Simulation of Process-Stress Induced Warpage of Silicon Wafers Using ANSYS® Finite Element Analysis

Aditi Mallik and Roger Stout
ON Semiconductor
5005E McDowell Road
Phoenix Az 85008

Abstract

Wafers warp. It is important to minimize warpage in order to achieve optimal die yield and potentially prevent future device failure. Although the word warpage is widely used in the literature to represent wafer bow (convex or concave shape), in the real world wafers are often seen into warp into saddle shapes. This complicates the characterization of both the sources of and solutions to warpage, because (as will be discussed) Stoney's formula (relating intrinsic stress and curvature) does not apply for structures warped with compound curvature, and standard wafer warpage measurements are not designed to measure compound curvature. During thin film deposition, wafer warpage occurs due to the intrinsic stresses and the coefficient of thermal expansion (CTE) mismatch of the different thin films and the substrate. Unfortunately, whereas the introduction of the thermal stresses due to CTE mismatch into a finite element model is easily understood, the introduction of intrinsic stress is not. Further, although a saddle shape is clearly a physically realizable (indeed, often preferred) equilibrium configuration for a circular disk (consistent with an appropriate state of stress), obtaining a saddle shape in a finite element solution turns out to be extremely difficult, as convex or concave shapes may also be stable and numerically preferred. In this paper, a finite element technique (using ANSYS software) to model wafer warpage is presented. Simulations have been done for silicon wafers with aluminum or standard UBM films on top. Saddle-shaped warpage has been successfully modeled, and the aggravating effects of thinning (back side grinding) have been reproduced.

Key words: Wafer warpage, wafer bow, saddle shape, wafer backgrinding

I Introduction

As electronic devices continue to shrink in size, the IC must be reduced in both footprint and thickness. This drives the semiconductor industry to produce thinner and thinner wafers. One of the major drawbacks of wafer thinning is increase in wafer warpage and fragility. It is important to minimize warpage in order to achieve optimal wafer deposition and die yield. Excessive wafer warpage can also potentially lead to die failure. Wafers warp to some extent during the deposition of the thin films; warp is then further highly aggravated during the thinning procedure. It is known that during thin film deposition, wafer warpage occurs due to the intrinsic stresses and the coefficient of thermal expansion (CTE) mismatch of the different thin films and the substrate. Unfortunately, whereas the introduction of the thermal stresses due to CTE mismatch into a finite element model is easily understood, the introduction of intrinsic stress is not.

In this paper a method to simulate saddle shape wafer warpage using the Finite Element (FEM) software, ANSYS® Version 12.0, is shown. There have been a few papers on simulation of bowing of wafers [1, 2] but to the best of the authors' knowledge there has been none of simulating wafer warping (by which we mean saddle-shapes specifically). Though the word warping has been used in these papers, what it actually represents is

wafer *bow* (convex or concave shaped wafers) and not saddle-shaped wafers. Vodenitcharova [3] *et al*. in their paper on 'the effect of thermal shocks on the stresses in a sapphire wafer' talk about observing a saddle shaped wafer on insertion of the wafer in the furnace and a bowl shaped wafer on withdrawal from the furnace. However no simulation on obtaining a saddle shaped wafer has been reported. Moreover, in their paper, the graphs showing the radial and circumferential stresses distribution doesn't resemble that for a saddle shaped wafer. There are few papers on obtaining analytical equations for the curvatures for saddle shaped wafers [4, 5]; however the pictures in all these papers represent bow shaped wafers.

We found that it is extremely difficult to produce a saddle shape in ANSYS as convex or concave shapes may also be stable and numerically preferred. In this paper, we discuss the ANSYS tricks that were performed to simulate saddle shaped wafer warpage.

II Procedure

We have reported our initial approaches and finally the method that leads to a saddle shaped wafer in the Appendix A, where we have also discussed the ANSYS techniques that seemed to be the correct approach, however failed to give a saddle shaped wafer. We think discussing all these approaches is going to help the readers save their time by avoiding those methods. Figure 1 below shows the ANSYS model of a 6" diameter Si wafer with a 5.4" diameter Al film on it. The hatched portion represents the Al film. The wafer is perfectly circular except for a flat of 2.6" at the top. The initial thickness of the wafer is 25 mils where as the Al film layer is only 5 um thick. The wafer is finally back-ground to a thickness of 6 mils. The Si substrate is assumed to be elastic where as the Al film is elastic-plastic.

