Design Rules and Scaling for Solar Sails

Dr. Glenn W. Zeiders*
The Sirius Group, Huntsville, Alabama, 35801

Useful design rules and simple scaling models have been developed for solar sails. Chief among the conclusions are

1. Sail distortions contribute to the thrust and moments primarily though the mean squared value of their derivatives (slopes), and the sail behaves like a flat sheet if the value is small. The RMS slope is therefore an important figure of merit, and sail distortion effects on the spacecraft can generally be disregarded if the RMS slope is less than about 10% or so.

2. The characteristic slope of the sail distortion varies inversely with the tension in the sail, and it is the tension that produces the principle loading on the support booms. The tension is not arbitrary, but rather is the value needed to maintain the allowable RMS slope. That corresponds to a halyard force about equal to three times the normal force on the supported sail area.

3. **Both the AEC/SRS and L'Garde concepts appear to be structurally capable of supporting sail sizes up to a kilometer or more with 1AU solar flux, but select transverse dimensions must be changed to do so. Operational issues such as fabrication, handling, storage and deployment will be the limiting factors.**

I. Sail Performance

Wrinkles and billow in the sail can reduce the thrust and even produce moments. It should be immediately obvious, however, that such sail distortion is a matter of degree and that a sail should ultimately behave as a flat sheet as the surface slopes become "small". In fact, analytical results shown in Fig. 1 were presented three years ago at the SSDM Conference in Denver (Ref. 1), predicting that wrinkle aspect ratios of as much as several percent should produce minimal effects under most conditions of interest. The findings appear to have gone unappreciated, however, and considerable effort has been spent on modeling and testing of what could be regarded as a "non-problem" since tension can be used quite effectively to flatten the sail (Section II.) It will be shown here, using a perturbation approach that has proved quite useful with optics, that their results can be generalized in terms of the RMS slope as a single "figure of merit" for sail distortion characterization and control.

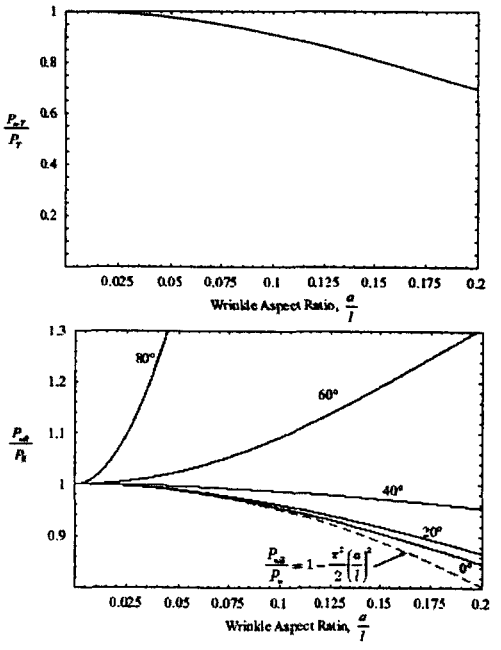

Figure 1: Tangential and radial thrust

First, considering for simplicity a sail with ideal specular reflectivity, McInnis (Ref. 2, Eqn. 2.20) has shown that the force per unit sail area is given by

$$\vec{f} = \sigma \left(\vec{U} \bullet \vec{n}\right)^2 \vec{n}$$

where σ = 9.12 N/km^2 = 2.05 lb/km^2 @ 1AU is twice the solar pressure p, $\vec{U}$ is its unit flux vector, and $\vec{n}$ is the unit normal to the surface. Force components are often chosen relative to the direction of flight or to the incident light (as was done for Fig. 1), but it is more convenient here to reference to a mean flat surface. We shall accordingly choose coordinates such that the mean sail lies in the plane z = 0, let the

* President, The Sirius Group, 2803 Downing Court SE, Huntsville AL 35801

incident flux vector lie in the plane y = 0 such that it is given by $\vec{U} = \vec{k}\cos\theta - \vec{i}\sin\theta$ as in Fig. 2, and let the sail surface be given by W = z – w(x,y) = 0. Thus,

$$\vec{n} = \frac{\nabla W}{|\nabla W|} = \frac{\vec{k} - \vec{i}\,w'_X - \vec{j}\,w'_Y}{\left(1 + w'^2_X + w'^2_Y\right)^{1/2}}$$

so that the total force on the sail is

$$\vec{F} = \iint \vec{f}\, dA = \sigma \iint \frac{(\cos\theta + w'_X \sin\theta)^2}{\left(1 + w'^2_X + w'^2_Y\right)^{3/2}} \left(\vec{k} - \vec{i}\,w'_X - \vec{j}\,w'_Y\right) dA\,,$$

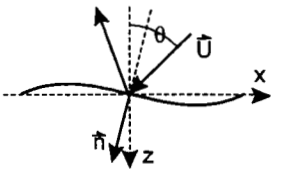

Figure 2: Thrust diagram

which, for small deflections w, can be expanded in series form to give

$$\vec{F} = \sigma \iint [\cos^2\theta + w'_X \sin 2\theta + w'^2_X \sin^2\theta - \frac{3}{2}(w'^2_X + w'^2_Y)\cos^2\theta + \cdots]\left(\vec{k} - \vec{i}\,w'_X - \vec{j}\,w'_Y\right) dA\,.$$

Integrals of the slopes alone tend to vanish for both local distortions (wrinkles and creases) and for symmetric global ones (large-scale billow, etc) symmetry in the force equation, so, defining mean values by $<X> A = \iint X\, dA$, we obtain

$$\vec{F} = \sigma A \vec{k}\,[\cos^2\theta - \frac{3}{2}\{<w'^2_X>(1 - \frac{5}{3}\sin^2\theta) + <w'^2_Y>\cos^2\theta\}] - \sigma A <\vec{i}\,w'^2_X + \vec{j}\,w'_X\,w'_Y> \sin 2\theta + \cdots$$

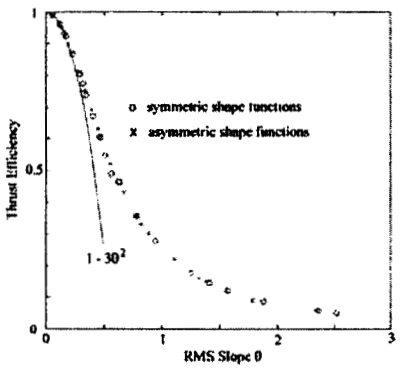

Figure 3: Thrust efficiency vs. slope

where the $\vec{k}$ component is the primary thrust normal to the mean sail plane while the other is the side thrust produced by distortions with oblique solar incidence. The distortions contribute to the force primarily though the mean squared values of their derivatives (slopes), and the sail behaves like a flat sheet if the values are small. This has been verified by a large number of numerical simulations by John Rakoczy of MSFC, and his results, displayed in Fig. 3 for ideal sails at normal incidence, actually correlate with RMS slope over a much wider range than might be expected. The effects of distortions on the normal and side thrusts are shown for various incidence angles θ in Fig. 4 in terms individually of the X (solid curves -- and in the plane of the incident light) and Y contributions (dashed curves.) *There is little effect of distortion when the RMS slopes (root-mean-square) are less than about 10%.* Small changes of incidence angle clearly have a far greater effect on thrust than do sail distortions, so it should be possible to easily compensate for even much larger distortions by "trimming" the attitude of the spacecraft (at least as long as the distortions are relatively static.) We shall purposely delay consideration of the moments until later.

