Tolerancing Primer

Marshall R. Scott

University of Arizona

marshallscott@email.arizona.edu

December 17, 2015

1 Introduction

The goal of the engineer is to design a system that meets a set of requirements while minimizing cost and complexity, knowing full well that once the thing is built everything will be wrong. Take, for instance, the lens shown in Fig. 1. In optical design software, each of the lens elements are chosen to have particular focal lengths and refractive indices, with very particular positions all set to some arbitrary level of precision. However, the actual lenses cannot be placed with absolute precision and accuracy, and so the real-world lens will not be as designed. Each lens will be shifted by some small amount from the optimal position and will be tilted with respect to the optical axis which will compromise the performance. To make matters worse, the focal length, diameter, and thickness of each lens

Figure 1: Cutaway of a Canon EF600mm F4L IS II USM lens. Photo by Morio[1].

will not be as designed, and the index of refraction will be wrong as well. The key, then, to designing a system is to carefully control how ‘wrong’ each of these parameters may be. The naive engineer may simply constrain tolerances to be very tight, but this can be difficult for the machinist, and generally results in a very expensive system. He or she may include in the assembly procedures a tedious and intricate alignment process where the position of each element is carefully measured and aligned by hand, but again the added complexity will increase the required labor, and increase the cost.

In general, a well-designed system has balanced cost with performance, utilizing the simplest design that still meets the requirements, and finding this optimum requires tolerance analysis. In this paper, I outline the very basics of the tolerancing process. First, a quantitative figure of merit (FOM) is chosen the reflects the performance requirements of the system. Then, a list is made of each parameter that could impact the performance. The effect of each of these parameters is calculated or simulated, and the system performance is modeled assuming loose tolerances. The tolerances on each parameter are tightened or loosened as required, and if necessary the design or requirements are modified. The process is iterated until the optimum design is found.

While the examples here will be optomechanical in nature, the principles may be applied to any system where uncertainties in components affect overall performance of an assembled system. An amplifying circuit for a photodiode is a good example. Such a system will have uncertainties associated with the values of resistors and capacitors and other components that will make the overall performance of the circuit uncertain. The concepts are the same.

2 Figure of Merit

The first step in the tolerancing process is to define the FOM, Φ, a number that reflects the performance of a system. For an imaging system, this could be the RMS spot size or MTF. In a fiber optic system, this could be the coupling efficiency between two fibers. In mechanics it may be the drift in position of a part, and in electronics it could be the gain or bandwidth of a circuit. The choice is limited by imagination, but should be simple, and should reflect the requirements of the system.

3 Parameters

Next, a list is made of all the things in the system that could go wrong. In an imaging system, this could be the focal length, diameter, thickness, or index of refraction of each lens element, or the positioning and tilt. In electronics, this could be the values of each capacitor and resistor. Other parameters could be related to the environment, like temperature, which can change the focal length of lenses, dimensions of mechanics, and values of electronic components. The FOM will be a function of this set of parameters $\{x_i\}$,

$$\Phi(\vec{x}) = f(\vec{x})\,. \tag{1}$$

4 Statistics

Since the values of the parameters cannot be known exactly, each one will have some uncertainty, Δx_i, and this will propagate to an uncertainty in the FOM, $\Delta\Phi_i$. Then as long as the set of parameters are independent from one another, the overall uncertainty in the FOM can be calculated by taking the root-sum-square (RSS) of the set of $\{\Delta\Phi_i\}$,

$$\Delta\Phi = \sqrt{\Phi_0^2 + (\Delta\Phi_1)^2 + (\Delta\Phi_2)^2 + \cdots}\,, \tag{2}$$

where Φ_0 is the residual uncertainty in the FOM, the value when there is no uncertainty in the parameters $\{x_i\}$ (due to diffraction, for instance, the RMS spot size of an imaging system will never be zero). The central limit theorem states that under certain conditions, the sum of many random variables follows a Gaussian distribution – the bell curve – regardless of how the individual parameters are distributed. This is not always strictly true, but it is usually a very good approximation for the purposes of tolerancing. If we assume that the $\Delta\Phi_i$ values are 2σ values – a sometimes pessimistic assumption – then the $\Delta\Phi$ in Eq. 2 will be the 2σ uncertainty in the overall system performance. This means that we can expect that 95% of systems built will fall within $\pm\Delta\Phi$ of our designed FOM. Similarly we could $\pm 1\sigma$ values which would would include 68% of builds, or $\pm 3\sigma$ values which would include 99.7% of the builds. Typically 2σ values are used.

As an example, consider 3 3 mm spacers, each with a thickness tolerance of ± 0.1 mm. We would like to put these together to get a 9 mm spacer. Clearly, an uncertainty of Δx in one of the spacers will result in an uncertainty in the total thickness $\Delta\Phi_i$ of the same amount. Using equation 2,

$$\Delta\Phi = \sqrt{\Delta x^2 + \Delta x^2 + \Delta x^2} = \sqrt{3}\Delta x\,. \tag{3}$$

Thus, the total uncertainty in the thickness of the combined spacers is about 1.7 times the tolerance on each spacer. In general, when N uncertainties of about the same value Δx are added in RSS, the resulting uncertainty is $\sqrt{N}\Delta x$.

