Self-Folding Metal Origami

Nathan Lazarus,* Gabriel L. Smith, and Michael D. Dickey

1. Introduction

Origami tends to bring back fond childhood memories of making cranes and boats from paper. It is easy to dismiss the field as “art” or “toys,” rather than a serious discipline. The reality is that we take for granted the importance and practicality of origami. We interact with structures folded through origami principles on a daily basis, from consumer packaging to automobile airbags and from shipping containers to the great solar arrays of satellites and space stations.1] Folding can seem so basic and natural that we fail to notice the complex series of folds that allows, for instance, our re-usable shopping bags to go from a flat, portable sheet into a well-supported 3D structure2] able to carry groceries.

1.1. Self-folding Origami

As the field of origami has grown and developed, there has been increasing focus on creating origami that is not just capable of being folded but can also fold itself. Removing the need for manual folding from the equation opens up compelling new possibilities in terms of capability and scale. Self-folding, defined as hands-free folding through an external stimulus such as light or heat,3] has opened whole new application spaces from robotics4] to structures at the micro- and nano-scale.5] Being able to rapidly fold and unfold, to re-configure at will, has brought science fiction concepts such as programmable matter,6] – that is, materials able to rapidly and dramatically re-shape themselves – to life in the laboratory.

Although we use the term “origami” in a general sense within this review, it derives from the Japanese words “ori” (folded) and “gami” (paper), and naturally folding of paper is the first material that comes to mind.7] In conventional origami, folding is largely independent of material choice beyond the basic mechanical properties, and there are few limitations on materials as long as the sheets are sufficiently compliant (typically accomplished using thin materials). While folding concrete8] or sheet metal9] is far less intuitive, most types of engineering materials are folded today. As the community has shifted toward self-folding, however, material choice becomes a critical parameter, and most actuation mechanisms are very much driven by the specific material used. Self-folding through magnetic fields,10] for instance, is heavily reliant on incorporating magnetic materials, and shape memory alloys11] rely on the intrinsic material phase properties for actuation.

Here, we will provide an overview of metal-based self-folding approaches and techniques as a complement to several recent surveys on polymer-based self-folding.12,13] Metals are fundamental to modern society, and as a material class, they are well known to provide superior conductivity14] as well as mechanical strength, toughness, and ductility.15] In origami, metals are most important for providing electrical conductors far superior than alternatives such as polymer composites or crystalline semiconductors. Creating intelligent systems with integrated electrical control or other functionality relies particularly heavily on folding of metal conductors. Important applications such as origami antennas and other electrical devices are difficult to envision by any other means. Similarly, working with metals provides new means of actuation such as shape memory alloys and magnetic materials, as well as new challenges in processing and design. The review will begin by surveying actuation methods for metal folding (Section 2), followed by discussing
commonly used materials (Section 3), control and design (Section 4), and applications (Section 5), finally providing outlook and concluding remarks (Section 6).

1.2. Definitions

Traditional origami consists of sharp folds (“creases”) with regions of unfolded paper between them; in engineering, this translates into hinges and rigid plates respectively.\textsuperscript{[15]} The requirement of a localized fold line or “hinge,” rather than a more gradual bend,\textsuperscript{[13]} is an important distinguishing feature of a self-folding origami structure (Figure 1a).

Strictly speaking, “origami” refers exclusively to folding of a sheet of paper without modification (i.e., no cutting), while “kirigami” refers to the related art where cutting is permitted.\textsuperscript{[16]} For illustration, Figure 1b,c shows a classic origami pattern, the Miura-ori,\textsuperscript{[17]} where the structure unfolds completely from its fully folded form to a flat sheet, while Figure 1d,e shows kirigami;\textsuperscript{[18]} with numerous cuts. Due to the ability to create higher complexity with cut lines, many practical systems and structures are “kirigami.” However, the distinction is not consistently made in the scientific community, and the accessibility and familiarity of the term “origami” has resulted in its use more broadly in many cases to refer to any folded structure. We will follow this convention here and do not attempt to rigorously differentiate between the two cases.

In paper folding, it is easy to fold in both directions, with folds typically classified into “mountain” or “valley” folds. As the names imply, a “mountain” fold looks like a mountain in which the hinge is a peak and the “panels” point downward. In contrast, a “valley fold” has its hinge at the bottom and the “panels” point upward.\textsuperscript{[19]} Many self-folding actuators can only deflect in one direction. While this is sufficient for many applications, more complex structures require folding both upward and downward. The Miura-ori pattern is one example of this type of geometry. Figure 1b shows the Miura fold pattern, with “mountain” and “valley” folds indicated as dashed and solid lines, respectively; this pattern cannot be created with single directional folds alone. The most powerful and general self-folding technologies therefore allow both.

2. Actuation Mechanisms

Self-folding origami can be remotely actuated. It is therefore fitting to begin by defining the means available for self-folding a metal structure. Broadly, actuation methods can be sorted into two distinct categories, namely bidirectional and unidirectional, defined largely by whether the actuator is capable of undoing its own movement. Bidirectional actuators are necessary for applications such as robotics (where a robotic gripper must be capable of releasing a held object, for instance), while unidirectional mechanisms incorporate a permanent change in the structure and are more typical in applications such as manufacturing and fabrication. Because unidirectional mechanisms typically change the structure permanently, these mechanisms often have the advantage of requiring no power to hold the folded structure in its final state, while many bidirectional actuators require holding power to remain folded.

The first three mechanisms here (bilayer thermal expansion, the shape memory effect, and magnetic actuation) are capable of being bidirectional, while the following two (laser forming and surface tension forces from a solid/liquid phase change) are typically unidirectional in nature. The section will then conclude with a broader survey of other, less commonly used actuation methods for folding metals.
2.1. Actuation through Thermal Expansion and Residual Stresses

Most materials have a tendency to expand with higher temperature, which is quantified by the "coefficient of thermal expansion" or CTE. The difference in the CTE within stacks of dissimilar materials results in out-of-plane bending, a classic result widely used in thermostats since the early 1900s. Figure 2a shows a basic schematic of this actuation mechanism. In a bilayer stack assuming that the top layer (material 1) has a higher CTE than the bottom layer (material 2), a decrease in temperature will result in the top contracting more than the bottom, thereby causing a deflection upward. Similarly, an increase in temperature causes the top to expand more than the bottom, causing a deflection downward. With sufficient temperature control, up and down motion is therefore possible for a thermal bilayer structure (these structures are also referred to as "thermal bimorphs"). Within the microelectromechanical system (MEMS) community, thermal actuation based on bilayers with differing CTEs is common and used in applications such as tunable electrical passives.

For metal origami, thermal bilayers are relatively uncommon, in part due to the popularity of shape memory alloys for thermal expansion discussed in more detail below. Most bimetal actuators generate a relatively large radius of curvature for bending, that is, a gradual curve rather than a "hinged" fold, as in the metal/polymer thermal actuator in studies by Kallaitzidou and Crosby. Bilayer actuators are most effective with stacks of dissimilar materials with large differences in expansion behavior, and many metals have CTEs that are too similar for making the small radii of curvature needed for true origami. While this can be combined with rigid sections as in the variable capacitor in studies by Reinke et al., the actuation sections remain relatively large and difficult to truly characterize as "hinges."

While thermal bilayer folding actuators have remained relatively uncommon in metals, there has been a more sizeable body of work in the area of residual stress-driven self-folding. Residual stresses are defined as stresses left within a material after processing. These stresses result largely from thermal mismatch between layers as well as changes in microstructure of the constituent material during specific processing steps. During microfabrication, different layers within a fabrication process experience repeated heating and cooling as they are taken through the steps of the process. Just like a bilayer thermal actuator, the individual materials attempt to expand and contract as the temperature rises and falls. However, the materials are constrained by the surrounding layers, resulting in a buildup of stresses within the layer. These stresses can be either compressive (i.e., pushing back against the layer to prevent expansion) or tensile (preventing the layer from contracting inward). If the constraint is removed late in the process (for instance, by etching away a supporting structural layer), the layers experience a release of the residual stress and resulting motion. For a compressive residual stress, the layer is now able to expand freely; removal of a tensile residual stress, on the other hand, results in the layer now being permitted to contract.
These residual stresses can be used to drive self-folding. A single isotropic layer will simply expand or contract without the out-of-plane deformation necessary for origami folding. Generating out-of-plane bending requires the use of two layers with different residual stresses or a single layer with a large stress gradient between the top and bottom. These differences in stress can arise through differences in processing or material CTE. In a single layer with a stress gradient developed during processing, the differential expansion between the top and bottom can create a 3D shape upon release, a technique that has been used to create helices and other complex arcs. A bilayer or multilayer stack, for instance of a highly compressed oxide stacked atop a less compressed or even tensile metal, will also create a 3D arc when released. If a flexible residual stress-driven bilayer structure such as chromium/copper hinges in studies by Bassik et al. is combined with thicker metal layers to define rigid regions, more traditional origami folding occurs upon release.

Figure 2b,c shows an example of copper/chromium hinges. Evaporated chromium is highly tensile relative to copper, as will be discussed in more detail in the materials section later. Coating regions of this film stack with nickel stiffens the “panels” of the substrate. The presence of photoresist in the hinged region prevents the substrate from folding until desired, used for triggering as will be discussed in more detail later. If the chromium is on the underside of the structure (Figure 2b), the hinge bends downward; likewise, shifting chromium to the top (Figure 2c) forces the hinge to bend upward.

Lithography or other microfabrication processes can define the geometry and location of the hinges, thereby allowing very precise control over the final structure. It is therefore straightforward to design more complex, multifold structures such as cubes (Figure 2d–f) and grippers on size scales as small as tens of micrometers. Adding a third layer also gives the potential of obtaining more complex folding behavior through the creation of mountain and valley folds. For example, additional chromium can be added to the chromium/copper film stack. As the chromium contracts upon release, this added layer creates a counter bending moment in desired regions, allowing folds in both directions.
Normally, residual stress-driven folding occurs upon releasing the structure from a rigid substrate. It is, however, not always desirable to have immediate folding for many applications (for instance, in the context of using a cube as a container or a gripper to grip), and there has been significant research on allowing controlled triggering of these types of structures. One possibility is to use temperature to trigger the folding motion. For example, a polymer layer on a hinge can constrain the hinge and prevent folding; when the polymer is heated to modest temperatures and allowed to soften, the folding is allowed to occur and the structure moves into its folded configuration. Likewise, a focused laser can selectively heat hinges and allow sequential folding of multiple nested cubes.

Thermal and other simple bilayer expansion techniques are relatively simple techniques for folding. Yet, the requirement of active heating to sustain folding is limiting. For this reason, it is more common to use residual stress (particularly for micro-scale applications) to induce folding as it lacks this drawback. This approach also offers the ability to be patterned lithographically and the ability to use standard cleanroom processes and materials. Residual stress-driven folding is, however, unidirectional and sensitive to processing environment and requires careful matching of material properties, which is its primary drawback.

2.2. Shape Memory Effect-Driven Actuation

Shape memory actuation has been widely investigated in part because it too does not require power to hold a shape in its folded state, despite being triggered by temperature. The shape memory effect is an intrinsic material property of specific alloy combinations as a result of a thermal transition from one solid phase to another solid phase with austenitic and martensitic crystal structure. Strains in the martensite phase occurs through repositioning of crystal twins within the material, defined as oriented crystals of the same chemical and crystallographic species. Because twinning realignment is reversible, unlike deformation through dislocations, this realignment enables repeated deformation without permanent material change.