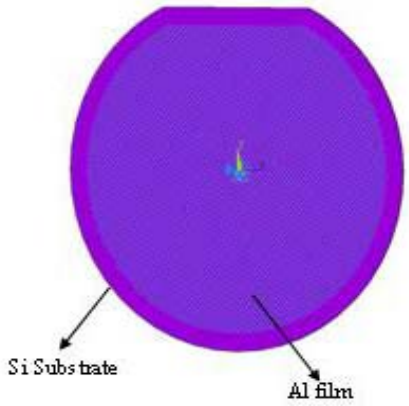


Figure 1. ANSYS model of Si wafer with Al Film; hashed portion represents the thin film.

ANSYS SOLID185 elements with "Enhanced Strain" option are used for the Si Substrate. SHELL181 elements are used for the Al film. A radial temperature distribution is given on both the film and the substrate. This is done by dividing the wafer into six different radial zones as shown in Figure 2. Each zone has the same material property except the stress-free reference temperature [TREF] which differs from the adjacent ring by 20 ºC. The innermost ring has the highest reference temperature of 410 ºC while the outermost ring has 310 ºC.

This could be a plausible assumption as the center of the wafer is almost certain to have a different thermodynamic equilibrium with the surrounding sputtering chamber than the edges. Deposition times are long enough to ensure z-axis temperature uniformity, but could be short enough *not* to ensure radial temperature uniformity.

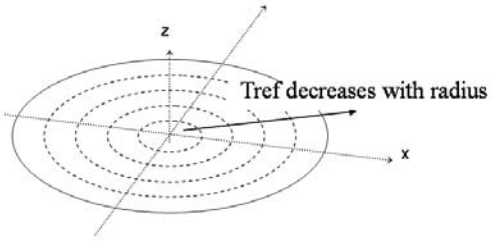


Figure 2. Wafer divided into 6 radial zones with increasing Tref.

We summarize the five solution steps in ANSYS that lead to a saddle shape:

1: Set each material to its own TINIT=TREF.
2: Apply the symmetry-breaking forces at ends of diameters.
3: Ramp (Cool) the temperature to the desired final uniform value (ambient).
4: Solve while ramping forces to 10% of initial value.
5: Solve again after removing the forces completely.

Finally, element "ekill" over Si elements is issued within the bottom 19 mil layer to simulate the back-grinding process. It was seen that the symmetry breaking forces were necessary to force the solution into a saddle shape in the initial state, and then have to be removed gradually. If the forces are "step removed" ANSYS failed to converge. Interestingly, and contrary to what may be conventional wisdom, the flat in the wafer only determines the preferential direction a wafer would bend of various directions possible. The significance of the flat is further explained in Appendix A.

Figure 3(a)-(d) shows the solution results obtained

from step 1 through 5 mentioned above. Figure 3(a) shows the radial temperature distribution when each zone has been raised to its Tref temperature. Figure 3(b) through 3(d) shows the out of plane or the UZ displacement of the wafer as the symmetry breaking forces are applied, cooled to ambient temperature, and finally ramping down the forces to zero, respectively. It is seen that even after the forces are removed to the wafer remains in the saddle shape. Finally, Figure 3(e) shows the out of plane or the z displacement after 19 mils of Si has been taken away. We note that the warpage more than doubles after the wafer has been cooled to room temperature, i.e. from step 3 to step 2. We saw similar results with standard UBM TiNiAg films. Multi-layer SHELL181 elements were used to simulate three-layer films. The results show that wafer warpage increases (the maximum deflection by 8 times) after backgrinding.

