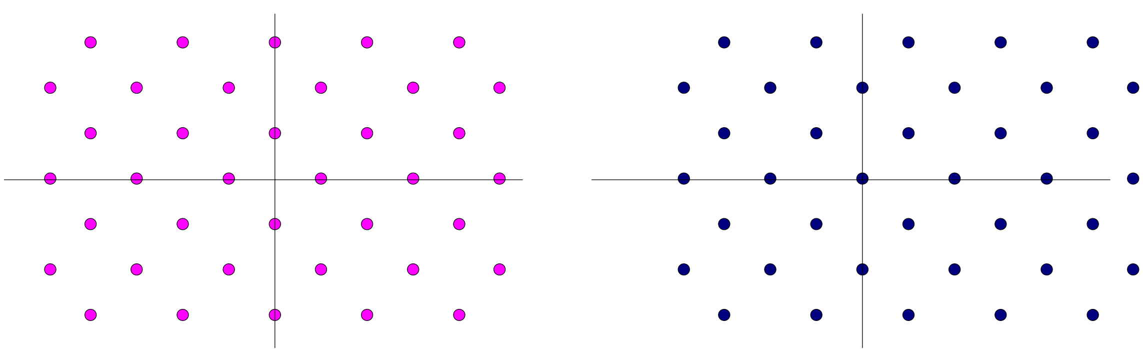

Fig. 7 – Two unique sub-lattice which when combined give the general Z^2 lattice.

The pink points and the blue points are cosets of $\mathbf{Z^2}$. Together the two coset make the original lattice of $\mathbf{Z^2}$.

A subset can have many cosets. All cosets are disjoint. Which means they do not share any points. The union of all unique cosets gives the parent subset. So we can see cosets as children of the parent sub-set, each unique.

Still not clear what a coset is? Let's take a look at the 16QAM constellation. We split the constellation in two unique cosets. Of course there are many ways this could have been done. But in signal processing we are neighbor-shy and want to get as far away as possible from our neighbor points. So please try other ways to portion the 16 points and see if you can get a better way to divide the subset such that;

1. The distance between each point in the divided subset is the same.
2. The distance is larger than the distance between the points in the previous level subset.
3. There are equal number of points in each coset.

I am pretty sure that you can not do better than the solution given here. The 16 point subset is partitioned into two cosets such that the minimum distance between its points is now 1.414, instead of 1.

