I'm probably unusually sensitive to image with a stereo system. I'm one of those guys who was always jumping up to tweak the balance control in pre-remote days, or moving myself to the left or right on the seat to keep all the instruments and the singer from sounding like they are crowding to one side. Even for single-listener sessions, sweet spots were just too small. And multiple listeners tended to basically get mono from whichever speaker was closer.

Most people who pay any attention to waveguides have heard about toe-ing them in, but a lot of comments I've seen on forums suggest that most people didn't actually get the point. So here are some graphics and words to (I hope) make the idea clearer.

One of the big advantages (in my opinion, the biggest) of waveguide speakers is the possibility of having a very wide "sweet spot" and very stable front images. To see how, consider what determines the apparent sound source direction in a stereo setup.

When the same sound comes at us from two speakers, the apparent source direction depends on which speaker's sound is more intense

If these factors are similar from both speakers, then the sound appears to come from somewhere in between the speakers. The intensity and arrival times, if of opposite effect in the two speakers, can also counteract. If a sound that arrives
earlier from one direction is not as loud as a copy that arrives later from a different direction, the perceived location can lie somewhere between the two actual source directions. (Google "time-intensity trading" for more about this).

Most audiophiles assume that the ideal speaker should be an omnidirectional point source. Here is a picture to illustrate the horizontal intensity patterns you’d get from two such speakers if you could achieve that kind of radiation:

A listener at position "m" will hear both speakers with similar intensity, and being about the same distance from each, a center image will appear between the speakers. But if he moves his head to position "o", a little off-center, he gets nearer to the right channel speaker so it is heard first (time factor). And the intensity from the right side also becomes stronger (intensity factor…in the graphic, sound intensity is indicated by color intensity).

Both factors, time and intensity, make the sound seem to come from the same place. So the image abruptly slides toward the right, and if the listener is more than a little off center, the image is mostly crowded near the right speaker. So, the omni point source is not so ideal! Use of a center channel can help a lot with this, but the side speakers are still working against keeping the image stable.

Another issue is that with omnidirectional speakers, the sound going backwards away from the listener will eventually bounce around and get back to him. If its spectral character is similar, that will add some ambience which if delayed long
enough is probably good. (If its character is different though -- non-constant directivity-- then the reflections will tend to betray the source as coming from a speaker, not good). In the graph above, though, the sound going to the side walls will reflect and arrive at the listener very soon after the direct sound from the same channel, resulting in a coloration from the filtering effect. That's not so ideal, either.

So let's assume we instead have speakers that are directional with a radiation pattern like this:

As your listening position gets off-axis (so that the speaker isn't pointing at you so directly), the intensity gets weaker. To be fully helpful, all frequencies, or at least the mid and treble ones, should do this similarly. For now, assume that the whole spectrum keeps the same general shape ("constant directivity") at different angles off-axis, but gets lower in intensity as you go further off axis ("controlled directivity"). This is usually accomplished with large-ish waveguide-based speakers.
Put two of those speakers in stereo and point them forward like this, and--:

Oops. **It's even worse!** As you move from the mid position "m", where not much direct sound reaches you now, toward "o", you not only get closer to the right channel, but the intensity from the right goes up rapidly while intensity from the left drops down just as fast. The image leaps to the right! Terrible. And the majority of reflections are near reflections from the close side walls coloring the audio spectra. That's what happens if you take controlled directivity waveguide speakers and set them up firing directly forward. This is the only way many listeners have ever heard waveguide speakers – no wonder they didn’t like them!

If you need to set up like that, you are probably better off **not using controlled directivity speakers**, but instead going for something with wider coverage.
Ok, so we'll toe them in some -- aim the speakers right at the center listener:

That's what people concerned about sound often do with speakers that are designed to have flat response on-axis, so that the response of the direct path sound waves is flat (though if the speaker isn't constant directivity, the reflected sound waves will have a different character).

Now, that's better, because as you go to the right, the intensity of both speakers drops, though the left speaker still drops faster (and its sound of course still arrives later). The image still shifts to the right, but not as badly as without the speakers toed-in. The near reflections off the close wall are much reduced, so that's a plus. And of course if the speakers are designed for proper response only on-axis, then toed-in like this is the only way to get the designed-for response for direct waves … even for the single centered listener.

So, is toe-ing in just a “fix” for some kind of defect with controlled directivity designs? No.
Try toe-ing in the speakers *considerably* further. In fact, point each speaker roughly at the furthest listening seat from it:

Now things get more interesting! When you move toward the right, the sound intensity from the closer right speaker now *decreases*, because you are getting further off-axis from it.

The intensity from the further left speaker *increases* even though you are getting further from it, since you getting more onto its strong main axis.

The intensity effects would make the image tend to go left. But you're closer to the right speaker, so time precedence would make the image tend to go right. The intensity effect goes opposite to the sound arrival time effect – “time-intensity trading” –

and the image tends to remain stable, something that omnidirectional speakers just cannot really do.

When properly balanced (achieved by adjusting the degree of toe-in), the image with controlled directivity speakers stays steady across a wide range of listening positions, and no matter how much the listeners move their heads or shift in their
seats. The image becomes independent of the listener’s movements, and provides a surprisingly relaxed listening experience. It is possible to get this effect in some installations even for listeners seated almost directly in front of the left or right speakers.

Interesting things happen with the reflections, too. The first strong reflections to reach you come from the wall but originate from the furthest opposite speaker, so it is much more delayed and provides an enveloping effect without coloring the timbre of the sounds.

The effect can be pretty amazing with controlled directivity waveguide speakers. In fact, in my system I now usually sit almost directly in front of my right speaker (because it’s nearer to a power outlet I can plug the laptop into, and nearer to the fireplace for warming feet!). The imaging is stellar and a solo singer stays right between the speakers. If you’ve “heard waveguides”, but they weren’t properly toed-in, then, well -- no, you haven’t heard waveguides.

One other pattern-related thing I’d like mention is obstructions in the room. The controlled directivity waveguides also control in the vertical directions, where they are often designed to have a narrower pattern (to avoid floor and ceiling reflections at high frequencies). If you put the speakers on the floor, with the waveguides at knee-level, then they will have strong radiation toward all the various things on the floor of your room: furniture, plants, lamps, etc., and may generate near reflections and general clutter off of these. But if you arrange the speaker so that the waveguide is near ear level, then horizontally there will fewer things up there for the sound to bank off of and the image will become expansive. In my system, that seemed to make a very significant difference. I originally thought I could put the speakers on the floor, tipped up toward my chair and all would be well. But when I put them up on stands instead, there was no way they were going back to the floor again. Try it.

Credit where credit is due:

I first heard of time-intensity trading with speakers in an AES paper about development of the DBX Soundfield One loudspeaker (google for more info).

Don Keele, Earl Geddes, Tom Danley, Wayne Parham, Duke Lejeune, and Zilch have discussed these concepts, as applied to horn/waveguide based loudspeakers, in various articles and forum posts..