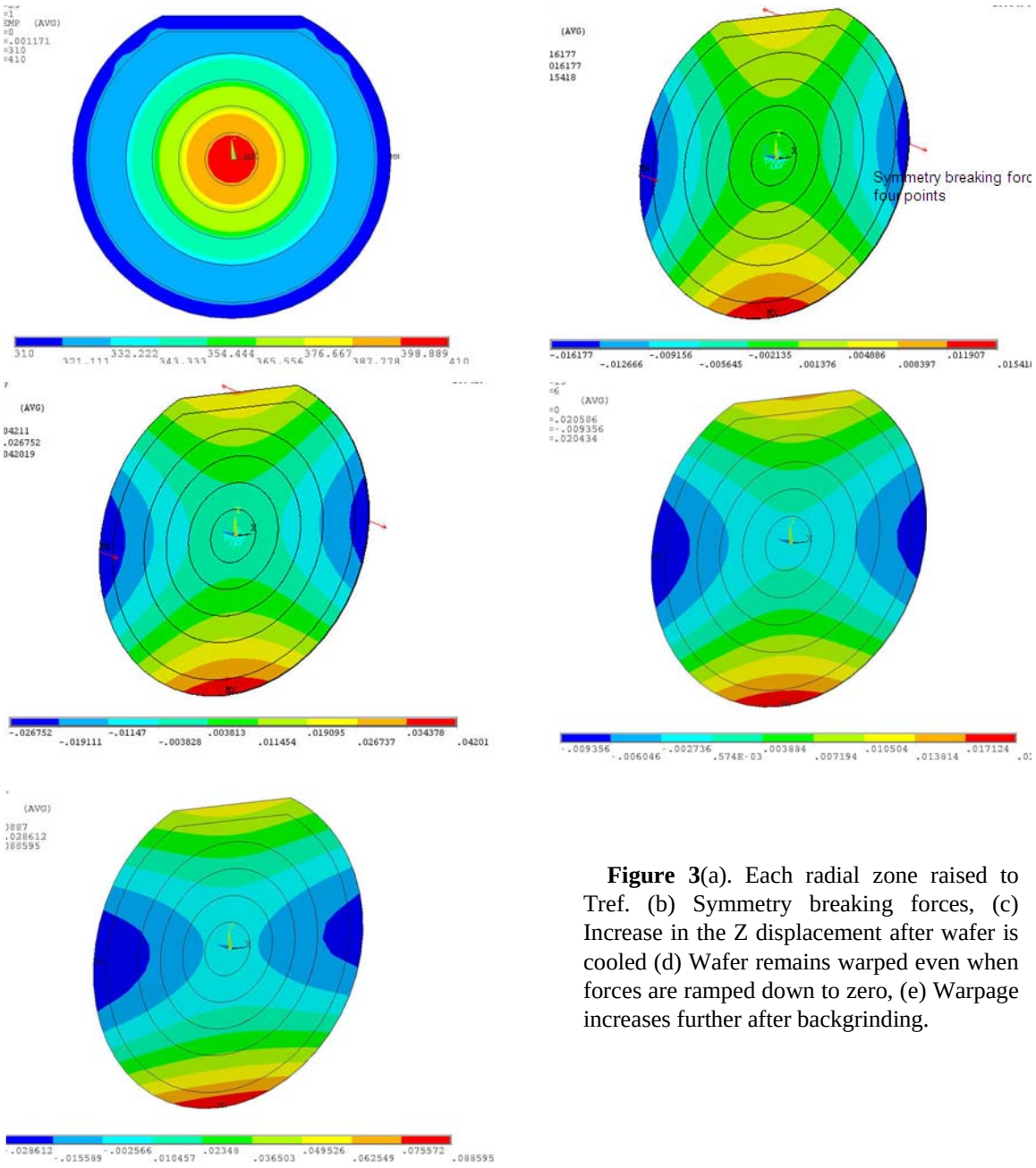


Figure 3(a). Each radial zone raised to Tref. (b) Symmetry breaking forces, (c) Increase in the Z displacement after wafer is cooled (d) Wafer remains warped even when forces are ramped down to zero, (e) Warpage increases further after backgrinding.