The mean square slope is related to the two-dimensional "wrinkle aspect ratio" of Ref. 1 by

$$<w'^2> = \frac{1}{2\ell}\int_0^{2\ell} (\frac{a\pi}{\ell}\cos\frac{\pi x}{\ell})^2\, dx = \frac{\pi^2}{2}(\frac{a}{\ell})^2 = (2.22\frac{a}{\ell})^2$$

and the results are equivalent when one accounts for the normalizations used in the paper. For some more interesting shapes (though not exactly representative of solar sails), the mean square slope for a parabolic strip of depth δ and length L is

$$\left\langle w'^2_x \right\rangle = \frac{1}{L}\int_{-L/2}^{L/2}\left[\frac{2\delta x}{(L/2)^2}\right]^2 dx = \frac{16}{3}\left(\frac{\delta}{L}\right)^2,$$

it is $$\left\langle w'^2_x \right\rangle = \frac{4}{\pi D^2}\int_0^{D/2}\int_0^{2\pi}\left[\frac{8\delta r}{D^2}\cos\theta\right]^2 r\, d\theta\, dr = 4\left(\frac{\delta}{D}\right)^2,$$

for a paraboloid of depth δ diameter D, and it is $<w'^2_x> = 4\,(\delta/L)^2$ for locally-plane wrinkled segments with surface slope 2δ/L. The reasonably close agreement of the numerical coefficients for the three simple shapes, as well as with the AEC wrinkle aspect ratio, gives confidence in our ability to relate the RMS slope distortion to the characteristic out-of-plane angle of the sail distortion.

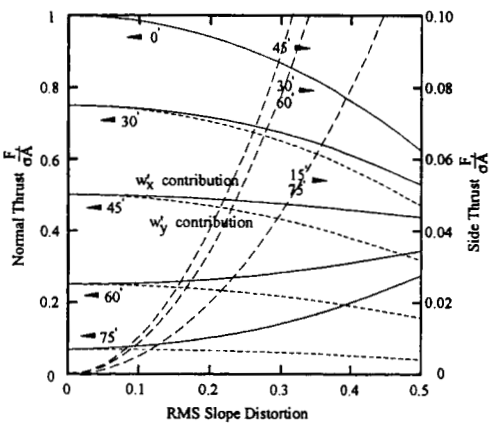

Figure 4: RMS slope effect on thrust

The analysis so far has been restricted to ideal sails with perfect specular reflection, and we now turn attention to nonideal ones. Returning to McInnis (Ref. 2, Sec. 2.6.1), the elemental optical force is the sum of vector components due to:

Absorption $\vec{f}_a = p\,(\vec{U}\cdot\vec{n})\vec{U}$, Specular reflection: $\vec{f}_{rs} = \rho\nu\; p\,(\vec{U}\cdot\vec{n})\vec{S}$,

Diffuse reflection: $\vec{f}_{rd} = B_f\rho(1-\nu)\; p\,(\vec{U}\cdot\vec{n})\vec{n}$,

and Thermal emission: $\vec{f}_e = \frac{\sigma'(T^4 - T_\infty^4)}{c}(\varepsilon_f B_f - \varepsilon_b B_b)\vec{n} = \frac{\varepsilon_f B_f - \varepsilon_b B_b}{\varepsilon_f + \varepsilon_b}(1-\rho)\; p\,(\vec{U}\cdot\vec{n})\vec{n}$

where $\vec{S}$ = unit reflection vector for which $(\vec{U}-\vec{S})\times\vec{n} = 0$ and $(\vec{U}+\vec{S})\cdot\vec{n} = 0$,

ρ = total reflection efficiency, ν = specular reflection efficiency,

B_f, B_b = front and back surface non-Lambertian reflection coefficients,

ε_f , ε_b = front and back surface emissivities.

Neglecting the small effect of sail distortion on $(\vec{U}\cdot\vec{n})$, the thermal emission equations predict the equilibrium sail temperature to be given by $T^4 = \frac{pc}{\sigma'}\frac{(1-\rho)}{\varepsilon_f + \varepsilon_b}\cos\theta + T_\infty^4$.

McInnis's diffuse reflection model directs the thrust component normal to the surface, but this doesn't appear to be an important restriction because some diffuse contribution can be assigned to the specular one.

Again referencing force components to a mean flat surface as in Fig. 2, the reflected unit vector is given by

$$\vec{S} = \vec{U} - 2(\vec{U}\cdot\vec{n})\vec{n} = \vec{k}\cos\theta - \vec{i}\sin\theta - 2\frac{\cos\theta + w'_X\sin\theta}{1 + w'^2_X + w'^2_Y}(\vec{k} - \vec{i}w'_X - \vec{j}w'_Y)\,,$$

so the total force per unit area on the sail is then

$$\vec{f} = \vec{f}_a + \vec{f}_{rs} + \vec{f}_{rd} + \vec{f}_e = p(\vec{U}\cdot\vec{n})[\vec{U} - \alpha\vec{S} + \beta\vec{n}] = p(\vec{U}\cdot\vec{n})[(1-\alpha)\vec{U} + (2\alpha\vec{U}\cdot\vec{n} + \beta)\vec{n}]$$

$$= p\frac{\cos\theta + w'_X\sin\theta}{\sqrt{1 + w'^2_X + w'^2_Y}}[(1-\alpha)(\vec{k}\cos\theta - \vec{i}\sin\theta) + (2\alpha\frac{\cos\theta + w'_X\sin\theta}{\sqrt{1 + w'^2_X + w'^2_Y}} + \beta)\frac{\vec{k} - \vec{i}w'_X - \vec{j}w'_Y}{\sqrt{1 + w'^2_X + w'^2_Y}}]$$

where we have introduced two new material property parameters, $\alpha = \rho\nu$ being the fraction of the incident flux that is specular reflected in the direction $\vec{S}$, and

$$\beta = B_f\rho(1-\nu) + (1-\rho)(\varepsilon_f B_f - \varepsilon_b B_b)/(\varepsilon_f + \varepsilon_b)$$

being the fraction of the incident flux that contributes to thrust from diffuse reflection and thermal re-emission. These two parameters totally account for the effects of absorption, specular and diffuse reflection, and thermal emission on thrust and moments. McInnis gives