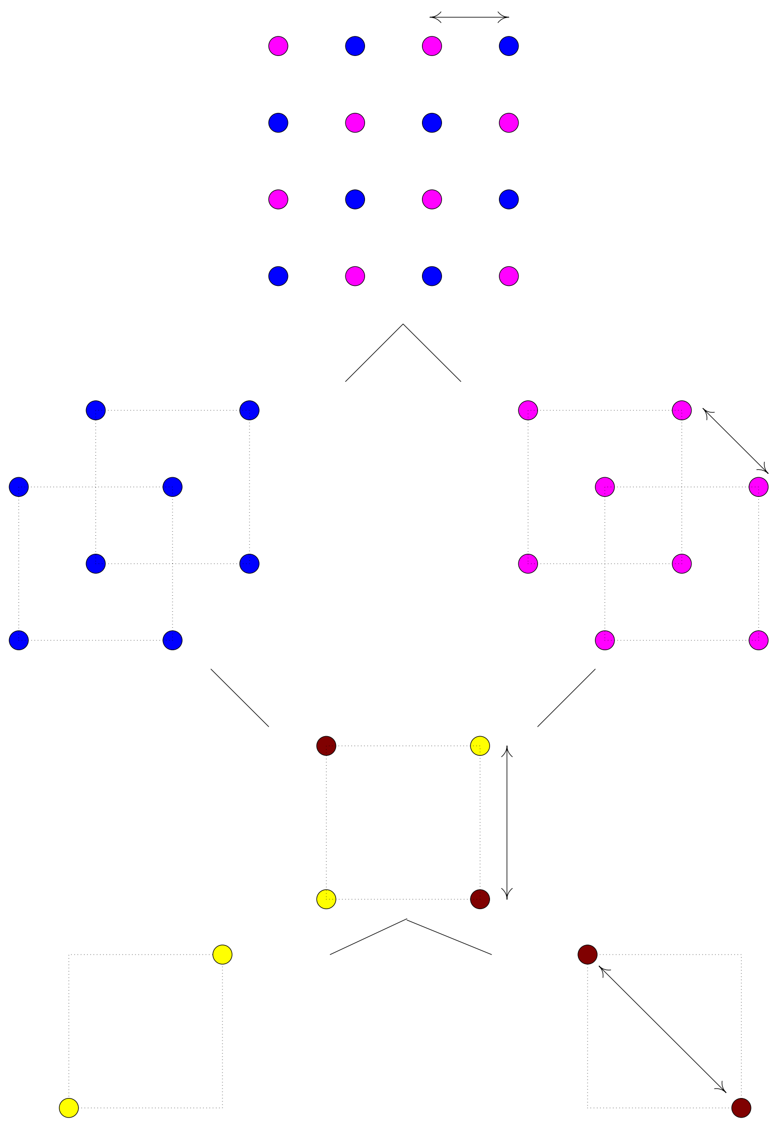

Fig. 8 – Partition of 16QAM lattice into cosets

The two cosets at the first level can be portioned into two more unique cosets, which is the QPSK constellation. And the QPSK constellation can be partitioned into two cosets again that contain two points only. The distance between adjacent points gets larger at each partition.

Now look at the case of 8PSK. As we did above, the coset partitions are easy as and as above at each partition, the nearest neighbor distance increases.