When a shape memory alloy (SMA) is elastically deformed beyond the elastic limit from external forces while it is in a twinned martensitic phase, it will remain in the new deformed shape once the external forces are removed. When heated above the transition temperature to the austenitic phase, however, the microstructure will revert to the austenitic phase. This ability to “remember” the original, undeformed state after heating is the key defining characteristic of a shape memory material.

Each alloy material and mass fraction combination has a transition temperature, above which the material is in an austenite phase and below which it will be in a martensite phase. A typical actuation cycle is shown in Figure 3a). Beginning in an austenitic phase, the shape memory alloy shifts to a twinned martensitic phase upon cooling. With a mechanical load applied, the martensite begins to detwin, leaving deformation when the load is removed. Upon further heating, the shape memory alloy returns to its original austenite phase and initial shape.

Shape memory alloys also have another important property known as superelasticity. When in the higher temperature austenite phase, an applied strain can cause a transformation to the martensite phase (stress-induced martensite). The phase transition accommodates the strain, allowing higher deformations before permanent, plastic strains start to occur. Strains as high as 8% have been demonstrated for NiTi, the most commonly used shape memory alloy. Relative to other actuation materials, SMA actuators have one of the highest energy densities measured in work/unit volume that is in the order of $10^7$ J/m$^3$. The superelasticity also makes shape memory alloys particularly well suited for self-folding, as a shape memory hinge can have a very tight radius for a given hinge thickness compared to alternative materials.

While the energy density is high in the shape memory effect, the frequency of operation is moderate relative to some other actuator mechanisms. This is because the phase change of SMA is inherently thermal and thus driven by thermal time constants. Unlike thermal bimorph-driven folding, which has linear thermal strain response, the shape memory effect is also quite nonlinear. While this allows rapid shape changing about the transition temperature, it also means that shape memory actuators do not have the same ability to selectively reach arbitrary displacements. It is also geometry dependent, scaling up exponentially as length scales drop due to the surface-to-volume ratio for thermal energy flux in and out of the material.

While the shape memory effect had been demonstrated in certain alloys as early as the 1930s, practical use dates from the 1959 discovery of the unusual characteristics of Ni/Ti alloy when developing heat shielding materials at the Naval Ordnance Laboratory. Equiatomic Ni$_{50}$/Ti$_{50}$ remains one of the most common shape memory alloys and is commonly referred to as Nitinol (an abbreviation of Nickel Titanium Naval Ordnance Lab). Since its discovery, it has been widely studied and used across commercial and defense applications. Though the ubiquity of Nitinol cannot be understated, other formulations of SMAs also exhibit important properties and will be discussed in the following SMA materials Section 3.1.

A shape memory alloy actuator works by shifting back and forth between the two crystal states, returning to and releasing from a remembered state. For a MEMS SMA actuator, a common structure uses the deposition and heat treatment of an SMA film, followed by structure release, at which point it deforms due to residual stresses while in the martensite state. When the structure is then heated to the austenite state, the structure returned to its flat as-deposited state, providing the means for actuation. For self-folding origami, on the other hand, it is more common to manually program, fixing the actuator in place during the programming process. For instance, springs for external control of origami folding were programmed by wrapping a Ni-Ti wire around a steel rod and annealing to 500 °C for 15 min to create actuator springs for driving an origami structure in studies by Salerno et al. [35].

Rotary actuators intended to be integrated with origami sheets can be fabricated very similarly. For instance, in studies by Paik et al. [46] several low-profile rotary actuator designs were investigated. The designs were programmed by physically wrapping the shape memory alloy around a steel rod intended to set the radius of curvature and heating to program the actuator.
Using this approach, the authors were able to demonstrate radii of curvature of a millimeter or less. The actuator is then heated for actuation either through joule heating of the shape memory alloy itself or an attached Ni-Cr heater element.

The same group also demonstrated, in their follow up work, a bidirectional rotational actuator by using two different shape memory alloy structures in close proximity that could be independently addressed with attached heater elements. The actuator was again programmed similarly through furnace heating, this time with a custom multi-peg fixture for wrapping to allow the resulting more complex motion patterns. Figure 3b shows a simplified version of this bidirectional shape memory hinge. Before heating, the actuator hinge and figures remain flat. However, the hinge “remembers” a curved s-shape defined by the pegs in the furnace fixture. When one of the heater elements is actuated, it turns on only one of the two shape memory actuators, causing the associated actuator to return to its curved shape and flip the attached plate. Actuating the second heater causes the full s-curve to be remembered, causing the plate to be flipped back and returned to its original orientation.

The biggest advantages of shape memory alloy actuators are high energy density and low actuation temperatures without a need for static holding power. Shape memory actuators can also be designed for bidirectional motion, enabling robotic applications of this form of actuation. The reduced power consumption for actuation has made this technique particularly appealing compared with thermal bimorphs in most origami folding applications. The primary disadvantages are the need for pre-programming of motion through high temperature annealing as well as the requirement for the integration of specific specialized materials.

2.3. Magnetic Actuation

Magnetic fields are another common means of actuating origami structures for rapid self-folding. Like thermal actuation and the shape memory effect, magnetic actuators are capable of bidirectional motion. Within the MEMS and actuator community, there are several variants of magnetic actuators in use, but both the

![Figure 3. a) Example of a NiTi microstructure undergoing thermal and stress/strain cycle. Reproduced with permission. Copyright 2007, SAGE Publications. The material starts in an untwinned austenitic phase (A), moves to a twinned martensitic phase (B) upon cooling, and then moves to a detwinned state due to loading (C). The unloaded state remains deformed (D). Once heated austenitic formation begins (E) and finally, the material returns to the fully untwinned austenitic phase and original shape (F). b) Representative shape memory actuator for origami self-folding.](image-url)
simplest and most common in self-folding origami consists of incorporating a magnetically responsive material into the foldable structure and applying a magnetic field externally.

Understanding the mechanics requires a brief review of magnetics. Magnetic materials can be magnetized, which means that the internal magnetic dipoles are aligned (either permanently or temporarily), causing the overall magnetic body to have an effective magnetic dipole. When the magnetic dipole is then exposed to an external magnetic field, it experiences a magnetic torque, causing its own magnetic moment to attempt to rotate to align with the external field direction. The torque on the magnetic body, \( T \), assuming the magnetization is uniform, is given by

\[
T = \mu_0 \mu M \times H
\]

where \( H \) is the magnetic field, \( M \) is the magnetization, \( \mu_0 \) is the permeability of free space, and \( \mu \) is the volume of the body. Larger torques are therefore possible with a larger volume of magnetic material. A high magnetization vector also leads to a high torque on the actuator; because the maximum magnetization within a material is set by a property known as the saturation magnetization, high saturation magnetic materials are generally preferred.

An early example of an out-of-plane magnetic actuator is shown in Figure 4. The actuator consists of a permalloy plate deposited on an elastic polysilicon beam that acts as a rotational spring (Figure 4a). The nickel must first be magnetized, orienting the dipoles within the material to set the magnetization vector through an applied magnetic field. When the magnetic field is applied at an angle relative to the original magnetization (Figure 4b), torque is generated on the actuator that causes the nickel plate to rotate away from the surface (Figure 4c). The plate with different levels of magnetic field applied is shown in Figure 4d-f. As long as the magnetic field is applied, the plate will remain folded; however, when the magnetic field is removed, the rotational spring provides a restoring force that causes the plate to rotate back to its original position, thereby allowing bidirectional motion.

Magnetic actuation is capable of very large forces (hundreds of Newtons or higher), although the forces scale poorly with small volumes; thus, the technique is most appropriate for structures that are at least a few micrometers in size. It is also possible to time a series of folds to allow complex folding, for instance, by designing the sheet of magnetic material to actuate when the magnetic field reaches different thresholds or combining with electrostatic forces to address individual magnetic devices. As with shape memory alloys, however, this type of actuator typically requires the integration of specific materials, in this case one that can be easily magnetized.

While shape memory alloys are capable of maintaining their shape without an applied heat source, the same is not necessarily true of a magnetic actuator. When the magnetic field is removed from the flap in Figure 4, the force on the nickel will be removed. For small angles of displacement as mentioned in, the deformation is elastic and magnetic forces can be used repeatedly for actuation without permanently changing the structure. For larger displacements, however, the flexible hinges begin to experience plastic deformation, experiencing permanent changes within the hinge. When the structure is rotated to an angle through a magnetic field, the actuator will remain at that angle as long as the magnetic field is applied. When released, the elastic portion of the strain will be removed, causing a partial fall back toward the initial position; the plastic deformation will remain, however, causing the structure to remain at a rest angle, a fabrication process known as plastic deformation magnetic assembly (PDMA). Used in this operating regime, magnetic actuation becomes a unidirectional process that is more suitable for fabrication than for reconfiguration.

In addition to relying on plastic deformation to hold the actuator in place, there has also been interest in holding the structure at its fully actuated position through other mechanisms, rather than allowing it to partially fall back to a rest position upon release. For instance, a magnetic plate can be folded followed by thermally welding through an embedded heater resistor. A similar latching effect was obtained by combining magnetic folding with using the magnetic field itself to melt an attached bead of solder for anchoring the folded panels in place.
The prior works use a magnetized block of magnetic material to generate the magnetic moment for actuation. While this is by far the most common means of creating folds magnetically, it is possible to generate a magnetic moment electrically. When a current passes through a wire, it generates a magnetic field; a loop of wire (an “electromagnet”) can also be used to create a magnetic moment. Just like the magnetic material case, a torque for folding can be generated by exposing the electromagnet to an external magnetic field. The electromagnet will rotate to attempt to align with the external field.

This approach was used for folding and latching MEMS corner cube structures, with the structure consisting of gold hinges and electromagnet, with the photoresist SU-8 used to create the body of the rigid plate.[59] The electromagnet is a single turn coil and was demonstrated to generate tens of microNewtons of force in a Tesla-scale external magnetic field, which is enough to fold plates vertically and latch them in place. The use of the electromagnet approach has the primary advantage of removing the need for specialized magnetic materials, thereby allowing easier microfabrication. It does, however, carry with it the disadvantage that the folding process requires electrical contact and significant currents within the electromagnet (up to 50 mA in studies by Shaar et al.[59]).

Overall, magnetic actuation has the advantages of being bidirectional and of being relatively easy to actuate through a neighboring permanent magnet or electromagnet coil. Its biggest drawback, however, is the requirement to integrate specialized magnetic materials, thus requiring additional processing and limiting options for the properties of the final structure.

2.4. Actuation by Laser Forming

Laser forming is nearly material independent and can be performed even on a blank unprocessed sheet metal. Similar to the other metals, it is a thermally driven process, yet the energy is delivered using light. Actuators based on both bilayer thermal expansion and shape memory alloys are typically actuated at low to moderate temperatures, with the intent of avoiding plastic deformation and permanent changes in the material. In laser forming, these permanent changes are fundamental to the mechanism itself. Defined as the use of a laser for localized heating to generate plastic deformations within a workpiece,[60] laser forming developed out-of-flame bending for ship construction in the mid-1980s[61] and remains the most commonly used technique for large-scale sheet metal applications. Other notable advantages include cutting that can be easily integrated into the process with the same laser and the ability to selectively deliver heat with the focused laser that removes the need to preprocess or pattern the target.