$\rho = 0.88$ $\nu = 0.94$ $B_f = 0.79$ $B_b = 0.55$ $\varepsilon_f = 0.05$ $\varepsilon_b = 0.55$

for the values of the parameters based upon JPL calculations for a comet Halley rendezvous, so $\alpha = 0.827$ and $\beta = 0.0417 - 0.0526 = -0.0109$ for that case, and the resulting 1AU equilibrium temperature for normal incidence and zero sink temperature is 264 ^{0}K. The net contribution of diffuse reflection and re-emission of thermal radiation is very small for this representative case, and it appears likely that it will be the case in general as long as ρ and ν are close to unity.

The force on a perfectly flat sail (w' = 0) is given by

$$\vec{f} = p\cos\theta[\vec{k}\{(1+\alpha)\cos\theta + \beta\} - \vec{i}(1-\alpha)\sin\theta]\,.$$

The thrust on a sail varies primarily as the square of the cosine of the incidence angle, but an in-plane component $f_X = -\,p(1-\alpha)\sin\theta\cos\theta$ arises when there is absorbed flux, and McInnis has pointed out that it directs the resultant thrust at an angle $\psi = \theta + \tan^{-1}(\frac{f_X}{f_Z}) = \theta - \tan^{-1}(\frac{(1-\alpha)\sin\theta}{(1+\alpha)\cos\theta + \beta})$

to the incident vector. The result is that pitch is less effective at directing thrust with a non-ideal sail, and there is, in fact, a maximum steering angle that can be obtained. The maximum thrust angle is $\psi = 55.5^0$ at a pitch angle $\theta = 72.6^0$ for the given parameters, but the thrust angle relative to the sail normal itself at that condition is -17.1^0, and it's magnitude continues to rise with pitch angle.

The force equation for the nonideal sail can be expanded as was done in a Taylor series for small values of the slopes w', this time producing the far more complicated expression

$$\frac{\vec{f}}{p} = \vec{k}\,[(1+\alpha)\cos^2\theta + \beta\cos\theta + \{\frac{1+3\alpha}{2}\sin 2\theta + \beta\sin\theta)w'_x - (\frac{1+5\alpha}{2}\cos^2\theta - 2\alpha\sin^2\theta + \beta\cos\theta\}w'^2_x$$

$$-(\frac{1+5\alpha}{2}\cos^2\theta + \beta\cos\theta)w'^2_y + \ldots] - \vec{i}\;[\frac{1-\alpha}{2}\sin 2\theta + \{\frac{1-\alpha}{2}\sin 2\theta + 2\alpha\cos^2\theta + \beta\cos\theta\}w'_x$$

$$+\{\frac{9\alpha-1}{4}\sin 2\theta + \beta\sin\theta\}w'^2_x - \frac{1-\alpha}{4}w'^2_y\sin 2\theta + \ldots] - \vec{j}\;[(2\alpha\cos^2\theta + \beta\cos)w'_y + (2\alpha\sin 2\theta + \beta\sin\theta)w'_x\,w'_y + \ldots].$$

The results remain complicated even after the area integration, but the essential behavior can be explored rather easily by working in steps, noting that the mean (i.e., integrated) values of the slopes w'_x and w'_y vanish by definition. Then, first letting the nonideal surface parameters α and β be constant, the total thrust is just

$$\vec{F}/pA = \iint \vec{f}/pA\,dA = \vec{k}\,[(1+\alpha)\cos^2\theta + \beta\cos\theta] - \vec{i}\,\frac{1-\alpha}{2}\sin 2\theta$$

$$-[\vec{k}(\frac{1+5\alpha}{2}\cos^2\theta - 2\alpha\sin^2\theta + \beta\cos\theta) + \vec{i}(\frac{9\alpha-1}{4}\sin 2\theta + \beta\sin\theta)] < w'^2_x >$$

$$-[\vec{k}(\frac{1+5\alpha}{2}\cos^2\theta + \beta\cos\theta) - \vec{i}\,\frac{1-\alpha}{2}\sin 2\theta] < w'^2_y > - \vec{j}\;(2\alpha\sin 2\theta + \beta\sin\theta) < w'_x\,w'_y > + \ldots$$

which reduces to the previous result for an ideal sail when $\alpha = 1$ and $\beta = 0$ and which reduces to the flat sail result when the mean square slopes $<w'^2_x>$ and $<w'^2_y>$ are small (e.g., less than about 0.01.) However, if they are small and if the surface parameters α and β are now allowed to vary, then linear slope terms of the form $<\alpha w'>$ and $<\beta w'>$ will also be small if the property variations $\delta\alpha$ and $\delta\beta$ are less than about 0.1. *It then follows that the thrust of even a nonuniform sail will act as though flat if* $\sqrt{<w'^2_x>}$, $\sqrt{<w'^2_y>}$, $\delta\alpha$, *and* $\delta\beta$ *are all less than about 0.1.*

The results for the moments $\vec{M} = \iint(\vec{x} - \bar{\vec{x}}) \times \vec{f}\,dA$ involve products of the moment arms, so the dimensions must be constant (or nearly so) across the scale of the disturbances if the same rules are to apply. This will be essentially the case for small-scale distortions like wrinkles and creases, and it will also be true for small nonuniform patches, but it may not hold true for billow or for large lossy regions of sail, either of which will require that attention be paid to the details of the disturbance.

While it might not guarantee flat sail performance for the moments when there are large-scale disturbances, the "10% rule" for RMS surface slopes and surface property variations will produce the most well-defined and repeatable behavior, and it is therefore appropriate to regard it as a fundamental design rule for solar sails.