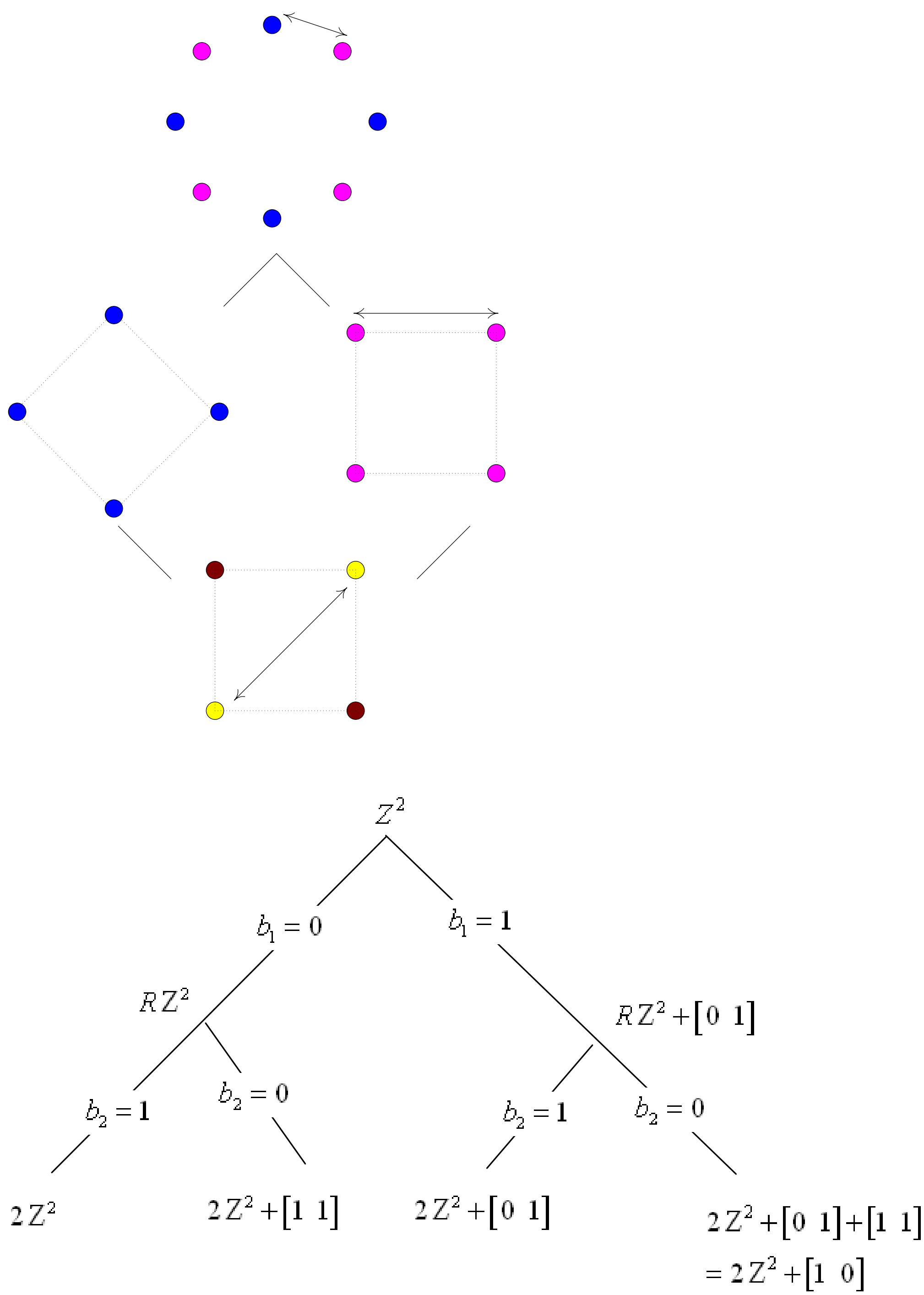

Fig. 9 – Partition of 8PSK lattice into cosets

The number of distinct cosets in a sub-lattice is called the order of the partition. Here in each case we have just two unique cosets at each level so the total partition has an order of 8, because to get to a single point, we have to partition the lattice 3 times, 2 of which are shown above.

The 16QAM constellation was partitioned 4 times to get to a single point, so its order is $2^4 = 16$.

Performance metrics for lattice

We define three parameters of a lattice that help to quantify its performance. These are

1. **Minimum distance**
2. **Fundamental Volume**
3. **Fundamental coding gain**

Minimum distance

The concept of minimum distance, d_{min} is pretty straight forward and clear as you see above. The distance between points in a coset is always larger than the distance of lattice points in the partition above it. This is an important fact used in TCM because the average distance then is larger than if the lattice was not partitioned at all.

Fundamental Volume

The points in a lattice have a density as seen by number of points in a space. The reciprocal of this density is called the ***fundamental volume*** associated with each point in the lattice. This is the space around each point. Looser the better. Let's call it V or $V(\Lambda)$ of a lattice. This volume is quantified by

$$V(\Lambda) = |\det G|$$

where G is the generator matrix of the lattice. This makes sense, since the generator matrix tells us how far apart (loose or dense) a unit shape is.

Fundamental volume of $\mathbf{Z^2}$ matrix is 1, it is 2 for $\mathbf{RZ^2}$ and $\mathbf{D^4}$. $\mathbf{RZ^4}$ has a volume of 4. A large volume is good, as we shall see next. This is why multi-dimensional lattice are used in TCM.

Fundamental coding gain

The fundamental coding gain of any lattice is given by a relationship between the above two parameters, the minimum distance d_{min} and the fundamental volume, V.

$$\frac{d^2_{\min}}{V^{2/N}}$$

A lattice $\mathbf{Z^2}$ has a volume of 1 and its d_{min} is also 1, so its coding gain is 1. This coding gain is not a function of scaling or transformation.

The coding gain of $\mathbf{D^4}$ lattice: the d_{min} is 1.414, its volume is 1.505,

So its coding gain is 10log (1.414) – 1.505

This is the best shape factor coding gain possible. In coding gain, then we have two things, the first factor is distance. The larger the distance, the better the coding gain. The second factor is called the shape factor and comes from packing the signals well. For rectangular constellations, this number is 1 so we gain nothing but if a honeycomb like shape can be used, then the fundamental volume of each point is small hence a net gain.

The largest shape gain possible is -1.505 dB.

Other tutorials at
www.complextoreal.com