Within the laser forming community, the term “laser forming” is an umbrella term encompassing several specific mechanisms reliant on temperature profile and workpiece geometry, with the most important known as the temperature gradient mechanism (TGM), the buckling mechanism (BM), and the upsetting mechanism (UM).[62] For the purposes of self-folding origami, only the temperature gradient and buckling mechanisms are relevant. The upsetting mechanism is defined by the shortening of the length of a metal plate through laser-induced plastic strains.[62] Because this mechanism is not generally used for origami folding, it will not be discussed further here.

In the temperature gradient mechanism, the laser is scanned across the target relatively quickly, creating a large difference in temperature between the top and bottom of the workpiece (Figure 5a). This temperature difference results in an initial thermal expansion of the top layer bending downward away from the laser source, a phenomenon known as “counterbuckling”[64] (Figure 5b). As the top expands, the cooler surroundings resist the motion, resulting in the building of plastic compressive strains within the workpiece (for example, the bottom of the sheet in Figure 5b).[65] After removing the laser, the top is able to cool and again contracts, but due to these additional compressive plastic strains, it contracts by more than the original expansion, resulting in a bending moment toward the laser source (Figure 5c). The TGM therefore always causes bending toward the laser source, that is, a valley fold.

For laser forming using the buckling mechanism, the laser is instead scanned relatively slowly, with sufficient time to allow the heat to propagate all the way through the thickness of the sheet (Figure 5d). The heated region again expands, forming compressive stresses in the plane of the sheet as the heated region pushes outward against the surrounding cooler regions. These stresses cause the membrane to buckle when the stress exceeds a value known as the critical buckling condition (Figure 5e).[66,67] The buckle then propagates along the travel path of the laser (Figure 5f), resulting in folding (Figure 5g). Unlike TGM, the direction of buckling is not predetermined, and BM bending can occur either toward or away from the laser beam.[68] For a controlled process, it is therefore necessary to induce a preferential buckling direction, for instance, by providing a prestrain in the workpiece.[62]

For origami folding, TGM is generally used for bending toward the laser source, while buckling is used to cause folding in the opposite direction. This allows up and down folding in the same part using the same laser. Because the TGM and BM modes are based on the type of temperature distribution in the material, it is possible to switch between them by using the laser travel speed.[68] A fast travel speed results in a vertical temperature gradient and upward folding based on TGM, while slowing down the travel speed gives time for the heat to propagate through the thickness leading to BM folding. When combined with laser cutting, laser forming allows complex 3D parts to be rapidly folded directly from metal sheeting (Figure 5h–j).[68,69]

Laser forming is most important for being material independent and capable of both up and down folds, without any need for preparing the substrate. It relies on plastic deformation and is a strictly unidirectional actuation mechanism. The other notable disadvantage of this technique is that it requires very high, localized heating, which can cause some limited material losses and warping; it is therefore most commonly attempted at larger scales for structures at least tens of micrometers thick and hundreds of micrometers wide.

2.5. Solid/Liquid Phase Change-Driven Actuation

At the microscale, laser forming is less practical. However, self-folding is still possible. At these size scales, one of the most
common techniques for origami self-folding is known as fluidic self-assembly. Surface tension is the primary driving force. When liquids and air come into contact, the liquid molecules are more strongly attracted to each other than to the air, resulting in an excess of surface energy at the interface, known as surface tension.[70] The surface tension results in a tendency in liquids to ball up to minimize surface area and the resulting surface energy. With units of force per meter (N/m), surface tension can exert a force on attached solids, for example, in causing deflection and stiction in MEMS fabrication.[71] At large size scales, the surface tension is relatively modest and easily overwhelmed by gravity and other forces. However, the surface tension becomes increasingly more important at smaller scales (typically sub-millimeter)[72] and is widely used for micro- and nanostructure assembly.

In fluidic self-assembly, melting of one material includes folding.[73] Figure 6a–c shows the basics of a fluidic hinge before, during, and after self-assembly. The structure consists of a more rigid plate such as thick electroplated nickel deposited on an underlying substrate, with an elastic hinge created using a thinner layer of metal. The solder for actuation can first be deposited as a solid (Figure 6a), allowing it to be incorporated into a conventional microfabrication process through standard photolithography[74] or as a solidified liquid through dip-coating.[75] When the hinge melts, the surface tension $F_s$ from the liquid pulls inward (Figure 6b), applying a bending moment and causing the suspended component to rotate upwards. The flexible hinge layer provides a counter force $F_e$. For the design in Figure 6, the component rises until the surface tension is completely balanced by the elastic force, leading to the final 3D shape (Figure 6c). Alternatively, if the elastic layer is omitted (in a so-called “hinge-less” design),[76] the final position is set primarily by the minimization of surface energy of the fluid hinge, removing the need to overcome elastic forces. Fluidic self-assembly of this form can be done with both metals and nonmetals, as long as the melting point is low enough to avoid damaging or warping the rest of the structure. In addition to sharp hinges, a similar approach can also be used with a continuous layer of meltable metal to create more gradual arcs as well.[77]

One of the primary advantages of fluidic actuation is the ability to control folding even at very small size scales. Figure 6d shows one example, a 15 μm-on-a-side cube consisting of nickel sides and solder hinges,[74] while Figure 6e shows a similar cube with 500 nm gold/alumina sides and tin hinges.[78] Melting of solder joints is fundamentally a heat-driven process, and if the overall structure is simply heated (through a hotplate or oven heating), all the joints will melt, resulting in complete self-folding in a single stage. It is, however, possible to selectively address the individual joints, for instance, through focused laser heating, to melt a single joint at a time,[79] thus allowing sequential folding.

Fluidic self-folding is important in MEMS and nanoelectromechanical systems (NEMS), and is particularly easy to incorporate into microfabrication. Hinges can be deposited in solid form, and the structure remains planar until final folding at the end of the process, thereby allowing spin-coating and other wafer-level processing. It is, however, fundamentally unidirectional, like laser forming, and is not suitable for larger size scales where other forces begin to dominate. Using fluidic self-folding can also limit complexity of a microfabrication process in certain ways; hinges must be deposited relatively late within a fabrication process, as it immediately prevents high temperature process steps to avoid folding until desired.
2.6. Ion Implantation-Driven Actuation

In addition to the earlier mechanisms, several other approaches have been investigated for self-folding of metals. One method involves inducing plastic strains through exposure to high energy ions to generate folding.[80] Focused ion beam (FIB) tools are common in semiconductor processing, widely used for both etching and material deposition. For folding, the ion beam is scanned across a hinge to inject ions into the structure. As with laser forming, two mechanisms have been demonstrated (Figure 7a).[80] At low energies, atoms are expelled upward (toward the source) from the bulk, resulting in the material folding upward into the vacated region. At higher ion beam energies, the ions inject deeper into the bulk, pushing material downward from the bottom of the sheet and triggering downward folding. The folding can therefore be used to create classic origami structures such as folded cubes.[81] In many respects, this has similar advantages to laser forming; it is relatively material independent, is capable of up and down folding and is based on permanent, plastic deformations in the target. Ion implantation folding is, however, limited to very small size scales, with demonstrations in the order of a few micrometers, and focused ion beam setups are relatively expensive and specialized, thereby limiting broader use.

2.7. Controlled Buckling-Driven Actuation

When a thin membrane or structure experiences significant compressive stresses, it buckles out of plane. In the case of laser folding, buckling is realized using locally focused heating. It is also possible to generate globally compressive stresses across a much larger structure and use this buckling to perform folding. One way to achieve these compressive stresses for controlled mechanical buckling is through release from a pre-stretched elastomer film.[82] As a thin ribbon is released from a stretched substrate, the ribbon experiences the significant compressive stresses needed. This does not necessarily create sharp folds; simple buckling of a metal layer generates more gradual bends. This can, however, be addressed by creating more flexible regions within the structure; in studies by Yan et al.[82], creases for folding are formed through the patterning of thinned regions for hinges that preferentially fold during the buckling process. One advantage of compressive buckling driven folding is that it can be performed with a variety of materials, not just metals, though it has been demonstrated for both metal films (Figure 7c)[83] and polymer/metal bilayers.[82] As with other stress-driven methods, however, it is also unidirectional and not capable of easily undoing the resulting motion.

2.8. Actuation through Chemical-Driven Folding

While bilayer metal actuators are most commonly thermally driven, it is possible to obtain differential expansion through a different stimulus such as chemical environment. This modality is relatively common for polymer-based folding,[1] but the lower absorption and therefore expansion of metals has limited this mechanism in the metal self-folding community. When metals are included, for instance, in the polymer-metal bilayer hinge used to obtain folding when exposed to electrolytic solutions in studies by Liu et al.[84], the polymer generally serves as the expansion layer while the metal remains stable. There have, however, been occasional exceptions. Perhaps the most notable is oxidation-driven actuation.[85] A copper/chromium hinge, when heated in the presence of oxygen, begins to oxidize the copper. The oxidation causes a volume expansion on the copper surface to cause bending. After shifting to a
hydrogen environment and again heating, reduction occurs and the oxygen leaves the surface, causing the surface to convert back to copper. Folding and unfolding of simple origami cubes and grippers could then be demonstrated. This technique is particularly notable for allowing a residual stress actuator, normally unidirectional in nature, to be used for bidirectional actuation. The oxidation-driven structure was therefore used for robotic structures such as folded grippers that would not have been otherwise practical.

2.9. Actuation by Folding Metals with Polymers

Although this review focuses on the folding of metals, and specifically on self-folding mechanisms driven by the metals themselves, it is important to mention that metals can also be induced to fold by incorporating them on other actuation layers such as polymers. One approach is to use a polymer film in a thermal bimorph film stack, as described previously (e.g., in studies by Luo et al. and de Leon et al. 

Because the shape memory effect in polymers is similarly based on temperature change, actuation relies on selective heating and cooling. It is possible to induce sheets of shape memory polymer to self-fold by locally delivering heat to hinges. This is possible by first patterning black ink onto the sheet, as shown in Figure 8a. When exposed to intense light, the black ink absorbs the light more effectively than the transparent sheet. Thus, it increases in temperature and allows the polymer to shrink below the inked patterns. The gradient in temperature through the hinged region causes folding to occur, as shown in Figure 8b. This principle has been harnessed previously to co-fold metal films attached to the polymer. For example, Figure 8c shows a copper film deposited on a self-folding polymer sheet. When exposed to light, the black ink absorbs the energy, causing the polymer to fold. Because the metal is attached to the polymer, it also folds.

Figure 7. a) Ion implantation driven folding mechanisms. Reproduced with permission. Copyright 2017, American Chemical Society. b) ion-folded cube. Reproduced with permission. Copyright 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, and c) buckling-driven assembly of an origami cube. Reproduced with permission. Copyright 2017, Elsevier.
Driving metal with polymer actuators has the primary advantage that it takes advantage of the large deformations possible for polymers, and there is a large body of work on self-folding with these materials. However, this method also does not take full advantage of the metal properties. Metals in a polymer actuator must be sufficiently thin or structured to avoid interference with the motion of the polymer layer, thereby limiting the minimum electrical resistance of the resulting conductor. Folding in this manner also does not take advantage of the mechanical properties of metals at all, as it requires the polymer to be dominant for actuation. The individual polymer folding mechanisms also have their own limitations; for instance, the technique in Figure 8 requires a bright light source and the requirement to have the absorbing material (here, black ink) accessible to the light, thus limiting possible geometries.