RMS slope is a useful mathematical concept, but one might ask whether it can actually be measured, and how, in fact, can one best measure sail distortion. Photogrammetry as done by LaRC's Optical Diagnostic System (ODS) (Ref. 3) can determine large scale distortions such as billow, but small-lateral-scale distortions such as wrinkles and creases can be missed completely -- although the degradation due to them could easily be just as important. However, the mean square slope itself is linked directly to the foreshortening of the sail material since, for example, the actual distance in the x-direction under a distorted sail with local surface slope w'_x is given by

$$x' = \int \cos w'_x dx = \int [1 - w'^2_x/2 + \ldots]dx = (1 - <w'^2_x>/2 + \ldots)x$$

where x is measured along the surface of the membrane. Thus, the RMS slope is related to the "apparent strain" by

$$<w'^2_x> = 2\,\Delta x / x \quad \text{and} \quad <w'^2_x + w'^2_y> = 2\,\Delta A / A \quad .$$

A 1% mean square slope corresponds to a 1/2% change in linear dimension, which, though small, should be easily measurable by a multi-megapixel digital camera with a good wide-angle lens (5 pixels out of 1000.) This might allow photogrammetry to make far more useful *in-plane* measurements of surface distortions with much larger spatial frequencies than those of the targets themselves. Resolution in the direction of viewing would be degraded with a shallow viewing angle, but complementary orthogonal cameras could solve that problem if practical locations could be found for them.

II. Sail Tension

The total force component in the attachments normal to the nominal plane of the sails is just equal to the normal force on a sail, whether produced by gravity on the ground or by the very small solar flux in space, but the resultant forces in the attachments, as well as the stresses in the sails, depend upon the specific geometry. Regardless of how much the tension in the sail is increased, the normal force components remain unchanged, the angle between the sail attachment and the plane of the sail decreases – and the sail is flattened.

Both the AEC/SRS and L'Garde systems use triangular sail quadrants, but the latter is more easily analyzed because it is composed of parallel stripes, each supported by cords through their centers as shown in Fig. 5. Defining the length of a stripe as L', the width as h, and the load per unit area as σ, the load on a stripe is $\sigma L'h$, and the opposing normal force at each end is then $\sigma L'h/2$ if adjoining stripes are considered to be essentially independent. If the sail shape is approximated by a parabola of depth δ (the limiting shape for a shallow catenary, sphere, etc.), then the tension acts at a shallow angle $\theta = 4\delta/L'$ so that the tension F in the cord is given by

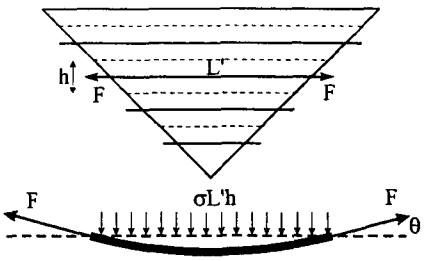

Fig. 5: L'Garde striped sail

$$\frac{F}{\sigma A} = \frac{F}{\sigma L'h} = \frac{1}{2\theta} = 0.125\frac{L'}{\delta}$$

where A = L'h is the area of a full stripe. The AEC/SRS quadrants, on the other hand, are each supported by halyards at the three corners, so the shape is more complicated. However, their Slide #84 from Ref. 4 gives the data below for the variation of sail peak displacement with halyard force, and Mathcad was used to produce the linear least-squares fit $\delta = 0.593 + 15.522/F$ for the L_S = 10m sail with a 1g normal loading of 6 gm/m^2. The small 0.6" distortion offset is likely due to measurement technique, and we'll choose here to simply subtract it from the deflection data. Since 10m = 393.7" and $\sigma L_S^2 = 1.32$ lbs for the given conditions, the result can be expressed as

F, lb.	1.5	3	4.5
δ, in.	10.93	5.81	4.01

$$\frac{F}{\sigma A} = \frac{4F}{\sigma L^2} = \frac{(4)(15.22)}{(393.7)(1.32)}\frac{L}{\delta} = 0.1195\frac{L}{\delta}$$

where $A = L^2/4$ is the area of a quadrant. The two results are plotted together in Fig. 6 using the quadrant length L as the characteristic length for the Able sail and the stripe length L' for the other. The remarkable agreement is somewhat fortuitous and arises in part from the particular choice of length for the AEC sail, but there is no question that the functional dependence on the characteristic slope is the same. We know from the preceding section that the RMS slope distortion for the L'Garde sail is 2.31 δ/L', so $F/\sigma A$ must exceed about 2.31/0.8 = 2.9 for the RMS distortion to be restrained to the 10% level acceptable for good and predictable sail performance. It's less certain how to relate the characteristic slope of the AEC/SRS sail to its RMS, but it is suggested in the absence of better information that $F/\sigma A = 3$ be applied as a design rule to all sails. It is worthy of note that all of the AEC/SRS data lies above this value.

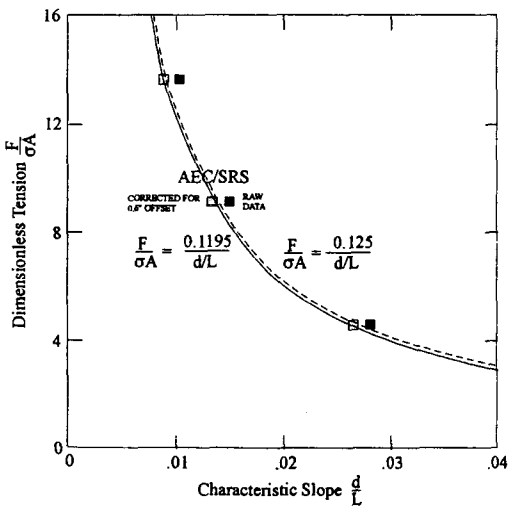

Fig. 6: Effect of tension on sail distortion

The halyard force is shown in Fig. 7 for operation at 1AU, and it is generally small because of the weak solar flux, but it can become important for very large sails, especially those operating near the sun and/or for ones with high tension to minimize sail deflection. We are ultimately interested not so much in the tensile load itself but rather in the buckling load that it imposes on the boom and on the stresses that are imparted to the sail. The force F in the AEC/SRS halyards tends to act between 14 and 22.5 degrees to the boom, whereas the L'Garde stripe wires (except that at the edge) are at 45^0 to the boom. The axial force for the former is then 92 – 97% of F, whereas it is 0.707F for the latter. Both corrections are

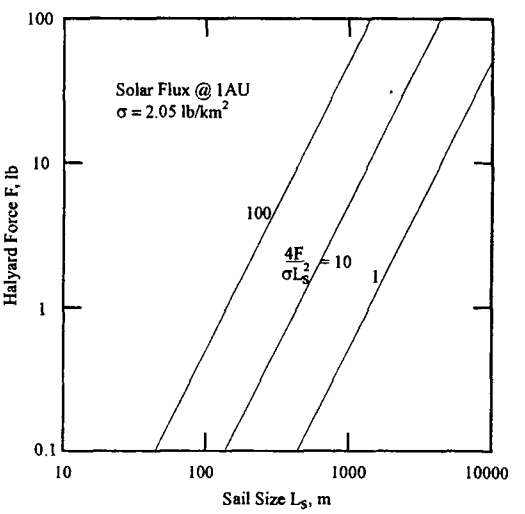

Fig. 7: Halyard force

relatively minor, and a more important difference is that the AEC/SRS system applies the full load at the end, whereas L'Garde distributes it along the length of the boom.