II Intrinsic Stress in Thin Films

As discussed earlier, a wafer warps because of the intrinsic stresses in the thin films. In this section simulation of warpage due to intrinsic stresses is shown. ANSYS has the option of applying an initial stress using the INISTATE command. One has to be careful that this inistate can be used only in the first load step and not in between a sequence of several load steps. Since the stresses at the last step (when the forces are taken away but before the back-grind step) of the previous model can be easily extracted from ANSYS postprocessor, this stress condition can be applied as an initial stress using the INISTATE command for a model instead of creating the radial bands of materials having different Tref. The stress in the previous solution can be written out to a file by issuing the command "INISTATE, WRITE, 1," before the solve command. Now for the model where the Tref is same throughout the whole wafer, we read this inistate from the original model in the first load step using the "INISTATE, READ" command. As a check we also write out the inistate to another file and compare to original input *.ist* file. It is found that the stresses on the elements are similar; however they are not exactly the same. This was due to the fact that in this case when in the initial step, the wafer is perfectly flat; but when the *.ist* file is written out in the previous model; the wafer already had some finite displacements. All the load steps are solved at ambient temperature (25 °C). In the next step we apply the symmetry breaking forces. However, we have to remember to reverse the direction of the forces, as inistate acts like an internal stress so the displacements would be in opposite direction with respect to the old model. The rest of the steps are same as before. We gradually remove the forces and find that the wafer remains in the warped shape.

Now, it would be of real interest to know if the model wafer will warp into a saddle shape solely by applying inistate to just the thin film. I.e. the wafer substrate doesn't have any initial stress of its own, but if the thin film imparts sufficiently large stress, can it cause the composite structure to warp? In the original model with different Trefs, we observe that the stresses in the plane of the thin film in X and Y directions (i.e. SX and SY stresses) are highly radial in nature ranging from 36 ksi to 46 ksi. The stress contours on the film are shown in Figure 4; also the out of plane stress component (i.e. SZ) is nearly zero. We next applied an inistate SX and SY on the shell elements as a radial function of the position of the centroid of the element which matches with the stress contour of Figure 4.

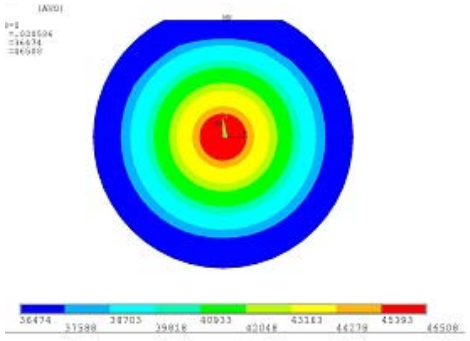


Figure 4. SX or SY stress contour on the thin film

As a check we wrote out inistate after issuing the solve command; however we found that the inistate on the shells were much less and ranged from only 15 ksi to 20 ksi. Increasing the magnitude of the stress or thickness of the shells to any amount failed to produce a saddle wafer. We also tried constraining the four nodes at 45° angles on the perimeter in the z direction; but the wafer wouldn't remain saddle shaped after the forces were removed. We concluded that it is absolutely necessary to have the stress on the Si substrate (the underlying solid elements), either through a radial temperature distribution or inistate command, in order to obtain a saddle shape.

The literature is replete with studies where researchers have used Stoney's [6] formula to calculate stresses of thin films, viz:

$$\sigma_f = \frac{E_s t_s^2}{6(1-\upsilon_s)t_f}(\kappa_e - \kappa_i) \qquad (1)$$

where σ_f is the biaxial thin film stress, t_f is the thickness of the film. E_s, υ_s and t_s represent the Young's modulus, Poisson's ratio and the substrate thickness. κ_i and κ_e represent the curvature of the wafer before and after the deposition respectively.