3. Materials

Once a suitable actuation mechanism has been identified for a given application, the next step is deciding what materials to use to build it. Most self-folding mechanisms are driven heavily by material properties, some so much so that very specific materials or material classes must be integrated into the structure. A review of fluidic self-folding would be notably incomplete without a discussion of solders, one on magnetic actuation sadly lacking without a discussion of magnetic materials, and so on. Even actuation mechanisms like laser forming still rely heavily on specific material properties such as thermal conductivity and reflectivity. Appropriate choice of materials is therefore one of the most important considerations for a designer and will be surveyed in the following section.

3.1. Materials for Thermal Expansion

Both thermal bilayers and residual stress-driven structures are primarily based on differences in expansion between layers, typically expressed in terms of their coefficients of thermal expansion (CTE). What, then, determines a material’s coefficient of thermal expansion? This property is closely related to how tightly the molecules within the solid are bound together, and therefore with another associated property, the melting point. As a general rule of thumb, the higher the melting point, the less is the expansion that results for a degree of temperature change (i.e., the lower the CTE). This result allows us to more generally predict which classes of materials will have larger CTEs. In general, polymers melt at lower temperatures than metals, and metals at a lower temperature than ceramics; the CTEs, as a result, tend to be highest in polymers and lowest for ceramics, with metals somewhere in between.

Because differences in thermal expansion, and therefore CTE, cause out-of-plane deflections for thermal bilayers, there is an incentive to make a bilayer consisting of a metal and a polymer, or a metal and a ceramic, to create a structure with very different CTEs on the top and bottom. While combining different material types can be one way of obtaining a large difference in CTE, it is also possible to obtain reasonable differences even with metals alone. Metals like mercury or certain gallium alloys are liquid even at room temperature; other metals can survive temperatures of thousands of degrees Celsius before melting, with corresponding differences in the corresponding thermal expansion.

Residual stress-driven systems are the most common bilayers seen in the metal origami field and deserve special mention. Residual stresses derive from two sources; differences in thermal expansion and other effects occurring during processing such as lattice mismatch or doping, with the components of the stress...
not directly explained by thermal mismatch collectively referred to as intrinsic stress.\[92\] The Gracias group has been very active in this area, making a variety of cubes and other polygons driven by residual stress in bilayer films.\[10\]–\[13\] Most commonly, chromium and copper are used for driving self-folding process because these films have sizeable differences in residual stress when deposited onto a surface.\[14\]

Chromium has a relatively high melting point for a metal, i.e., 1890 °C,\[93\] while that of copper is far lower, i.e., only 1083 °C.\[94\] This difference results in very different CTEs; chromium has a CTE in the order of $5 \times 10^{-6}$ near room temperature,\[95\] compared with about $15 \times 10^{-6}$ for copper.\[96\] However, the deposition process also plays a major role in the intrinsic stress of the resulting thin films. Evaporated chromium is known to experience significant tensile residual stress when deposited on a substrate at room temperature, well above what would be expected from thermally induced stress alone.\[97\] This is not true of all chromium films; the same study showed that sputtered chromium has negligible stress beyond what would be expected from thermal stresses. Combining chromium’s intrinsically high tensile intrinsic stress with evaporated copper, which has significantly lower stress, can therefore achieve sizeable differences in expansion, thus driving the self-folding behavior.

In general for both residual stress and more conventional bilayer actuators, the primary material goal is to maximize the difference in expansion. For a conventional bilayer, this is done most simply by selecting very different coefficients of thermal expansion, either by selecting metals with very different melting points (as in the chromium-copper bilayers above) or by combining a metal layer with a different material such as a polymer with significantly different CTE.\[24\] This aspect remains a major consideration for residual stress-driven devices as well, due to the dependence of residual stress on thermal strains, but variations in intrinsic stresses due to different deposition processes also play a major role.

### 3.2. Shape Memory Alloys

The first shape memory effect was demonstrated in a gold-cadmium alloy in 1932, but it was only with the discovery of the large strains possible for the nitinol and the nickel-titanium system that shape memory alloys became attractive.\[98\] Crystal structure of the metal alloy (the presence of solid martensitic and austenitic phases in the phase diagram in a reasonable temperature range) is the main factor that determines whether a given material experiences the effect. The martensitic transformation that drives the shape memory effect is a shear-driven displacement of the crystal lattice, rather than atomic diffusion.\[39\] The number of different shape memory alloys in active use remains relatively small. Binary Ni-Ti, in various atomic percent compositions, was the first major shape memory alloy in active use and remains the most common commercial SMA material today.\[141\]

Actuation temperature is generally the most important consideration for self-folding origami. A higher temperature requires more power to actuate, while using a temperature too low can make the actuator sensitive to environmental conditions. One of the biggest advantages of the nickel-titanium system has been that the possible range of temperatures spans from significantly above room temperature to far below, thus allowing use in a relatively broad application set. The relative atomic percentages of nickel and titanium can therefore be used to adjust the nominal solid-solid phase transition temperature. The phase transition temperature is the temperature required to convert the material from the martensitic phase to the austenitic phase. The transition temperature is, however, highly sensitive, requiring careful control of the atomic percentages to achieve a desired actuation temperature. Between 50% and 51% Ni, small changes in relative atomic percent Ni/Ti ratio result in drastic changes in transition temperatures from 60 °C at 50% to −100 °C at 51%.\[99\] The transition temperature is ≈60 °C from 48.5% to 49.7% Ni.

Therefore, the desired activation temperature must be well understood prior to obtaining and specifying the material composition. For example, common activation temperatures for medical use are just below the body temperature for use in cardiovascular stents or orthodontic archwires.\[100\] Another issue with Ni-Ti material is that to set the desired shape, it must be held in the desired position to crystallize at a temperature of around 500 °C. For practical applications, these thermal considerations must be compatible with other materials in a folding metal system, like polymers or circuitry. When designing a folding system, often times the shaped NiTi is assembled after it is shape set. Finally, a major advantage of binary Ni-Ti is its proven biocompatibility and FDA approval for in vivo use.\[100\]

Though binary Ni-Ti is the most understood and commonly used shape memory material, other alloys also exhibit the shape memory effect. Adding a third element can be used to shift the actuation temperature range higher as well as to tune other properties such as fatigue resistance.\[31\] Alternative material systems have also been investigated to attempt to reduce cost,\[11\] with the copper shape memory system particularly studied. Several binary and ternary copper-based formulations include Cu-Zn, Cu-Zn-Al, Cu-Al-Ni, and Cu-Sn.\[11\] Of these copper-based formulations, only Cu-Zn-Al and Cu-Al-Ni are of significant commercial importance relative to Ni-Ti.\[99\] Due to the lower cost of these materials relative to Ni-Ti, they show promise but only where a lesser shape memory effect can be tolerated.

Because nickel-titanium is by far the most commonly used shape memory alloy available, and one that is easily available commercially, most efforts within the self-folding community has been limited to this material. Shape memory alloys are also highly sensitive to precise atomic ratio, and purchasing sheets or wires eliminates the need to develop such capabilities elsewhere. The shape memory rotary hinge actuators discussed previously used commercial NiTi sheets with a transition temperature of 65 °C for an easily obtained actuation temperature but one significantly above room temperature, while commercial NiTi wire was used in the external shape memory actuator in studies by Salerno et al.\[45\] There are, however, potential benefits to alternative shape memory alloys including lower cost, higher temperature range, and better fatigue resistance,\[41\] and investigation of a broader set of materials is a promising future research area. Magnetic shape memory alloys,\[106\] alloys that use magnetic energy instead of heat for transformation, also could open up new possible actuation mechanisms.
3.3. Magnetic Materials

Most materials respond at some level to magnetic fields, but creating a meaningful actuator generally requires the use of more specialized magnetic materials. One of the defining characteristics of a magnetic material is its ability to be magnetized, to align magnetic dipole moments within the material in the presence of a magnetic field. Magnetic materials are traditionally categorized into two sub-sets: soft and hard magnetic materials. Both can be magnetized in magnetic fields; however, a soft magnetic material demagnetizes when the applied magnetic field is removed, while a hard magnetic material, once magnetized, remains aligned even when the applied magnetic field is removed.\textsuperscript{102} Both are widely used in different applications: for instance, hard magnetic materials used as permanent magnets in motors and generators, while soft magnets are critical for inductor and transformer cores.\textsuperscript{103}

From the perspective of the origami designer, one important question is which is preferable, a hard or a soft magnetic material? The answer comes largely from understanding the forces on a magnetized body. When a magnetized material on a hinge, as in an origami fold, is exposed to an external magnetic field at an angle relative to the magnetization direction, a torque results. The magnetization direction of the material is pulled closer to the angle of the external field.\textsuperscript{104} As both hard and soft materials can be magnetized, it is possible to make an actuator from either one. From the mechanics perspective, the primary distinction is that a harder magnetic material, a permanent magnet, does not change its magnetic moment in response to the magnetic field, allowing simplified analysis of the resulting motion.\textsuperscript{104}

Based on the earlier section, it may come as a surprise that magnetic self-folding origami is almost universally done with soft magnetic materials. The reason is that even though both are functional and hard magnets are simpler to design, soft magnetic materials are easier to fabricate, particularly at the micro-scale where magnetic origami is most commonly used. Hard magnetic materials such as rare-earth permanent magnets\textsuperscript{105} are challenging to make both because their performance is very sensitive to microstructure and chemical composition, and because it is necessary to “pole” the magnetic material before use.\textsuperscript{106} "Poling" means to set the direction of the magnetization, for instance, by applying a large magnetic field during the fabrication process.\textsuperscript{107} Common soft magnetic materials, on the other hand, include the single element metals iron and nickel as well as a number of their simple alloys, and can be deposited easily onto micrometer-scale features through electroplating.\textsuperscript{108}

Most of the examples of magnetic self-folding origami surveyed in this review were done with one of two materials, either nickel or permalloy. Nickel is often chosen as single element materials are easier to electroplate and nickel has far better corrosion resistance than iron.\textsuperscript{108} Permalloys are iron/nickel alloys with between 40% and 80% nickel composition,\textsuperscript{109} with particularly useful magnetic properties. Notably, permalloy with 79% nickel and 21% iron is important for having an order of magnitude higher permeability than iron or nickel alone, as well as very low magnetostriction.\textsuperscript{110} Magnetostriction is deformation in the magnetic material resulting directly from the applied magnetic field; because this mechanism is not the driving mechanism of motion for these devices, minimal magnetostriction is preferred for self-folding.

As the strength of the magnetic moment sets the amount of torque within a magnetic actuator, magnetic response is the primary consideration. There is, however, much scope for improvement in material selection. The self-folding origami community has preferred a handful of classic materials, but there has been much research elsewhere in new magnetic materials in recent years, including a number with superior performance in many aspects.\textsuperscript{103} Incorporation of materials with larger magnetic response could enable larger forces and new capabilities in the future.

3.4. Laser Formed Metals

Based on simple thermal expansion and the resulting plastic stress buildup, laser forming can be performed on a relatively broad material set. Numerous metals have been laser formed such as aluminum,\textsuperscript{111} titanium,\textsuperscript{112} nickel,\textsuperscript{68} stainless\textsuperscript{113} and other\textsuperscript{114} steels, copper alloys,\textsuperscript{115} and nickel-titanium shape memory alloy.\textsuperscript{116} Although rarer, there have been demonstrations in nonmetals such as crystalline semiconductors and glass and ceramics.\textsuperscript{117} It is therefore not nearly as reliant on integration of a few specific materials or material classes as many of the other origami processes.