Another important consequence of Equation 2 is that big terms tend to dominate the sum since they are first squared. For example, if two uncertainties of value 1 and 10 are added, then the root sum square is 10.04. The uncertainty of 1 has increased the total uncertainty by only half a percent.

5 Sensitivities

In order to efficiently tolerance a system, we need to calculate the relationship between the tolerance on each parameter, Δx_i, and its effect on the merit function, $\Delta\Phi_i$. This is easiest when we make the assumption that figure of merit is linear with respect to each parameter. This is typically a good approximation as tolerances in the parameters are small, though it is good practice to verify that the parameter is sufficiently linear over the expected tolerance range of each parameter. If that is the case, then the sensitivity for a parameter x_i is given by the partial derivative of the merit function with respect to the value of the parameter, $\partial\Phi/\partial x_i$. Then the effect on the merit function can be calculated using

$$\Delta\Phi_i \approx \frac{\partial\Phi}{\partial x_i}\Delta x_i\,. \tag{4}$$

If we know the equation, then it is straight forward to apply Eq. 4 to find the sensitivity, but for more complex systems like a photographic lens, this can be difficult. Instead, we can use a model from a simulation program such as Zemax to do the calculation. Programs like these usually have a tolerance feature that can be very useful, but it important to really understand what is going on, or it may give wrong or misinterpreted results. The sensitivity for a given parameter may be estimated by perturbing it by some small amount, and then recording the effect on the merit function. Then the sensitivity may be calculated using

$$\frac{\partial \Phi}{\partial x_i} \approx \frac{\Delta \Phi}{\Delta x_i} \tag{5}$$

When the residual merit function Φ_0 is much smaller than the perturbed merit function, then it may be ignored and $\Delta\Phi \approx \Phi(x_i + \Delta x_i)$. If it not small, and it is correlated with with the perturbed parameter, then $\Delta\Phi = \Phi(x_i + \Delta x_i) - \Phi_0$. In the more common case where Φ_0 is independent of the perturbed parameter, then the net effect on the merit function is the RSS, $\Phi(x_i + \Delta x_i) = \sqrt{\Phi_0^2 + (\Delta\Phi_i)^2}$, which can be inverted to find $\Delta\Phi = \sqrt{\Phi(x_i + \Delta x_i)^2 - \Phi_0^2}$.

Fig. 3 shows how Excel can be used to calculate sensitivity of the doublet shown in Fig. 2. The tilt of the first lens is changed in 1 mrad steps, and the figure of merit is computed for each. The results may be plotted to verify linearity, and the slope of the line is the sensitivity.

In some cases, a model may not be available and an experiment is required to find the sensitivity. The procedure works the same way, except parameters are perturbed in a real system.

6 The tolerance table

When the parameters and sensitivities have been calculated, the tolerance table, Table 1 may be constructed in a spreadsheet such as Excel. Loose initial tolerances should be chosen using rules of thumb and knowledge of the fabrication process.

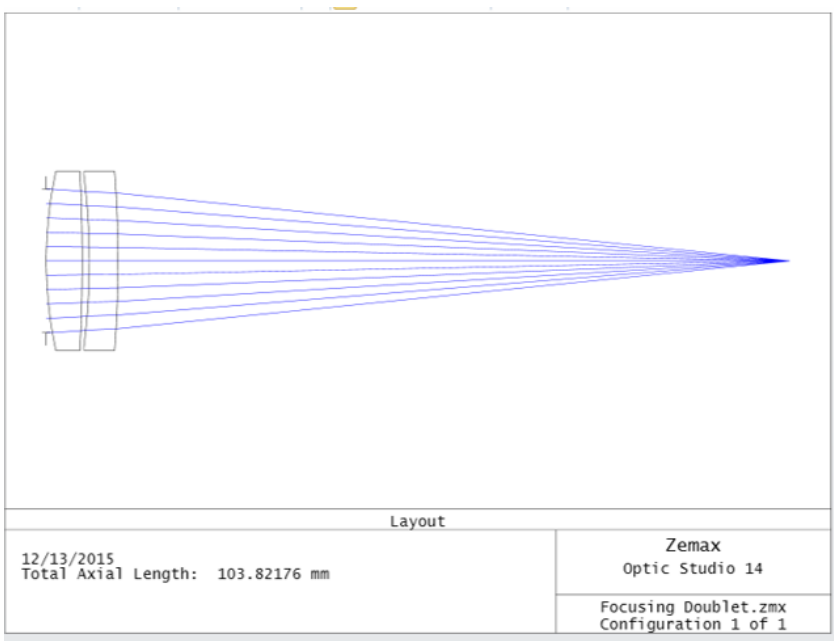

Figure 2: A simple doublet requiring tolerance analysis.