But, we concluded that the above Stoney's formula for intrinsic stresses cannot be applied to the warped or saddle shaped wafers.* It is not fully clear if the

* Stoney's original derivation was actually even somewhat simpler, based on uniaxial stress, linear beam theory; the commonly used equation quoted has been extended to uniform biaxial stress. A planar membrane, however, *may* have a much more complicated relationship between stress and curvature, even in a uniform disk [4]. A simple thought experiment would lead one to conclude that if Stoney's formula applied to saddle-shaped wafers, one could quickly identify lines of zero stress by locating the lines of zero curvature; similarly, maximum stress would appear in directions perpendicular to these lines of zero curvature (i.e. maximum curvature), and would alternate in sign from quadrant to quadrant of the wafer. Yet there is no reason to believe

formula gives the magnitude of average film stress, and as seen from the simulation results, the wafer alternates between regions where curvature in two perpendicular directions (say tangential and radial) changes both in sign and orientation whilst the stress itself is constant between said regions. This implies that some experimental method other than simple correlation between stress (in a predefined direction) and curvature (in that direction) is required to determine film stresses. As we have as yet been unable to find such methods, further experimental verification of our models has been deferred. Basically, we have concluded that if you have a convex or concave wafer, with similar magnitudes of curvature in perpendicular directions anywhere on the surface, you can infer film stress from Stoney's formula. If you have a saddle-shaped wafer, you're out of luck.

Conclusions

We have demonstrated a way to simulate wafer warpage in ANSYS. We emphasize again that warpage here refers to saddle shaped and not bowl (convex or concave) shaped, as warpage is often interpreted as bowl shape in the literature. Our simulations also show that the warpage increases with the back-grinding process as seen in the real world. This has been demonstrated with Al film as well as standard TiNiAg films. In the Appendix A of our paper, we have also discussed the ANSYS approaches which seemed correct, however failed to produce a saddle shaped wafer.

References

[1] S. Irving and Y. Liu, "Wafer Deposition/Metallization and Back Grind, Process-Induced Warpage Simulations", *Electronic Components and Technology Conference*, 2003, pp1459-1462.

[2] R.B.R van Silfhout, W.D. van Driel, Y.Li, G.Q. Zhang and L.J. Ernst, "Prediction of Back-End Process-Induced Wafer Warpage and Experimental Verification", *Electronic Components and Technology Conference*, 2002, pp1182-1187.

[3] T. Vodenitcharova, L.C.Zhang, I. Zarudi, Y.Yin, H. Domyo and T. Ho, "The Effect of Thermal Shocks on the Stresses in a Sapphire Wafer", *IEEE Trans. On Semiconductor Manuf.*, vol 19, no 4, 2006, pp449-454.

[4] K.A. Seffen and R. A. McMahon, "Heating of a Uniform Wafer Disk", *International Journal of Mechanical Sciences*, vol 49, 2007, pp230-238.

[5] Y. –L. Shen, S. Suresh and I.A. Blech, "Stresses, Curvatures and Shape Changes Arising from Patterned Lines on Silicon Wafers", *J. Appl. Phys, vol 80(3), 1996, pp1388-1398*

[6] G.G. Stoney, "The Tension of Metallic Films Deposited by Electrolysis", Proc. Roy. Soc. London A Meth, 82, 1909, pp172-175.

[7] W.C.Young, "Roark's Formulas for Stress and Strain", 6th edition.

APPENDIX A

A 150 mm diameter wafer only 0.5 mm thick may easily warp 3 mm or more. When considering how to model such warping, even though the deflections may be locally large compared to the initial undeformed geometry, the slopes are small and it was not entirely clear that a linear analysis wouldn't yield reasonable results. Nevertheless, *Roark & Young* [7] have shown that any time the deformations exceed half a membrane's thickness, it needs to be handled as a large deformation problem, and typical warping clearly falls into this category. Therefore, throughout our wafer modeling efforts, we turned on ANSYS' "NLGEOM" switch (non-linear geometry, meaning large deformation analysis).

Our initial approach, because we were thinking that the problem directly related to film stress (and film thickness and film CTE) differing from the silicon substrate, was to model a bi-material wafer, one material being the bulk, and the other being a thin film (i.e. shell elements). But as we mentioned in the introduction, no matter what we did, all we'd achieve were convex or concave shapes (Figure A1). We tried all sorts of variations of CTE, stress free temperatures (TREF), initial stress inistate's, etc. We even tried creating an initially deformed geometry (a saddle shape with small amplitude), but still ANSYS gave a concave/convex displacement field superimposed on the initial geometry.