This does not, however, mean that material selection is not important for laser forming. Although the process is relatively general, there is a significant dependence on bending rate with material choice. Laser forming is a sufficiently complex process that simple analytical models able to completely capture the behavior have remained difficult and computational modeling has been preferred.\textsuperscript{68} However, it is possible to make observations on the relationship of material properties to bending performance. For instance, the bending angle per laser pass in TGM bending is related to the material heating by\textsuperscript{60}

\[
\theta_{\text{pass}} \propto \frac{\text{CTE} \cdot P \cdot A}{\rho \cdot c_p}
\]  

where CTE is the coefficient of thermal expansion, \( P \) is the laser power, \( A \) is the material absorptivity, \( \rho \) is the density and \( c_p \) is the material specific heat capacity.

Understanding the expression is straightforward. The heating is given by the total energy inputted into the system divided by the heat capacity (specific heat multiplied by mass), the heating per unit of energy. While the total energy is directly related to the power level of the laser, the absorptivity (the fraction of the incident power absorbing into the workpiece) ultimately sets the energy that makes it into the target. The composition of the metal and the wavelength of the laser dictate absorbance. Thus, switching target materials can have a sizeable effect on bending at a given laser power level. It is also possible to surface treat the samples to obtain better absorption. CO\(_2\) lasers are relatively cheap and widely available, but absorb poorly into metals; as a result, it is common to coat metals with an absorbent coating such as sprayed graphite to allow their use for laser forming of sheet metal.\textsuperscript{61} The CTE then defines the amount of expansion per degree of heating.
From the perspective of material choice, therefore, materials with high reflectivity (one minus absorptivity for an opaque material) require a higher laser power to obtain the same degree of folding. This reflectivity can vary significantly for different metals at the same wavelength. For instance, with an Nd:YAG laser, one of the most common laser types for metal ablation, the reflectivity at lower laser powers is 0.9, 0.72, and 0.6 for copper, nickel, and tungsten, respectively. A copper sheet therefore requires four times as much power to fold at that wavelength than a tungsten one.

Laser forming is a thermal process, and from Equation (2), it is natural that thermal behavior also plays a prominent role. A high CTE combined with a low heat capacity will therefore tend to bend most rapidly. While not present in (2), the thermal conductivity is also a prominent factor in laser forming. Laser forming relies on thermal gradients in the workpiece, with TGM bending in particular heavily based on the ability to create a vertical temperature gradient across a thin metal plate. Scanning the laser across a workpiece too slowly can allow heat to propagate all the way through the thickness, resulting in buckling mode bending instead. The needed laser scan rates to generate this vertical gradient is therefore significantly higher for a high thermal conductivity material like copper than a lower conductivity material such as stainless steel.

Notably, the mechanical properties are not present in the most basic laser forming models such as (2). While accurate to some degree, these models do neglect important second order effects. For instance, differences in the elastic strain at the yield point have a sizeable effect on the resulting bend per laser pass. Because the laser forming mechanism depends on plastic deformations, a material able to undergo large elastic strains without plastic deformation is therefore going to bend less than the one that yields more easily.

In summary, a wide range of metals can be laser formed, although with sizeable variability in their individual folding performance. Laser reflectivity and thermal properties are the dominant parameters defining laser forming behavior, with mechanical properties playing only a secondary role.

3.5. Solders and Fusible Alloys

With fluidic self-folding, the most important material choice by far is the fluid selection. For metal origami, this almost invariably means solder. A solder is formally defined as a metal alloy with a melting point in the range 90 to 450 °C, a range set largely by electronics industry needs (stable at or near room temperature, but easily melted when desired). Fusible alloy is a related term that refers to alloys that are easily melted (i.e., fused); the term is sometimes used more generically, but the American Society for Metals (ASM) defines fusible alloys as a metal within specific families of alloys that melt at a lower temperature than tin-lead solder (183 °C). While fusible alloys and solders include slightly different temperature ranges, there is sufficient overlap that the terms are sometimes used interchangeably in practice. Many of the metals used for fluidic self-folding qualify as both.

Several aspects come into play in selecting a solder. The most important is typically the melting temperature. Solder-driven folding is most often used as a step in fabrication for structures such as electrical devices. The folding temperature must therefore be sufficiently high to avoid folding during standard microfabrication steps such as photore sist processing. It is also important to avoid melting during regular operation, because the hinge is intended to become a solid structural component of the final device, and melting points significantly above room temperature are preferred. It is, however, desirable to avoid very high temperatures to minimize thermal effects on the rest of the final system. The needed temperature range is often similar to what would be expected of solder for electronics packaging. Most work in fluidic self-folding is in the range of about 100 to 350 °C, most commonly various forms of lead- and tin-based solders.

The other major consideration is the generated force. Solder-driven self-folding is based on the surface forces of the melted solder, with the magnitude of these forces set by the surface tension of the melted fluid. The surface tension derives from an imbalance between the repulsive and attractive forces between the individual molecules on the surface of the liquid, and can therefore be correlated with other physical properties of the metal. Notably, the surface tension in a liquid significantly above its melting temperature generally falls as the temperature rises; in a pure metal, higher melting points also correlate with higher surface tension. Unfortunately, the surface behavior of multicomponent alloys is far more complicated, and impurities with lower surface tension tend to build on the surface; thus, there is no satisfactory theory for the surface tension of a liquid multi-component metal alloy. Due to the significant research in this area, however, detailed tables of surface tension and melting points of commonly used alloys have been compiled. In general, low melting point metals have a relatively high surface tension, between 0.3 and 0.6 N m⁻¹, compared with 0.073 for liquid water and even lower for many common organic solvents and polymers. Melted metals can therefore be capable of nearly an order of magnitude higher actuation forces than common alternative fluids.

Aside from melting point and surface tension, another major concern has been toxicity. In particular, leaded solders were historically very common in electronics packaging, but due to environmental concerns are in the process of being phased out. Due to these concerns, non-leaded solders have increasingly been preferred, with tin and tin-based solders the most common. Because solder-based folding is most common at the MEMS/NEMS size scales, ease of processing is a final important consideration, and drawing on the work in the electronics community, many solders can be sputtered or electroplated for integration into a process flow.

Selecting a solder is often a matter of deciding a desired temperature of actuation. Once the temperature range has been determined, it is relatively straightforward to select from the more limited range of solders that melt at the appropriate temperature for obtaining maximum force and maintaining low toxicity and process compatibility.

3.6. Liquid Metals

In addition to the actuation materials discussed earlier, there has also been a need within the self-folding community for highly flexible or foldable metals, particularly within the context of
polymer-driven actuation. Because most metals are significantly stiffer than polymers, incorporation typically requires the metal layer to be made very thin to allow the polymer to freely fold. Most commonly, these metals are intended for electrical conductors, and thinning results in relatively poor electrical performance.

To address this concern, there have been efforts to instead substitute an inherently stretchable conductor such as a room temperature liquid metal within the origami hinge. Liquid metals are compelling materials for origami because of 1) they maintain electrical conductivity during deformation and 2) they provide essentially zero mechanical resistance to folding. In addition, liquid metals remain compliant even for significant cross-sections, allowing a low resistance electrical conductor to be integrated without preventing folding.

Mercury is a common and well-known liquid metal, but its toxicity limits its utility in many applications. Gallium and gallium-based alloys have melting points near or below room temperature, and do not have the toxicity concerns associated with mercury. Furthermore, they have water-like viscosity (suggesting that they can flow readily) and negligible vapor pressure (suggesting that they will not evaporate). Importantly, these metals rapidly react with air to form a very thin (≈3 nm) oxide layer on their surface. This solid-like “skin” helps to hold the metal into non-spherical shapes that would normally not be allowed by interfacial tension. Thus, it can be patterned into electrodes and interconnects by a variety of techniques such as 3D printing, inkjet printing, and stencil printing. Consequently, the metals can be patterned across hinges and be utilized for stretchable circuits, sensors, and foldable devices that maintain conductivity when folded.

Using a liquid conductor can also cause several additional challenges. One consideration when using such materials is that they must be encased in polymer to avoid metal leakage or smearing when contacted. For example, the metal can be encased in an elastomeric coating, as a soft elastomer can also freely deform during hinge motion. The common liquid gallium alloys, galinstan and eutectic gallium indium, are also more resistive than the best conventional conductors such as copper by about a factor of 17. The ability to maintain stretchability for a liquid metal trace even for significant cross-sections means that the resistance of a trace over the hinge can still be made significantly lower than traditional metal alternatives.

4. Control and Design

When we look at a simple origami cube, it takes little imagination or design to understand the necessary sequence of folds necessary to bring it to reality. The same very likely does not hold true for origami with hundreds or even thousands of folds. Even the simple cube itself can be folded using a variety of different fold patterns. It has been proven that even without resorting to cuts, it is possible to form any connected polygon with or without holes using a single piece of paper, but that result is of little use without technology to enable us to understand the sequence. With the great advances in computing power in recent decades, it has become increasingly possible to use computers to design far more complex origami than would be possible with the human brain alone. In this section, we examine design methods for self-folding of origami, from the design of individual fold lines to the techniques used for creating complex origami through computational methods.

4.1. Fold Design

Understanding origami folding begins with the individual fold. It is difficult to imagine an ornate self-folded shape without deep insight into how much a plate will move for a given stimulus. It is therefore critical in most cases to obtain consistent and repeatable angles, and to know these angles in advance. Fold design and modeling has been widely studied, and there is no lack of useful resources in understanding the folding for a given mechanism.

The first step is developing an analytical understanding of the fold, developing simplified models of the folding forces and behavior. Each specific mechanism is based on distinct physical phenomenon, and a full analysis in each case is beyond the scope of the current review. Both surface tension and magnetic-driven actuation can be modeled as a simple rotational motion driven by a discrete torque; analysis for one simple geometry is given in studies by Syms et al. and Judy and Muller for fluidic and magnetic actuation respectively. Bilayer expansion models suitable for the thermally driven mechanisms are also widely available, for instance in Timshenko’s original thermostat paper. Laser forming has proven more difficult to capture with simplified models, but the review in studies by Shen and Vollertsen provides an overview of currently available analytical approaches.

Capturing the behavior completely, however, requires the use of computational modelling tools. One major example is in the area of fluidic self-assembly. To make good use of a fluidic hinge, it is very important to be able to develop insight into the resulting angle as the solder melts. Fortunately, modelling of fluids is of great value in many applications, and many tools are available for understanding the shape and forces of liquids. One example is the powerful software package known as Surface Evolver. Developed at the University of Minnesota and distributed as freeware, Surface Evolver performs rapid calculations of the surface energy of fluids and resulting liquid shape. Because these energies directly set the behavior of self-folded solder assembly, the software is ideal for modelling solder hinges and obtaining accurately simulating the resulting hinge angle using the known fluid properties and original size of the unmelted solder.

While analytical modelling for laser forming is challenging, thermomechanical modelling can capture much of the laser forming process. Laser absorption can be treated as a thermal input, and the approximate angle of deformation on a laser fold line can be estimated using conventional finite element programs. Unfortunately, however, the large temperature range and nonlinear effects such as plastic deformation and the resulting strain hardening complicate the behavior, particularly for large numbers of laser passes. While the plastic deformation within the materials makes more complete modeling of laser forming complex, there have been a number of efforts in the field, with a study comparing and contrasting different available computational methods.
It is also possible to design more complex bending patterns using these types of thermomechanical models. Optimization of the thermomechanical models in studies by Peraza-Hernandez\cite{133} were used for designing the fold patterns for shape memory alloy composite sheets. The technique used relied on targeting specific bending maxima and iterating the temperature inputs to minimize the positional error of the maxima. While there remain some challenges on fully understanding specific mechanisms or geometries, overall there is very good understanding within the field on predicting individual fold behavior. Precise actuator control is necessary for many applications beyond self-folding, and the community has been able to build on the associated resources. Moving beyond a simple fold, however, has been more difficult.