SRS has measured the yield stress of coated CP-1 sail material and found it to be about 9500psi, so a 5μ sail can support about 25 lb/m with a factor of safety of 3. Fig. 7 shows that the halyard force is typically no more than a few pounds, so sail stresses become an issue only in concentration zones like those near the attachment points where the stresses tend to vary radially with distance from them. SRS uses grommets and doublers to reinforce the corners of the sail, and their shear compliant border further helps to spread the load. The stress should not produce a limitation even with extremely large (or thinner) sails as long as careful attention is paid to the corners (or wherever force might be concentrated) and to bonding techniques.

L'Garde's present sail is fabricated in stripes from 2.5μ Mylar, a material that has been used in both space and industry for many years and whose general properties are well characterized as exemplified by the Dupont stress-strain curves of Fig. 8. There is appreciable temperature dependence, but the strength is not significantly different from that of CP-1. The stripes are in turn bonded along their lengths to Kevlar cords that are suspended periodically from the booms, so stress concentrations are minimized, and the mean stress in a stripe of width h and thickness t is

$$S = (F/\sigma A)\,\sigma\, h/t\,,$$

which is generally far less than 0.1 psi for operation with solar flux – and is also far less than the values encountered with most tests and simulations. The same is true away from the corners and edges of the AEC/SRS sail.

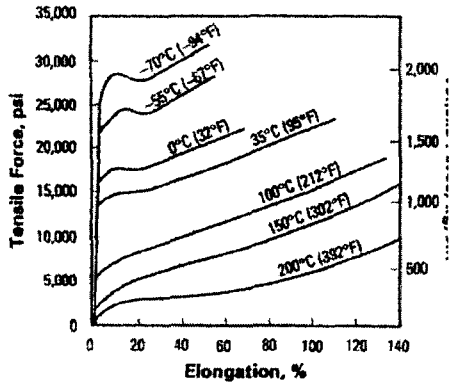

Figure 8: Mylar strength

III. Dynamics

AEC, L'Garde, and LaRC have performed varying degrees of dynamics modeling, but there has been little application to actual system design. Some of the most useful data was presented by Able in Ref. 4. Typical FEM results for mode frequency and shape are shown in Fig. 9 for the 10 m "workhorse" sail, and the frequency results are summarized on the next page:

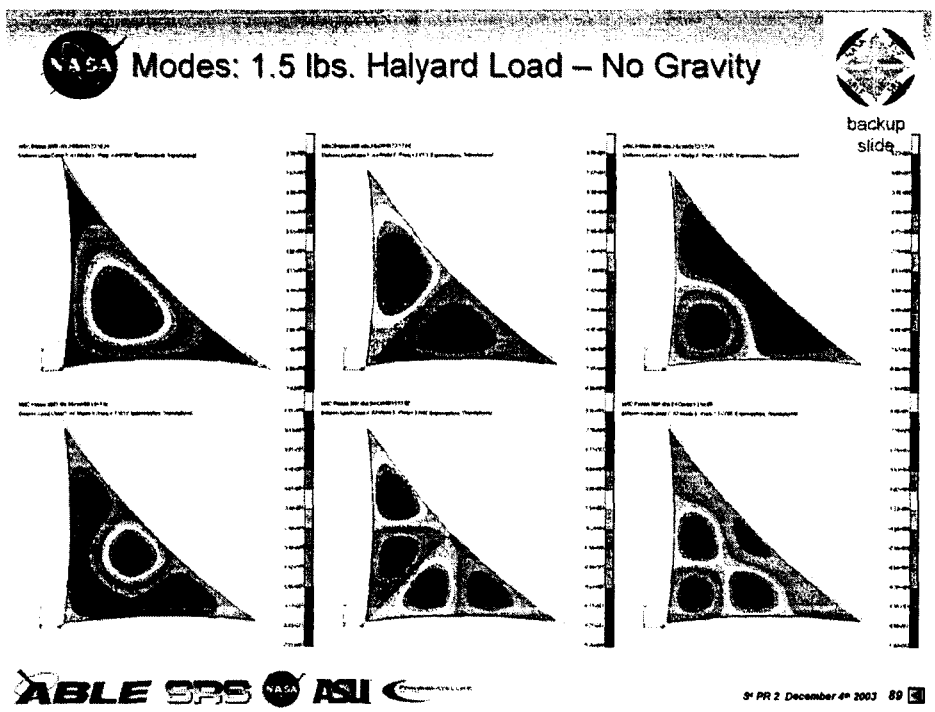

Fig. 9: Predicted Able sail dynamics

Gravity, g	Tension T, Lbs.	F_0, Hz	F_1, Hz	F_2, Hz	F_3, Hz	F_4, Hz	F_5, Hz
0	1.5	0.971	1.471	1.625	1.922	2.159	2.280
0	3	1.378	2.091	2.318	2.737	3.078	3.253
0	4.5	1.696	2.573	2.862	3.375	3.799	4.018
1	1.5	1.188	1.939	2.431	2.738	2.816	3.255
1	3	1.551	2.473	2.871	3.377	3.559	3.926
1	4.5	1.855	2.936	3.308	3.939	4.170	4.552

Classical membrane theory predicts that resonant frequencies are proportional to $1/\lambda\ \sqrt{(\sigma'/\rho)}$ where λ is the characteristic mode size, σ' is the tension per unit length, and ρ is the areal density of the membrane material. The above results normalized by multiplying by $\sqrt{(1.5/T)}$ and dividing by the lowest frequency at T = 1.5 lbs are:

Gravity, g	Tension T, Lbs.	F_0, Hz	F_1, Hz	F_2, Hz	F_3, Hz	F_4, Hz	F_5, Hz
0	1.5	1	1.515	1.673	1.979	2.223	2.348
0	3	1.004	1.522	1.688	1.993	2.242	2.369
0	4.5	1.008	1.530	1.701	2.007	2.247	2.389
1	1.5	1	1.632	2.046	2.304	2.370	2.740
1	3	0.923	1.471	1.709	2.010	2.118	2.336
1	4.5	0.902	1.427	1.607	1.914	2.026	2.212

The results with gravity obviously fail to follow the simple scaling and show that strong normal loading will somewhat increase the natural frequencies, but the 1g gravitational load of the 5.96 gm/m^2 material is 58400 N/km^2 or 6400 times the 1AU solar flux, so it can safely be concluded that external loading will have only a negligible effect on the dynamics of a solar sail in space. On the other hand, the frequency results for no gravity certainly agree well with the simple scaling law, and it appears reasonable to suggest that the resonant frequencies of other sail quadrants in space with three-point halyard support will be given by

$$f_{Hz} = \alpha \frac{9.71}{L_m} \sqrt{\frac{T_{LBS}/L_m}{0.15} \frac{5.96}{\rho_{g/m\wedge 2}}}\ ,\ \alpha = 1.00, 1.52, 1.68, 1.98, 2.23, 2.36, \ldots\ ,$$

and that the corresponding modes will be those shown in the AEC slide. It should be noted that the "classical membrane theory" rigorously applies to membranes uniformly supported along their periphery, not discretely supported as are solar sails, but the close agreement with sail data shows the value of extending such models after verifying that they may indeed be appropriate,.