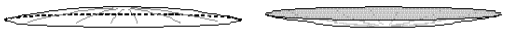


Figure A1: Convex or concave deflection field

We then made some paper models of wafers (disks) with wedge shaped overlaps or wedge shaped insertions, and realized the saddle shape arises easily simply if you add area faster at the outer diameter than you do at the center as shown in Figure A2. Note that if you remove material faster at the perimeter, as in Figure A3, you get the proverbial

that films, nominally uniformly applied over the surface of a wafer, should magically build up quadrants of positive and negative intrinsic stress; nor further, indeed, that the film stress should be zero along particular radial lines. In fact, our simulations of saddle shaped wafers clearly show that stresses are actually locally isotropic everywhere, that in-plane stresses are everywhere of the same sign, and that they vary, in fact, only in the *radial* direction. These conditions are completely consistent with common-sense understanding of the nature of the intrinsic film stresses believed to arise during film deposition, and coexist with the obviously highly varying local curvature from quadrant to quadrant. In other words, we have shown clearly that *simple* curvature is essentially independent of the intrinsic planar (biaxial) film stress.

Chinese-hat shape – an approximation of the convex/concave solutions.

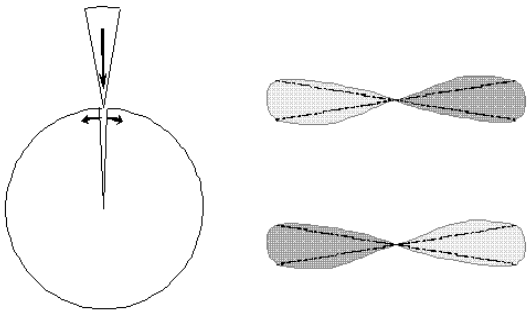


Figure A2: Convex or concave deflection field

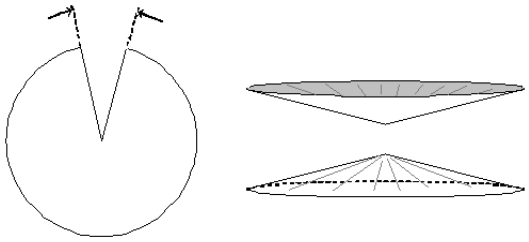


Figure A3: Remove a wedge of material from a flat disk; result is the "Chinese hat" – also a bi-stable geometry

Temporarily setting aside the bi-material wafer problem, the question now arises as what needs to be done to make ANSYS distort a uniform, flat, disk into a saddle shape? Taking this simpler approach, we immediately discovered (perhaps unsurprisingly) that if you start with a flat wafer, and are doing purely thermal loading, ANSYS has no reason to generate out-of-plane deformations at all. (The irony is that at least with the original bi-layer model, we had non-planar deformations, even if they weren't what we were looking for!) Indeed, if one takes a uniformly thick wafer and varies the CTE from center to outer diameter, and then changes the temperature from TREF, ANSYS shows that the wafer will get thinner in the middle and thicker at the edge, but it will never *warp* out of plane at all. (Depending on the relative "polarity" of the CTE gradient vs. the temperature change from TREF, you can get thicker in the middle and thinner at the edge, but again, you won't get any warp.) These results are depicted in Figure A4 (a)-(b).

But again, we know that in the real world, you can easily create a saddle shape, and at least in real wafers (and thin paper disks), it tends to be bi-stable (in that you can take the originally warped wafer and simply by pressing your fingers down on the two high edges, it'll pop through to an oppositely symmetric, stable, warped configuration). So we started adding small forces at the edges. It turned out the best way to get out-of-plane saddle shapes was to apply +Z forces at the ends of the x-diameter, and –Z forces at the ends of the y-diameter as shown in Figure A5.