4.2. Origami Design

While structures with one or two folds can be useful, making use of the power of origami requires an understanding of how to make more complex structures to incorporate many folds to build a truly 3D geometry. Origami design has become an important academic discipline, and there have been great advances toward modeling and understanding the underlying mathematics, advances that the self-folding community can build on. A number of broad families of approaches have now been developed for approximating complex 3D shapes such as a curved surface. The first set of these relies on making the structure globally compliant by incorporating repeated folds, allowing the overall structure to be deformed into a desired shape. The most important example of this is through tessellation, repetition of a simple fold pattern across the entire sheet. Several of the best-known examples are the Miura-ori pattern (Figure 1b,c) and waterbomb pattern (Figure 10a–c). A related approach uses pleated folding, alternating mountain and valley folds,\cite{134} to obtain similar compliance and global shaping (Figure 10d). Tesselation is particularly powerful because it is a relatively intuitive approach to complexity. Because the individual tessellations are not overly complex, the folding behavior can be understood relatively easily, which is a major advantage for incorporating actuation and control. The Miura-ori pattern in particular is well known for its ability to transform from compact 3D volume into a flat 2D surface with a simple compressive force. Its ability to minimize volume makes it useful for transporting and deploying solar arrays on spacecraft.\cite{7}

While tessellation and pleating are powerful, they are also limited in complexity. Creating more intricate patterns becomes much more difficult to design intuitively. The second set of approaches instead relies on the use of computational design tools to both determine whether a structure can be folded and how to do so. Unfortunately, it turns out that solving the origami folding problem completely is very difficult, and it remains computationally unfeasible to determine the optimum folding in many cases.\cite{135} Obtaining at least one folding for a given polyhedron, however, is a simpler and more solvable, and one for which powerful tools exist.

Within computational origami, there are several techniques available for creating an arbitrary polygon. One of the simpler and more universal approaches for doing so is to create the polygon from a long strip of material, effectively building the structure surface layer by layer. It has been proven that this approach is universal and can be used to build any polyhedron,\cite{137} with a simple example (a folded alphabet) shown in Figure 10e). It does, however, have a major limitation; building a complex polygon from a long strip results in slits along much of the resulting surface, undesirable for a real engineered structure.

The more practical computational alternative is to use what are known as tucking molecules, that is, re-mapping the surface onto a plane and incorporating flat foldable segments (known as tucking molecules) in between to allow folding.\cite{136} Perhaps the
best-known and most widely used example of this approach is a freely distributed computer program known as the Origamizer.\cite{135} Unlike the strip approach, the Origamizer relies on targeting a property the authors have defined as “watertightness,” where there are no paper slits within the defined outer boundary of the paper. While maintaining this property limits the ability to obtain negative curvature vertices, the added tuckable features make these structures foldable.

Beyond designing hypothetical fold patterns, the realities of self-folding add an additional constraint. With manual folding of paper, it is possible to flex or bend individual segments to allow them to pass others during folding. Self-folding, on the other hand, cannot always rely on this ability. Because self-folding occurs without direct human involvement, when two segments collide, it will lead to one of two outcomes: either the segments will stop or one or the other will be damaged as it contacts the other segment. Both outcomes are clearly undesirable for self-folding, and designed complex folds that do not consider this constraint will fail. There has therefore been important research on incorporating collision-free folding into computational origami algorithms. For instance, one approach is to require individual folds to be linearly foldable,\cite{136} defined as having a collision free straight path between its original and final positions. Adding this constraint has a powerful impact on the folding performance, with demonstration that the resulting optimized patterns are almost 90 times faster to fold than more arbitrary designs.

Moving from computation to real self-folding structures brings in other considerations. One consideration is sheet rigidity; paper is flexible because it is thin and made from fibers, yet most metals of appreciable thicknesses are significantly less flexible. Sheet rigidity has therefore been incorporated into computational origami planning tools, for instance, in the Rigid Origami Simulator tool for use in modelling of shape memory folded sheets.\cite{139} Real actuation mechanisms and fabrication processes are far from ideal as well, and there have attempts to incorporate these imperfections into the computational design algorithms. For instance, a shape memory alloy driven origami process was computationally designed\cite{140} to include considerations of finite thickness and non-idealized beveled hinges, and was used to build self-foldable parts. Figure 11a–d shows the algorithm producing one pattern, a 3D dome.

There has also been a prominent effort to study the effects of folding sequence on reliability for solder-driven self-folding
Computational design algorithms were used to generate all possible folding sequences for the desired shapes. The algorithms factor in both solder hinges for actuation, as well as additional solder locations for locking the feature in place after folding. Representative patterns were then chosen, fabricated and allowed to self-fold, allowing a study of the fabrication yield for patterns with different properties. The study determined that patterns with relatively high compactness correlated with higher process yield for the same polyhedral.

Laser forming provides a different set of constraints, most notably in being a line-of-sight process; folding is only possible if the laser reaches the hinges. It is, however, possible to add manufacturing constraints such as line-of-sight into a computational algorithm as in. Starting with a 3D CAD file, the needed folding sequence was meshed and verified, followed by generation of a 2D SVG file with the laser path, which can be imported into the laser software and used to rapidly cut and fold a desired pattern such as an antenna design. Techniques such as these for directly feeding 3D files into the needed machine input have the potential to become the self-folding equivalent of software the 3D printing industry is built on, software such as the slicers for layering 3D designs for printer toolpaths.

Perhaps the most complete attempt at automated origami folding has been for programming of shape memory-driven sheets in studies by Hawkes et al., aimed at automatic reconfiguration and shifting between shapes. As in the earlier work, this begins with developing planning algorithms allowing understanding of the needed folds to shift between states. Just like the laser forming example, the next step is implementing in hardware;
in the shape memory sheets, this was done through the integration of heater resistors into the individual hinges and the development of control hardware designs to allow shifting between states.\cite{144} Sets of conductive connectors called stickers are used to program the actuation sequence. Figure 11e shows the complete approach, from origami planning to control of design inputs and circuit design. Although the mathematical and computational origami community has made great strides in understanding design of origami folding, incorporating into real self-folded systems has been a newer and less-explored field. Many self-folding specialists have come from the materials, fabrication, and mechanical design fields and have had less exposure to the kind of computational resources that can enable greater degrees of complexity. There is a significant potential for making these types of algorithms and design more widely available, as in the earlier examples, and this may be a major growth area in the coming years.

5. Applications

The most important question for self-folding origami, is “what can I do with it?” The answer is driven by understanding the benefits of an origami approach in general, and self-folding in particular. Origami at its heart allows a 2D form factor to be transformed into a more complex 3D one, and therefore provides the benefits of ease of storage and transport, simpler manufacturing, and possibility of reconfiguration.\cite{16,19} Self-folding provides further benefit of enabling these transformations in situations or size scales where it might be difficult, inconvenient, or impossible to fold by other means.\cite{19} Taking advantage of these benefits has opened up a variety of important applications from reconfigurable robots to tiny stents ballooning to hold open blood vessels.

5.1. Robotics

One of the most exciting application areas has been in the area of making new reconfigurable robot technologies. As with other applications of origami, the primary advantage of origami robots are for simplifying manufacturing by keeping the structure 2D until the final folding steps, allowing design complexity to be made independent of the underlying fabrication process.\cite{144}

The simplest origami robot component is a gripper, a structure intended to grasp and then release a target. While a gripper is a relatively simple geometry, creating one that is functional and useful requires bidirectional motion, i.e., the ability to close and open. One actuation mechanism with this capability is a copper oxidation-driven bilayer, which is developed using electrochemistry to selectively oxidize and remove copper oxide from a copper/chromium bilayer. The oxidation bilayer structure in studies by Randhawa\cite{85}, for instance, was used to fabricate an approximately millimeter size gripper (Figure 12a), and repeatedly open and close the structure for tens of cycles, and perform simple operations like grip micro-scale tubes.

Shape memory alloy actuators have been a particularly popular choice for robotic origami as a simple and bidirectional actuation mechanism easily controlled with resistive heating. Through this, there have been several approaches derived for making use of the SMA actuator for driving origami folding. In one variant, SMA actuators are used as an external force driving the collapse or folding of an origami sheet,\cite{145} while the actual origami structure is made of alternative materials such as laser cut polymers. By compressing and expanding the origami sheet, a simple worm robot can be created, sliding across surfaces at speeds as high as tens of millimeters per minute. Adding additional shape memory actuators to give more direct control over hinged plates allows even more complex motions with worm robot bodies.\cite{146} Moving beyond worm geometries, a folding shape memory-actuated rolling and jumping robot has also been demonstrated, namely the three-armed Tribot.\cite{147} Shape memory wires are used to rapidly deploy the Tribot from its flat orientation as well as to jump by more than 200 mm and roll using controlled leg motion.

Although external springs can be a powerful approach, direct folding through the hinge itself can allow more powerful and versatile motion, and shape memory alloys have again been an important mechanism. One notable example has been the use of self-folding structures for the purposes of programmable matter, defined as materials with reconfigurable sub-elements used to change shape or stiffness.\cite{19} The demonstrated programmable matter in this case can be thought of as a dramatic case of

**Figure 12.** a) Oxidation-driven microgripper (scale bar 500 μm). Reproduced with permission.\cite{85} Copyright 2010, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. b) shape memory driven programmable matter assembling into origami shapes. Reproduced with permission.\cite{148} Copyright 2014, Elsevier.
reconfigurable origami. Shape memory alloy hinges control an array of 1 cm rigid elements, allowing the sheet to completely redefine itself in a few seconds, going from a flat sheet to other origami shapes such as a folded “shuttle” or a “hat” (Figure 12b).\cite{143}

From this starting point, a number of efforts have been made to create more functional robotic systems. One of the key needs for proper control is the integration of positional sensing into the origami folds. In studies by Firouzeh\cite{148}, strain-sensitive resistors made of carbon black-loaded rubber were integrated with shape memory alloys. Adding strain measurement allows feedback and allows controlled motion. Further development and integration of control circuitry allowed this technique to create a stand-alone integrated robot.\cite{149}

Origami robots have become one of the most exciting areas within metal-driven self-folding and remain an active discipline of research. There are, however, sizeable challenges in the field, notably in the best means for miniaturization to the micro and millimeter scale, improved sensing and power management, and more complete computational modeling capabilities.\cite{144}

5.2. Electrical Devices

Origami robots rely heavily on bidirectional actuators such as shape memory alloys, with the intent of reconfiguring a folded structure. The unidirectional folding mechanisms, on the other hand, are primarily used in fabrication and manufacturing, and the best-known application is in the creation of small-scale electrical devices. Self-folded electrical systems are an important example of origami self-folding for low-cost and simpler manufacturing, and one where metal conductors are fundamental to functionality. Large-scale versions of many common electrical passives such as inductors and antennas are complex 3D structures. Folds\cite{150} or corrugations\cite{151} on an antenna can reduce size, for instance, and many possible coil designs for inductors such as toroids are fundamentally 3D. Microfabricated versions, however, are often limited to two dimensions due to fabrication constraints. Self-folding has opened up a new design space to the electronic designer as a powerful approach to generate these 3D designs, and their corresponding advantages, while limiting additional design complexity.