It's also worth noting that the simple static model predicts that the sail tension should be proportional to the normal load for a given level of slope distortion, so that would produce a $L^{-1/2}$ dependence for the sail natural frequencies – as long as both the slope and the loading remain the same. A word of caution here: the tension used for the AEC data corresponds to that required for a reasonable billow at 1g (regardless of the "no gravity" description), and is over 6000 times larger than really needed in space, so it appears that the actual natural frequencies in space could be almost two orders of magnitude lower.

Comparable information has not yet been found for the L'Garde striped sail, but they have reported 1.073 HZ for the lowest order mode of their 10M sail. No cord tension was specified, but it appears likely that their multi-point support shouldn't significantly affect either the mode frequency or shape for at least the lower orders (for given loading and billow – again probably corresponding to ground operation.)

A key point to be remembered for either type of sail is that internal stresses (even with 1g loading) are very small in the sails and in their attachment cords, so elastic strain will be virtually negligible compared to the apparent strain due to sail distortion. Thus, neither static nor dynamic loads will cause the sails to stretch, and their peak displacement will be determined by the details of their distortion distribution and their attachment to the booms. Global thermal expansion due to overall heating or cooling of the sail can also be important: heating can cause significant wrinkling and/or billow, while cooling can produce excessively large thermal stress if there is not enough billow to accommodate it. Tension is essentially inversely proportional to the billow (or RMS slope distortion), and thermal contraction can therefore significantly limit the allowable tension.

IV. System Scaling

Although the functions of the booms are the same, i.e. to deploy themselves and the sails in a fully controlled manner from a collapsed state and to thereafter support the sails and associated controls during spacecraft operation, the designs from the two contractors are otherwise fundamentally quite different.

The AEC CoilAble boom is a truss structure that deploys like an unwinding coiled spring under its own stored energy from a collapsed state with a length typically about 1% or less of the deployed length. It is composed of three continuous longerons (graphite in the ISP design), periodically spaced equilateral triangles of battens that bend and spring-load the longerons in the deployed state, and diagonal tension elements (Kevlar wires) connected at the batten/longeron junctions. The design bears some semblance to a tensegrity structure, but it is not a tensegrity because of the internal moments experienced by the longerons under applied load. The boom is deployed first and the sail is raised thereafter using halyards from the tips.

The principal failure mode of the structure is global buckling, for which the classical critical load for a column fixed at one end and free to rotate at the other is $P_{CR} = \pi^2 EI/4L_B^2$ where the moment of inertia $I = 3R^2h^2/2$, R is the "radius" of the structure, h is the diameter (width) of a longeron, and $L_B = L/\sqrt{2}$ is the total boom length. Local buckling of the longerons between the battens can also be an issue, especially if all of the compressive load should be taken on one of the longerons; it corresponds to a column free to rotate at both ends, and is described by $P_{CR} = \pi^2 EI'/l^2$ where $I' = h^4/12$ is the moment of inertia of a single square longeron and l is the unsupported length between battens. Local buckling can therefore be avoided by sizing the system such that $l < \sqrt{2}\, hL_B/3R$. Column theory is far less than an exact science, so a safety factor η of least 3 should be applied. Longeron compressive stress must also be no more than about 1/3 of the maximum allowable working stress σ_w. The collapsed package will have a radius of about R, while the absolute minimum stored length $L_0 = hL_B/2\pi R$ allowing no room for fittings and the like, and AEC (Ref. 5) recommends $L_0/L_B = (3/2\pi R)(h+0.02")$ for a lanyard-deployed CoilAble. A key factor in the design is the maximum strain $\varepsilon = h/2R$ in the longerons in their coiled state, and AEC reports that $\varepsilon = 0.015$ has served them well for their many previous booms that have used S-glass/epoxy, but that a lower value of 0.0072 is recommended for the newer high strength pultruded graphite booms with E = 27,300 kpsi. It should be cautioned, however, that the peak working stress $\sigma = \varepsilon E = 200$ kpsi in the stored longerons seems to be very high for most materials, especially for a pultruded composite for which high axial compressive loads might lead to delamination. Finally, it can be shown that the stored energy $\mathcal{E}_s$ in the coiled boom is proportional to $E\varepsilon^2 h^2 L_B$ so that the self-deployment force is proportional to $E\varepsilon^2 h^2$.

Noting that the boom mass tends to be dominated by the three longerons, its mass can be represented by $M_B = 3\rho_B h^2 L_B$ where ρ_B is the effective mass density, so general scaling laws for the Able boom are

$$\frac{h}{L} = \left(\frac{8\sigma\eta\varepsilon^2}{\pi^2 E}\right)^{1/4}$$

and

$$\frac{M_B}{L^3} = \frac{3}{\sqrt{2}}\frac{\rho_B h^2}{L^2} = \frac{6\varepsilon\rho_B}{\pi}\sqrt{\frac{\sigma\eta}{E}}$$

where we have set $P_{CR} = 3\sigma L^2/2$ in accordance with the observations of Section II for two quadrants per boom. Using $\varepsilon = 0.00719$ and $\rho_B = 2.89$ gm/cc obtained from the 70.1 gm/m Able mast with 2.84 mm graphite longerons and a diameter of 395 mm, these are $h/L = 0.00884$ mm/m and $M_B/L^3 = 0.000478$ gm/m^3 for operation at normal incidence at 1 AU. The results are plotted in Fig. 10 in terms of the areal density. The masses for the booms and the sails add, and an optimum system corresponds to roughly equal contributions from the two because effort should generally be directed towards reducing the larger and/or easing requirements on the smaller. It should be immediately obvious from the plot that the linear scaling of the AEC boom, while advantageous from the standpoint of design experience, would yield a beam that would be far lighter than necessary and weaker than desirable for large sails – unless the mass of the sails themselves could be