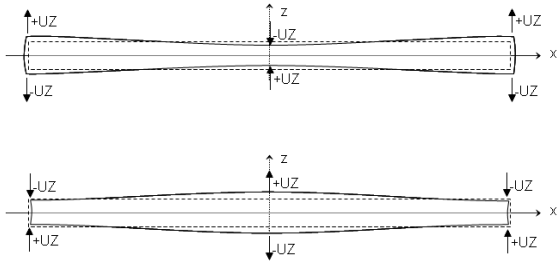


Figure A4: Wafer shape with nothing to break the Z-symmetry (a) if solution temperature is above TREF (b) if solution temperature is below TREF

If we applied just two forces, e.g. one up at +x, and one down at +y, we tended to have convergence problems. If we simply did two +Z forces on a single diameter, we'd get Chinese hats again. So having at least one up and one down seemed necessary, and having two up and two down converged pretty well. (Concern about displacement constraints possibly messing us up was never an issue. Throughout this whole process our displacement constraints were always very carefully applied at the origin – UX=UY=UZ=0 at the origin, UY=UZ=0 one node out on the x-axis but as near the center as we could get, and UZ=0 one node out on the y-axis as near center as we could get. And of course, once we started to apply forces to the model, it was mandatory to constrain the model in displacement somewhere.)

Here's where it really got interesting, because we know that real wafers are stable in the warped geometry even with no external forces applied. Our simplified approach to this point was basically to define the model, set CTE's in zones that changed with radius (Figure 2), set a common TREF for everything, and then with load step 1, solve the model (with those forces applied) at some temperature that should be consistent with the saddle shape (i.e. higher than TREF, with CTE's that increased with radius, thus *adding* material faster with increasing radius). With load step 2, we'd then remove the forces, with the hope that the warp would reduce slightly, but not go away. Unfortunately, all that would happen was that as the load would go away (if it converged at all, which sometimes it wouldn't), the deflections would also gradually go away, until we'd end up again with that flat wafer that was thicker at the edge and thinner in the middle

(Figure A5) - except often it wouldn't converge at all. Then we learned that even when you have KBC=ramp, a *deletion* of forces is applied as a *step*, so to truly *ramp* from a big value down to zero, you may have to do it in two steps (one to ramp it to some value that's non-zero but small enough, and the second to then delete that final little bit of force). Then we noticed something odd – the converged solutions after the first load step (with the forces applied), were always like those shown in Figure A5, where the stable solution is actually with the forces *countering* the curvature of the deformation field! Of course, once we saw this, we realized that this is quite plausible physically, because you've really just got two sets of arches that can readily support compression loads and are simultaneously providing equilibrium reactions for each other. In our real world of wafers, however, this is not likely going to be the result (and indeed, we're looking for something like Figure A5c). The solution to this dilemma is to introduce the saddle-shaping mechanical pre-load first, and impose the temperature changes second. This order "locks" the deformations into the A5c configuration, so that even when the pre-loading forces are then (carefully) ramped away, the saddle shape remains as desired.

Having now proven that we can reproduce in a relatively simple and straightforward ANSYS model something very similar to what we observe experimentally, we set about to make the problem setup a little more realistic. First off, we don't think it's likely that thin films of material on the surface of a wafer are going to have non-uniform CTE's. They don't even have enough variation in thickness (perhaps only 15% or less) to have the *effect* of having non-uniform CTE's – mainly because the films are already a small fraction of the thickness of an otherwise fairly uniform substrate. (e.g. 3E-6 mm of film, ±5E-7 mm, on top of a 1E-4 substrate.) We also don't believe that things like crystal orientation in the substrate, or grain orientation, etc. of the deposited layers, are a first-order effect. All these things may have a second order effect, which in particular may dictate the particular theta-orientation of the most stable saddle-shape. (One thing we've omitted, for simplicity in the preceding exposition, is that our wafers aren't perfect circles. They are perfect circles *except* they have a small "flat" along one edge, typically at the +y axis as the pictures are oriented. We found that the location of the flat is quite sufficient to force the saddle shapes to be symmetric in x and y. Attempts to solve with the symmetry-breaking forces applied along axes rotated 30° or 45° from the flat simply resulted in longer solution times with convergence challenges, and when they did solve, the extrema of displacement would not coincide perfectly with the forces; further, when the forces were then removed, and subsequent load steps converged at all and after much difficulty, the extrema would have rotated to coincide with the flat and the x-axis. Indeed, it stands to reason that with a perfectly axisymmetric wafer, there should be an infinite number of stable saddle shapes, and it's not clear how ANSYS would behave. That's a problem we're not interested in solving!)