Inductors are common electrical passive devices that store energy in the form of magnetic fields generated by currents through coils of wire. They are often incorporated into integrated circuits (ICs) as components of power, communication, and filtering systems. Large-scale inductors are made through wrapping enamel-coated wire around a preform or magnetic core. In an era of increasing miniaturization, however, inductors are often integrated directly onto silicon in system-on-chip designs. One challenge, however, is that the magnetic fields generated by the inductor also can generate losses within neighboring conductors and semiconductors. This includes the underlying silicon of many ICs, limiting performance of power devices on-chip. While one possibility is to etch away the underlying semiconductor to minimize these losses,\cite{152} this approach is destructive and eliminates possible regions that could instead be used to make active electronics.

A clever alternative is to instead fold the inductor up, away from the silicon\cite{153,144} using a micromachined process based on electroplated solder hinges and copper coils. During fabrication, the system is held in-plane for ease of processing; however, after processing, the inductors are heated to a moderate temperature (Figure 13a), causing the solder to melt and reflow, applying surface tension to the inductor. The inductor then folds up out of plane to a final position that is roughly vertical (Figure 13b), minimizing coupling to the substrate. Inductors are nonradiative devices, meaning that most of the magnetic energy is kept relatively close to the structure. Flipping an inductor on its side is sufficient to significantly reduce the fields in the silicon substrate. As one measure of the resulting performance, the quality factor (the ratio of the stored energy to the resistive losses of the inductor) was characterized before and after folding, showing a factor of five increase in quality factor.\cite{154} Similar folded-up spirals have also been demonstrated using other self-folding approaches such as magnetic assembly.\cite{155}

While the folded-up inductor is a clever approach, the actual inductors are still classic 2D designs, serpentine and planar coils, which are relatively easy to make using standard microfabrication approaches. Self-folding origami is also a promising avenue for creating truly 3D inductor designs on chip. A solenoid (an inductor where the turns are wrapped in a cylindrical shape) is a traditional 3D inductor, commonly used in macroscale wire-wrapped inductors but far less so in microfabrication. This structure is particularly difficult to make on-chip, at least with a large area enclosed by the turns of the inductor, due to the limited thicknesses of individual layers. It is, however, possible to create such a shape by patterning and depositing the traces in-plane, followed by folding traces up to create a solenoid with a large enclosed core area. One way of doing this is through PDMA, using an applied magnetic field to fold up electrical traces.\cite{156} The inductor traces consist of gold traces, with the traces intended to be out of plane coated in a layer of permalloy to give a magnetic response. When a magnetic field is applied, the traces fold up to vertical to create a solenoid structure. The same work also demonstrated that electroplating of copper was possible after folding, allowing the hinges to remain relatively thin to allow folding and then thinned to reduce electrical resistance for improved performance.

While the creation of high-performance 3D inductors on chip has been one driver for using self-folding, there are also possible advantages in larger scale inductor manufacturing. Larger inductors are widely used, particularly as a component of power supplies and converters. One of the most commonly used inductor designs is known as a toroidal inductor, which corresponds to wrapping metal wires around a toroid (donut)-shaped preform. Wrapping an inductor is tedious without specialized equipment, and being able to rapidly manufacture quality toroids in a machine shop or laboratory environment could be a significant advantage. As with a solenoid, however, toroids are fundamentally 3D and therefore relatively difficult to make using conventional 2D processing. However, self-folding can be used to create high-performance toroids using simpler 2D tools.

Laser formed origami inductors\cite{156} are notable examples of this approach, by using a laser cutter for rapid folding of electrical components. With laser patterning, it is possible to both cut and self-fold from blank metal feedstock. Traces are first cut in
copper sheeting (Figure 13c), followed by folding them vertically up out of plane (Figure 13d) and further cutting to define the remaining inductor and release from the foil (Figure 13e–h). Rather than building up the vertical portion of the toroid, as would be necessary using alternative techniques such as 3D printing or micromachining, laser forming enables the vertical traces to be rapidly folded using the laser itself. The inductor was then measured (Figure 13i) and demonstrated to have performance and quality factor competitive with commercially available inductor designs.

Antennas are a necessary component for radiofrequency (RF) communications, and as with inductors, self-folding is a promising avenue for creating complex 3D versions at small length scales. Inductors and antennas are similar in many respects, and the laser forming technique shown above has also been demonstrated for cutting and folding antennas as well.\textsuperscript{[158]} There have also been notable efforts to use self-folding to fold 3D antennas at the micrometer scale. In studies by Mao et al.\textsuperscript{[158]}, ion-induced stresses were used to create repeated fold lines in 3 μm wide strips of aluminum to approximate one of the most common antenna designs, a helical antenna of less than 2 μm in diameter. In addition, self-folding polymer sheets can also induce films of copper conductors to fold out of plane. For instance, the polymer-driven structure shown previously in Figure 8, a shape memory polymer actuator driving a thin piece of copper foil, is being used to create a reconfigurable antenna. Folding the copper foil up out of plane switches the design between two common antenna structures, a flat “microstrip” to a “monopole” antenna. The resulting changes in geometry change the antenna performance, thereby allowing to develop shape reconfigurable antennas that respond to light.\textsuperscript{[90]}

Although less common, complete circuit boards can also be folded remotely for compactness or new capabilities. It is possible, for instance, to fold a populated flexible printed circuit board, a copper/polyimide stack, by using compressive buckling assembly.\textsuperscript{[159]} The resulting folded system incorporates a number of useful capabilities including a microcontroller and near field communication (NFC) coil.
5.3. Biomedical Applications

For robotics and inductors, ease of fabrication has been a primary driver for relying on self-folding techniques. The next example focuses on a second advantage: the ability to go from compact form factor for ease of transport or storage to a more complex and functional unfolded state. Perhaps the most important example is being able to unfold from small packages for use in the human body. The body is filled with narrow passageways from blood vessels to the digestive tract providing routes to critical organs while limiting surgical trauma, and a variety of biomedical origami devices and structures have been reported. An important example is the origami stent graft.\cite{160} Stent grafts are deployable devices that can be folded into small dimensions to move up tubes such as the veins and arteries of the cardiovascular system, then unfolded in situ to open blocked blood vessels or treat/protect damaged patches. In the stent in,\cite{160} NiTi shape memory alloy is used to create an origami structure that expands at human body temperature from the smaller diameter of a delivery catheter into a larger diameter upon delivery (Figure 14a–i). NiTi is chosen due to both its shape memory characteristics and its biocompatibility.

A stent provides a simple expansion/contraction motion, but more complex movements are also possible. In studies by Salerno et al.\cite{45}, an origami gripper designed for use in minimally invasive surgery was demonstrated, taking advantage of origami reconfigurability for biomedical applications. Shape memory actuators are used to provide fine control of segments, allowing four degrees of motion within the surgical actuator. For the purposes of assisting or performing surgical operations, similar SMA actuators are again used for controlling a hinged gripper for manipulation.

In addition to use in minimally invasive surgery, an alternative application is in drug delivery. Self-folded boxes, cubes, and other containers of length scales from the nano to millimeter scale or larger are used as examples throughout the origami field. Using those containers to carry a medically active agent is relatively straightforward. In studies by Park et al. and

Figure 14. a–j) NiTi origami stent graft as it is placed using a smaller tube and allowed to unfold to full size within a larger channel. Reproduced with permission.\cite{160} Copyright 2006, Elsevier. j) Controlled chemical release from a folded micro-container. Reproduced with permission.\cite{162} Copyright 2006, American Chemical Society.
Leong et al.\cite{161,162}, melted solder hinges were used to fold nickel walls around a dissolvable gelatin material. The use of magnetic nickel then allowed magnetic fields to be used to guide the resulting delivery vehicle to a desired location. Both chemical environment\cite{162} (Figure 14j) and inductive heating\cite{161} were demonstrated to trigger dissolution of the payload.

Due to its ability to rapidly unfold from compact packages, biomedical origami has been a major area of research for much of the community, and one with a particularly large potential impact if successful. There does, however, remain very significant challenges, largely due to the high standards necessary for any device deployed inside the human body. Controlling complex surgical tools remotely is a very difficult problem, and there is also room for improvement in material choices and patterning.\cite{163} For real medical use, the final product must be safe, biocompatible, and easy to use;\cite{164} and achieving the needed regulatory approvals is perhaps the largest and most expensive barrier to overcome.

5.4. Origami Manufacturing

In addition to more specialized applications, there has also been interest and significant work toward making origami folding into a more general process for creating 3D parts. Much of this work has been inspired by the success of additive manufacturing, also known as 3D printing. 3D printing has been revolutionary in many ways in the manufacturing sector, allowing complex and custom 3D structures to be rapidly prototyped, often at lower cost than traditional technologies.\cite{165} There are, however, several important limitations of 3D printing, particularly in the case of creating metal structures. Commercial metal 3D printers, which are based on technologies such as selective laser sintering of metal powders, are generally very expensive, often marketed at over half a million dollars.\cite{166} The high cost inevitably has limited use outside of larger scale commercial manufacturing.

Because 3D printers build parts layer by layer,\cite{167} additive manufacturing is also slower than origami folding for the creation of shell-type structures, with thin features jutting up from the build plane.

One of the key advantages of 3D printing is the ability to rapidly create custom parts ("mass customization")\cite{165} rather than numerous copies of an identical part as in traditional manufacturing. There have been notable efforts toward attempting general purpose manufacturing using self-folding metal technologies. Laser forming is the most commonly used, largely due to its ability to be easily combined with laser cutting to both pattern and fold using the same piece of equipment. In one early example, steel sheeting was cut and formed into practical parts for a vehicle, including cover plates and a scale model door panel (Figure 15) through manual alignment and flipping.\cite{168} Building on this work, other research studies demonstrated the possibility of using a motorized stage for positioning and flipping during the laser cutting and folding process to obtain up and down folds\cite{69} or using the TGM and BM modes together with laser cutting to create complex parts without repositioning the part.\cite{68}

For the purposes of rapidly creating a series of custom parts, one of the most important challenges is feedstock. To build up parts, 3D printing makes use of some form of general purpose material such as filament, a pool of liquid resin, or a bed of powder to create the required volume. Building repeated parts with self-folding origami also requires rapid bed clearing and material replacement, and one way to do so is to shift to a roll-to-roll format. Roll-to-roll processing\cite{169} is a widely used, high-throughput technology in printing, where material from one roll is sent through a series of processing steps and then recollected onto a second roll (a take-up roll). A roll-to-roll laser forming origami process is shown in Figure 15b.\cite{63} Metal foil off of a one roll are fed through a foil guide that holds the foil in the focal plane of the laser system. The laser is then able to cut and fold a series of parts

![Figure 15](image-url). a) Laser cut and formed 1:8 scale steel car door panel. Reproduced with permission.\cite{168} Copyright 1997, Emerald Publishing Limited. b) Roll-to-roll laser forming setup and c) after cutting and folding of series of parts. Reproduced with permission.\cite{63} Copyright 2018, Springer Nature.
(Figure 15c), followed by performing a final cut to remove them from the foil and allow them to fall down a chute to the user. The foil is then taken up by a second motorized roll used to pull the foil through the system.

Creating assemblies with free moving parts is another powerful capability for a manufacturing process. Laser cutting and forming was used to position and cut a simple moving part, a flywheel with axles, to drop into position in folded axle holders.\cite{170} The part remained freely able to move and was successfully spun remotely with the laser cutter itself using laser ablation propulsion, where the jet of ablated material is used for force generation.