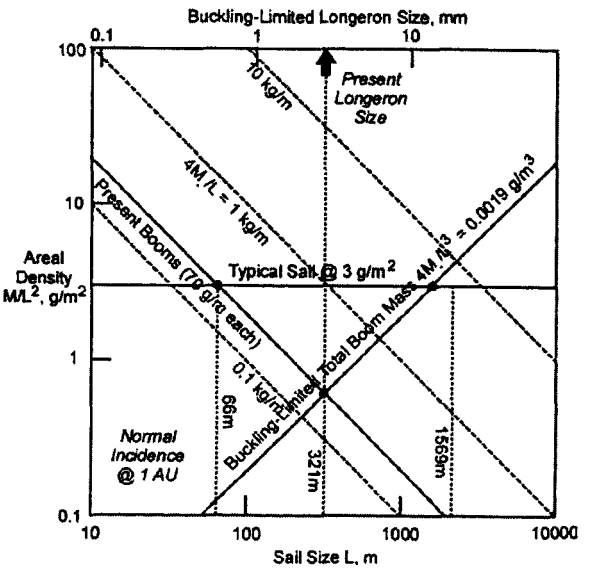

Figure 10: Able Scaling

significantly reduced. The mass of a 70 g/m boom would be less than that of a relatively advanced 3 g/m^2 sail above a sail size of about 66 m, the same design could be used in space at 1 AU up to a sail size of about 321 m, and a geometrically scaled version could be used up to 1569 meters without the mass exceeding that of the sail! The present Able boom is significantly over-designed for many solar sail missions -- though not for operation or testing at 1g.

Simple theory predicts that the resonant frequencies for lateral vibrational bending of a uniform beam are given by $f_n = \beta_n/L_B^2 \sqrt{(EI/\rho L_B)}$ where $\beta_0 = 0.56$ for a clamped-free configuration (Ref. 6.) This predicts a lowest natural frequency of 12 Hz for the 7.1 m CoilAble boom with a rigidity of 82,407 N-m^2 and mass of 70.1 g/m, whereas AEC has measured 5.6 Hz with a 0.4 lb tip mass and with tip loading via a negator spring. The agreement seems reasonable considering factors like the tip loading, and lends credence to the use of the simple scaling form with a "semi-empirical correction factor" 5.6/12 = 0.47.

To complete the scaling for 1 AU solar loading, the lowest resonant frequencies of a 3 g/m^2 sail and a buckling-limited Able boom are given by

$$f_S = \frac{0.044}{\sqrt{L_m}} Hz \quad \text{and} \quad f_B = \frac{1.52}{L_m} Hz$$

which are plotted in Fig. 11. One can argue with the details of the assumed numerical values, but the point here is that buckling-limited scaling for space operation causes the boom frequency to be higher than that of the sail and its mass to be less for sail sizes up to about a kilometer. Damping may be very difficult to estimate for these systems, but complicated coupling between the boom and sail loads could possibly be avoided by widely separating the natural frequencies, and it does appear that the lowest sail and boom modes will in fact be reasonably well separated for all but the largest systems of interest for practical applications. Although it is unlikely that the modes can be separated further by reducing the natural frequencies of the sails themselves, the rigidity of the booms can be increased somewhat to increase their frequencies without significantly increasing the overall system mass.

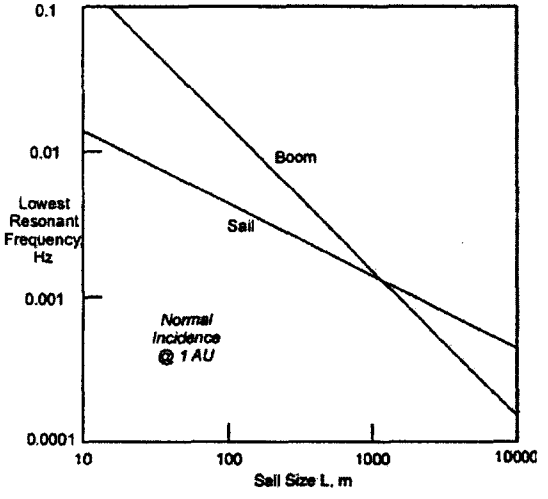

Fig. 11: Natural frequencies for Able system

The L'Garde boom is a telescopic structure of cylindrical tubes that are deployed via internal gas pressure and are then thermally rigidized. Sailboat-like battens and stays provide additional strength with minimal added mass, and the "striped" sails are attached periodically to the booms. L'Garde and LaRC have performed detailed computer modeling of the system, and Greschik (Ref. 7) has provided a set of structural equations with some numerical results for a Solar Polar Imager @ 0.48 AU, but useful values for scaling are meager at best, so an independent simple model is presented here using available design information.

Except for the outer edge with its "half-width" stripe and "catenary scallop", the L'Garde sail quadrants load the stepped booms with a discrete set of linearly-distributed 45^0 force components. Greschik gives base diameter D_L = 10 cm, tip diameter D_0= 7.5 cm, wall thickness t = 0.0513 mm, and modulus E = 4,160,000 psi for the L_B = 77.3 m L'Garde boom, so the equivalent tube rigidity

$$EI = \pi E/4 \int_0^1 [D_0 + (D_L - D_0)\xi]^3 \, d\xi = \pi E/16 (D_0 + D_L)(D_0^2 + D_L^0)$$

is 808 n-m^2. This is far smaller than the value of 11315 calculated by Greschik including the effects of the additional supporting structure of stringers, etc., but he does degrade it by 40% to 6789 N-m^2 to account for less-than-rigid connections. Thus, the tapered boom apparently contributes only very little to the bending/buckling stiffness of the system and apparently serves primarily as a compression/torsion member, so it will be assumed here that the rigidity scales geometrically with D^3t and that the mass scales with DtL_B. However, because of the stacked nature of the telescoping tubes, this boom is unlikely to scale simply by increasing the length. It is more likely that stripe widths might remain the same because of experience gained with fabrication of smaller systems, and we shall assume for added strength that the wall thickness will scale with the diameter. Greschik gives 50.14 gm/m for the specific mass of the L_B = 77.3 m boom, and 3.03 g/m^2 as the areal density for the Mylar film and Kevlar cord sail combination.

Assuming that buckling is in fact the proper failure criterion for the L'Garde boom, the deformation of a uniform boom with linearly distributed axial loading $q(x) = q_L\, x/L_B$ can be modeled, and the results are well represented by $\bar{q} = 16EI/\eta L_B^3 = 6\sigma L$ where

$$EI\frac{d^2y}{dx^2} = -M(x) = -\int_0^x q(\xi)[y(x) - \eta(\xi)]d\xi \quad \text{or} \quad EI\frac{d^3y}{dx^3} = -\frac{dy}{dx}\int_0^x q(\xi)\,d\xi = \left[q_0 + (q_L - q_0)\frac{x}{2L_B}\right]x\frac{dy}{dx}$$

subject to y = 0 and y" = 0 at x = 0 and y' = 0 at $x = L_B$.