So the more realistic, plausible problem statement is to presume that it's TREF that changes during the film deposition process. For instance, the center of the wafer is almost certain to have a different thermodynamic equilibrium with the surrounding sputtering chamber than the edges. Deposition times are long enough to ensure z-axis temperature uniformity, but far too short to ensure radial temperature uniformity. (By extension, it's not even clear that there are necessarily *circular* isotherms; however, our preliminary work showed that saddle shapes are at least consistent with circular, and not radial, isotherms.) We do know that the deposition process takes place hundreds of degrees above room temperature, and that when the wafers cool to room temperature, they're warped into saddles. Unfortunately, in changing the model construction to impose radially varying TREF's with uniform CTE's (in contrast to the previous radially varying CTE's with uniform TREF's), we found that ANSYS insisted on again defaulting to solutions where the preload forces *opposed* the warp direction, instead of amplifying the warp. This meant introducing yet one-more artificial step, in recognition that *any uniform* temperature field (above, below, or at average of TREF's) results in the wrong solution (forces oppose deflections), where each band of material is first "solved" to its own unique TREF value before the mechanical forces are applied. We have in the end, the following *five* load step technique for creating stable, warped (saddle-shaped), wafer solutions:

1: create circular bands of materials, setting each material to its own TINIT=TREF and solve with no forces applied

2: apply the symmetry-breaking forces at ends of diameters and solve

3: ramp the temperature to the desired final uniform value (e.g. ambient)

4: solve while ramping forces to 10% of initial value

5: solve again after removing the forces completely

Figure A5: Coordinate definitions and applied forces

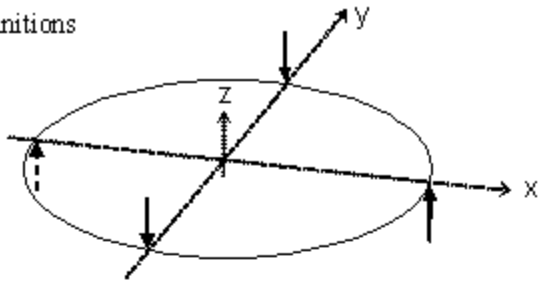
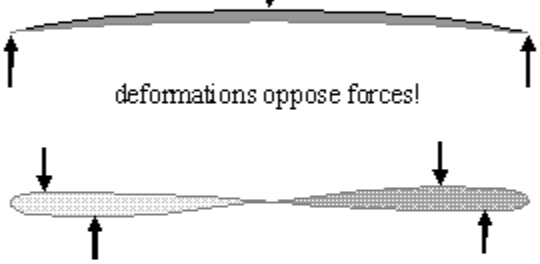


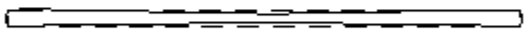
Figure A5b: Temperature and forces applied concurrently

this is what you *get*



deformations oppose forces!

and if you remove the forces:



This procedure doesn't work:

set TUNIF to desired final temperature, apply symmetry-breaking forces and solve
2: remove forces and solve

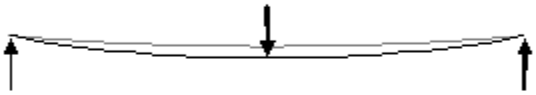
Often doesn't converge on step 2

Even this procedure doesn't work:

1: set TUNIF to desired final temperature, apply symmetry-breaking forces and solve
2: ramp forces to 10% and solve
3: remove final 10% of force and solve

but this is what you *want* ...

Figure A5c: Forces applied first, then temperature field (then forces removed)



deformations have same sense as forces

and if you remove the forces:

This procedure works:

1: Set each material to its own TINIT=TREF.
2: Apply the symmetry-breaking forces at ends of diameters.
3: Ramp (Cool) the temperature to the desired final uniform value (ambient).
4: Solve while ramping forces to 10% of initial value.
5: Solve again after removing the forces