mrad	degree	waves rms	
Perturbati	Perturbation	WFE	slope
0	0	0.002	
1	0.05729578	0.012	0.01
2	0.114591559	0.023	0.011
3	0.171887339	0.034	0.011
4	0.229183118	0.046	0.012
5	0.286478898	0.057	0.011

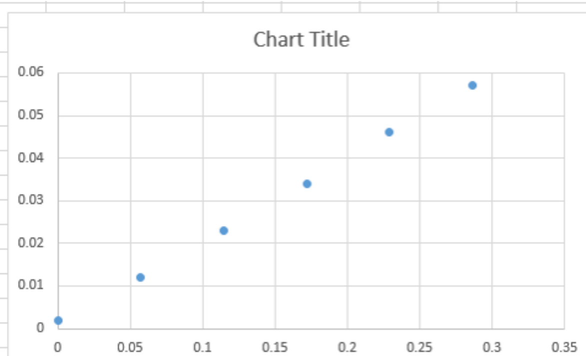

Figure 3: Excel is used to quickly calculate sensitivities and verify linearity.

Table 1: The tolerance table

Parameter	Tolerance	Sensitivity	Effect on merit function
x_1	Δx_1	$\partial\Phi/\partial x_1$	$= \Delta x_1 \cdot \partial\Phi/\partial x_1$
x_2	Δx_2	$\partial\Phi/\partial x_2$	$= \Delta x_2 \cdot \partial\Phi/\partial x_2$
x_3	Δx_3	$\partial\Phi/\partial x_3$	$= \Delta x_3 \cdot \partial\Phi/\partial x_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
RSS			=sqrt(sumsq(·))

Sensitivies are loaded in the third column, and the effect on the merit function is automatically computed in the fourth column, and finally the root-sum-square is calclulated. The tolerances may now easily be adjusted until the RSS meets the required specification. Tolerances whose effect on the merit function is high compared to the others should be tightened, and those whose effect is small should be loosened. A well tolerenced system will exhibit similar effects on the merit function from each parameter.

An Excel assembly tolerance table for the doublet lens in Fig. 2 is shown in Fig. 4. In this case a compensator was used–the position of the detector was allowed to be adjusted to compensate for focusing errors. When compensators are used, they must be accounted for when calculating the sensitivities. Therefore, when each parameter was perturbed, the focus position was adjusted to simulate the focus adjustment. Then the resolution of the adjustment appears in the tolerance table, in this case as ‘Detector Axial Position’. The RSS shows that performance of the system is better than $1/20\lambda_{rms}$ which meets the requirement. The tolerances are reasonable, and the values of the effect on merit function are all close in value.

		mm or mrad	waves rms/mm or waves rms/mrad	waves rms
	Parameter	Tolerance	Sensitivity	WFE
	Residual			0.002
Lens 1	Tilt	1.5	0.008	0.012
	Decenter	0.1	0.2	0.02
	Lens spacing	0.5	0.018	0.009
Lens 2	Tilt	1.5	0.011	0.0165
	Decenter	0.1	0.2	0.02
Detector	Axial position	0.005	2.3	0.0115
RMS				0.03786159

Figure 4: Tolerance table for assembly of a simple doublet (Fig. 2.

7 Conclusion

Tolerancing is a complex, yet essential, skill in engineering that becomes easier with practice. It allows us to predict how a system will perform when it is actually built, and allows us to minimize the cost of the design. The principles listed herein are very basic. Each field, and each application, will pose its own challenges. Sources for further reading are listed below.

References

[1] File:canon ef600mmm f4l is ii usm cutaway 2012 cp+.jpg. `https://commons.wikimedia.org/wiki/File:Canon_EF600mmm_F4L_IS_II_USM_cutaway_2012_CP%2B.jpg`. Accessed: December 14, 2015.

[2] Jim Burge. Tolerancing optical systems, opti521 lecture notes, 2015.

[3] Warren J Smith. Fundamentals of establishing an optical tolerance budget. In *1985 Los Angeles Technical Symposium*, pages 196–205. International Society for Optics and Photonics, 1985.

[4] Robert H Ginsberg. Outline of tolerancing (from performance specification to toleranced drawings). *Optical Engineering*, 20(2):202175–202175, 1981.

[5] Richard N Youngworth and Bryan D Stone. Cost-based tolerancing of optical systems. *Applied optics*, 39(25):4501–4512, 2000.

[6] Richard N Youngworth. Twenty-first century optical tolerancing: a look at the past and improvements for the future. In *Contract Proceedings 2006*, pages 634203–634203. International Society for Optics and Photonics, 2006.

[7] Jessica DeGroote Nelson, Richard N Youngworth, and David M Aikens. The cost of tolerancing. In *SPIE Optical Engineering+ Applications*, pages 74330E–74330E. International Society for Optics and Photonics, 2009.

[8] Simon Magarill. Optomechanical sensitivity and tolerancing. In *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pages 220–228. International Society for Optics and Photonics, 1999.

[9] JH Burge. An easy way to relate optical element motion to system pointing stability. In *SPIE Optics+ Photonics*, pages 62880I–62880I. International Society for Optics and Photonics, 2006.