It follows then that the scaling is given by

$$N_{FS}\frac{F}{\sigma A} = 151.8\left(\frac{77.3}{L}\right)^4\left(\frac{D_0}{7.5}\right)^4\left(1+\frac{L/3}{77.3}\right)^2\left(1+\frac{L/6}{77.3}\right)$$

and

$$\frac{M_B}{L_B} = 42.98\left(\frac{D_0}{7.5}\right)^2\left(1+\frac{L/6}{77.3}\right) gm/m$$

Fig. 12 shows that, with the prescribed scaling, a 357m sail can be used at 1AU with a L'Garde boom having D_0 = 7.5 cm as it is now with N_{FS} F/σA=9 (3 for distortion and 3 for safety factor) and that its mass would remain less than that of the sail even if the latter were reduced up to 5.7X, e.g. from 2.5μ thickness to 0.44μ. Doubling the tip diameter to 15 cm would allow a 1293m sail, and would still allow a 2.7X margin for sail mass reduction.

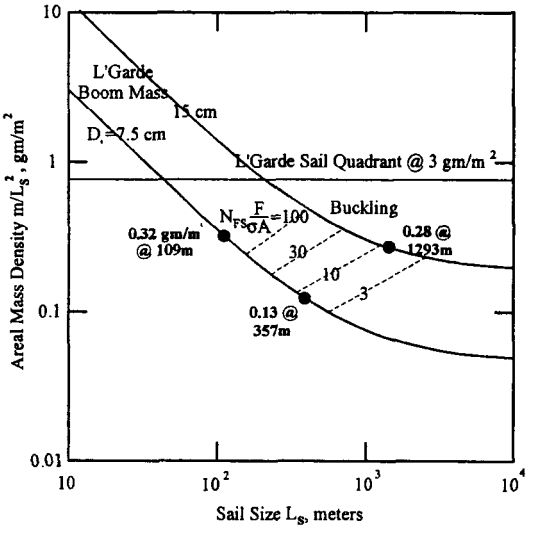

Figure 12: L'Garde System Scaling

The ultimate issue with larger systems will probably not be mass, however, but rather sail fabrication, packaging techniques, and dimensions. Sizes larger than about 100 meters or so may extend beyond the envelopes of all but the largest available launch vehicles with the present solar sail "cruciform" configurations. A storage length of about 1% is typical for the AEC boom (actually 0.81% for e = 0.0072 and h = 2.84 mm), but the L'Garde scaling effectively maintains constant stored length while the diameter increases. If square sail designs are to continue to be used, it may prove necessary to store both the booms and the sails in a parallel configuration, folding both 90^0 into a common plane for deployment and operation, perhaps using a concept akin to a sailboat's "roller furling" for the sails. .

V. Conclusions

It has been shown here that the design and the scaling of solar sail systems proceeds in an orderly manner from requirements on allowable sail distortion, to the tension required to produce that level, to the strength and mass of the booms needed to support the sails.

The fundamental parameter (or "figure of merit") for describing sail distortion is the mean sum of the squares of the surface slopes. This quantity is strikingly similar to the RMS phase distortion $<\delta/\lambda>$ that is used to describe wavefronts and surfaces in optical imaging systems, and the simple reason is that both are based on integrals over the surface of quantities that must be small for the systems to behave consistently and properly. Much as one would always use precision fabrication techniques, adaptive optics, or the like to reduce optical system wavefront errors to an allowable amount for good overall performance, so also can surface billow and wrinkles of a solar sail be reduced to an allowable level by applying tension to the sail to effectively flatten it. Creases are likely to persist, but the affected fractional area will invariably be small. The analysis in fact shows that an RMS slope of about 10% will be sufficient to assure good performance, and that it can be obtained with a halyard (AEC) or cord (L'Garde) load about equal to three times the normal force on the supported area. Therefore, solar sail distortion-produced thrust and moment uncertainties can and should be *prevented*, not modeled and measured except to show that they can in fact be discounted.

Present sail designs can easily support such loads essentially regardless of size, but the tensile loads are directly transferred to the booms which must be designed to prevent buckling, their principal potential mode of failure. It is

shown that the Able and L'Garde booms are both capable of remarkable strength and low mass for lengths of hundreds of meters in their present configurations, and that increases of transverse dimensions can extend their use to kilometer-size sails for use in the 1 AU or so space environment. These would be true space structures, however, and would not support the weight of a sail at even hundredths of one g.

The conclusions of this report are an important beginning to answer the questions regarding scaling to very large solar sail systems, and they provide credibility to the approach being taken by the NASA In-Space Program. However, this scaling study has not addressed the major operational and logistical issues of gossamer-like sail system fabrication, handling, storage, and deployment – perhaps done in space itself -- that will require carefully planned and conducted demonstrations, not just modeling, with small-scale systems to prove that very large ones can indeed be trusted to perform properly in the final environment.

Acknowledgment

The work described in this paper was funded in whole or in part by the In-Space Propulsion Technology Program, which is managed by NASA's Science Mission Directorate in Washington, D.C., and implemented by the In-Space Propulsion Technology Office at Marshall Space Flight Center in Huntsville, Ala. The program objective is to develop in-space propulsion technologies that can enable or benefit near and mid-term NASA space science missions by significantly reducing cost, mass or travel times.

References

1. T.W. Murphey, D.M. Murphy, M.M. Mikulas Jr., and A.L. Adler, "A Method to Quantify the Thrust Degradation Effects of Structural Wrinkles in Solar Sails", AIAA Paper 2002-1560, 43rd AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference, Denver CO, April 2002

2. C.R. McInnis, *Solar Sailing: Technology, Dynamics, and Mission Applications*, Springer-Praxis Series in Space Science & Technology, 1999

3. R.S. Pappa, et al, "Optical Diagnostic System for Solar Sail: Phase I Final Report", NASA TM-2004-213511, December 2004

4. Able Engineering Progress Review #2, ISP S^4 Phase II Contract No. NAS8-03043, 4 Dec 2003

5. D. Murphy, Able Engineering, Personal communication

6. L.S. Marks, *Mechanical Engineer' Handbook*, 5th Edition, McGraw Hill, New York, 1951, p. 499

7. G. Greschik, "ROSS NRA Cycle 1 – L'Garde Solar Sail General Structural Design and Analysis", L'Garde Structural Report 2003/3, U. of Colorado, November 2003

8. S.P. Timoshenko and J.M. Gere, *Theory of Elastic Stability*, McGraw Hill, New York, 1961