Despite significant promise, origami manufacturing remains far from the mature technology that 3D printing has become. A casual user can draw up a CAD file, send it to a 3D printer, and reasonably expect their part to appear in minutes or hours, and origami-based processes to move beyond a highly specialized niche it must begin to do the same. The biggest challenges are in the computational design space and in improved process control to make origami-driven processes sufficiently stable and repeatable for general use.

5.5. Optics

Self-folded optical systems have much in common with their electrical counterparts. As with electronic components such as antennas, many optical systems rely on a third dimension, and self-folding can be a convenient, low-cost method for moving parts out of plane. At large size scales manual positioning is relatively easy, but aligning a MEMS mirror or filter can often be done much more easily and consistently using self-folding.

Perhaps the simplest example of this is the creation of an angled mirror reflector.\cite{68} In many marking lasers, there is only a single write head, allowing patterning only from the top surface, while the bottom surface is inaccessible without flipping over the part. In the marking laser used, the laser was designed to focus in a single cutting plane, which is common for systems designed to mark flat metal or polymer surfaces. Laser forming was used to rotate up a flap of metal to $45^\circ$ to create a mirror element focused at the vertical plane at the base of the mirror.\cite{68} Additional laser forming is then used to bring a portion of the previously inaccessible underside against the mirror and to use the reflected laser light to etch the letters “ARL” into the foil, followed by moving away for further assembly into a cube.

More complex and functional features are also possible using self-folding. The micromirror in studies by McCarthy et al.\cite{171}, for instance, uses a solder reflow hinge to move a MEMS mirror up out of plane. Unlike the previous mirror reflector, however, the refloved solder mirror is designed to use several solder balls to position a mirror frame and actuation electrodes at desired angles (Figure 16a), while the mirror itself retains ability to rotate on torsional springs. Electrostatic actuation then used once the mirror is in position to provide fine positioning for light control. Arrays can also often be created relatively easily using self-folding; in one example, compressive buckling driven folding has been demonstrated for assembly of arrays of reflectors with controllable angle for an optical transmission window.\cite{172} Alternatively, shape memory alloys can allow bidirectional control of optical elements. Shape memory alloy wires controlled through resistive heating have been used for beam steering of several origami and kirigami reflector designs.\cite{173}

Beyond simple reflectors, more complex optical systems have also been created using self-folding. One notable example is magnetically assembled devices.\cite{174} Based on traditional micromachining techniques, magnetic permalloy is electroplated on the ends of rigid plates of polysilicon before final release. After release, the plates are left free on rotary hinges, allowing rotation to vertical upon exposure to magnetic fields above a designed threshold. Actuation of additional plates with appropriate slots are then used to lock the structures in place, allowing precise positional control. Using this approach, a series of common optical devices were demonstrated, including a Fresnel lens (Figure 16b), an optical grating, and a corner cube reflector.
5.6. Sensors

Sensors have been another area of interest for the self-folding community. Many important physical quantities have a directional component, and obtaining complete information requires knowledge about the out-of-plane direction. Self-folding can be a convenient mechanism for rapidly creating large vertical displacements for MEMS sensors, and as with electronics and optics, the primary advantage is in reduced fabrication complexity. Folding is therefore used to obtain a third axis of measurement without resorting to building up large stacks using layered microfabrication.

Flow measurement is one interesting example. Placing a chip with a flow sensor on the edge of a fluidic channel will not give as much flow information as if the flow sensor juts out into the middle of the channel. Being able to place an array individual flow sensors at different heights from the edge is also important for learning not just the flow at one point, but the complete flow distribution. This is difficult using conventional micromachining, where individual layer thicknesses must be well defined and consistent across a chip or wafer. For self-folding, on the other hand, different lengths can be patterned in plane using standard patterning methods such as photolithography and then folded up to achieve different heights. This method was used to make arrays of one common type of flow sensor, a hot-wire anemometer (a suspended and heated wire whose temperature reaches different levels depending on flow due to variations in convective cooling) before folding, the structure is plated with permalloy, allowing it to be folded using magnetic assembly.

Many cleanroom deposition processes are directional, and planar measurements of the deposited thickness can neglect important information about sidewall coverage and other parameters. One way of measuring deposition thickness is through changes in frequency of a resonant device. Deposited material both increases stiffness (causing the resonant frequency to rise) and increases the mass (causing the frequency to fall), and which behavior is dominant (and whether the frequency ultimately goes up or down) dependent on the individual design geometry. A three-axis self-folded deposition monitor based on cantilevers aligned in each orientation was demonstrated and used to measure copper deposition rates. Deposition of material resulted in a stiffer cantilever, resulting in a shift upward in frequency as more copper was deposited. The monitor was based on solder hinges for folding and nickel cantilevers to create the individual sensing elements. The sensor was then used to monitor evaporation of copper, a highly directional deposition process, and was able to demonstrate clear variability in deposition on the individual elements depending on orientation relative to the line-of-sight of the evaporation.

Sensing can also be incorporated for added functionality within a more complex system. Light is also highly directional, the primary reason for self-folding in optical systems. In a final example, light sensing was one of the major functionalities of the folded electronic system made using compressive-buckling-based self-folding. Multiple photodiodes are the primary sensor input for the system, with demonstrations of variations in response before and after folding up relative to a directional laser input. This input then fed into data processing circuitry for near-field communication, sending the sensor signal to the user.

Overall, the field of self-folded metal sensors comprises several very promising demonstrations, but is not as mature as many other application areas discussed here. There is significant scope to grow, particularly in identifying new sensing modalities that benefit from being able to place sensing elements out of plane.

6. Conclusion and Outlook

This review is intended to provide a survey of the field of self-folded metal origami, including design, material selection, and potential application areas. Self-folding of metals has opened up new opportunities from creating cheap and reconfigurable robots, electronics, sensors, and optics while enabling rapid fabrication of complex 3D structures at the micro- and nano-scale. While metals have desirable mechanical and thermal properties, this field is very much driven by the value of metals as electrical conductors, with orders of magnitude higher electrical conductivity than most other materials. Metal self-folding has therefore been the most important in applications such as electrical devices and robotics where the ability to move electrical signals efficiently is fundamental to the operation of the device. These electrical

Figure 17. a) Self-folded three-axis mass deposition sensor and b) frequency shift with copper deposition. Reproduced with permission. Copyright 2008, AIP Publishing.
advantages have helped metals hold their own in these applications, even though polymer actuators are typically more responsive and easier to drive to large displacements. It does, however, help to explain that metals have not been as competitive in areas like sensors where polymers are often far more active than their metal counterparts.

There have been great advances in self-folding metal origami, particularly in the past decade, and many recent developments have been highlighted here. There remain significant challenges and opportunities within the field going forward. Maturing the field, and in particular getting self-folded origami from the laboratory to the marketplace, is a difficult problem, and one that will likely require much collaboration, often between groups that do not normally communicate. For instance, there has been wonderful work on creating algorithms for complex origami design within the computational community, adding powerful new functionality and showing great future promise. The MEMS/NEMS community has similarly developed a variety of useful fabrication approaches for self-folding at size scales where alternatives to self-folding are far and few between for making complex 3D structures. Surveying the literature, however, the vast majority of practical MEMS self-folded structures are simple, with a small number of folds to perform a basic function like moving away from the substrate. Origami algorithms with idealized views of folding and material behavior are more common than ones that feed in the real constraints of fabrication approaches, and improved communication is likely to be the best solution. The groups able to bridge this gap, and to identify clear scenarios where the complexity is valuable at small size scales, will likely have the most impact going forward.

Similar opportunities are possible in the materials realm. For instance, it is quite reasonable that those researching self-folding magnetic actuators have largely limited themselves to a few classic, easily processed magnetic materials. When investigating new actuator designs and complex fabrication processes, it is reasonable that researchers have chosen not to try exotic and unfamiliar materials as well. A variety of exciting new devices have emerged from this strategy. Optimizing performance and developing these technologies further will require incorporating state-of-the-art magnetic materials, taking advantage of decades of materials development and process development. Similar statements could be made about many of the other material classes. There are likely solders that could push the boundary on fluidic folding forces or conductivity or cutting edge shape memory alloys with improved elastic or thermal properties. Collaboration between the self-folding designers and the materials community will enable better performance and open up new application areas.

A final major opportunity is in integration and system complexity. The robotics community has been the leader in adding control circuitry and other additional functionality driven by folding metals, but this has been easiest to do at larger size scales and very understandably the community has chosen to focus their efforts at this scale. At some levels the path is clear; the circuits community has for decades prioritized system-on-a-chip implementations, and shifting from circuit boards to integrated circuits with active hinges, the obvious path to miniaturization, is certainly possible. Integrated circuit processes are, however, also very expensive and often difficult to modify without a pressing need. The members of the community able to identify such a need and commercial market sufficiently large for self-folding to be taken seriously by the commercial semiconductor foundries will similarly have a prominent role in making this technology practical. There will also be major challenges in making self-folded origami processes sufficiently stable. A thorough study of yield and manufacturability of folding complex polygons using liquid solder assembly, for instance, showed yields as high as about 80%. While this result is impressive for a laboratory environment, significant process and design improvements will be necessary for commercial use.

Even though these challenges are significant, they are also straightforward and likely can and will be addressed as the self-folding metal community matures and grows. There is, however, grounds for optimism, and very real promise that self-folding will emerge from being a small but useful niche technology into becoming a broader and more universal one. The field builds on decades of materials, physics, and controls innovations, and making full use of these resources will allow major strides forward.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

active materials, actuation, metal folding, origami, self-folding

Received: June 26, 2019
Revised: July 30, 2019
Published online: August 29, 2019

[1] N. Turner, B. Goodwine, M. Sen, Proc. Inst. Mech. Eng., Part C 2016, 230, 2345.
[2] W. Wu, Z. You, Proc. R. Soc. A 2011, 467, 2561.
[3] Y. Liu, J. Genzer, M. D. Dickey, Prog. Polym. Sci. 2016, 52, 79.
[4] D. Rus, C. Sung, Sci. Robot. 2018, 3, 0938.
[5] J. Rogers, Y. Huang, O. G. Schmidt, D. H. Gracias, MRS Bull. 2016, 41, 123.
[6] E. Hawkes, B. An, N. M. Benbernou, H. Tanaka, S. Kim, E. D. Demaine, D. Rus, R. J. Wood, Proc. Natl. Acad. Sci. U. S. A. 2010, 107, 12441.
[7] R. J. Lang, Phys. World 2007, 20, 30.
[8] R. J. Wheen, R. Q. Bridge, J. Ferrocem. 1981, 11, 311.
[9] Q. Shi, X. Shi, J. M. Gattas, S. Kitipornchai, J. Constr. Steel Res. 2017, 138, 235.
[10] S. Ahmed, C. Lauff, A. Crivaro, K. McGough, R. Sheridan, M. Frecker, P. von Lockette, Z. Ounaies, T. Simpson, J. Lien, R. Strzelec, Proc. ASME 2013 IDETC-CIE, Portland, OR, USA 2013.
[11] J. M. Jani, M. Leary, A. Subic, M. A. Gibson, Mater. Des. 2014, 56, 1078.
[12] D. H. Gracias, Curr. Opin. Chem. Eng. 2013, 2, 112.
[13] R. Resta, J. Phys.: Condens. Matter 2002, 14, R625.
[14] K. Lu, Science 2010, 328, 319.
[15] R. J. Lang, K. A. Tolman, E. B. Crampton, S. P. Magleby, L. L. Howell, Appl. Mech. Rev. 2018, 70, 010805.
[16] C. D. Santangelo, Annu. Rev. Condens. Matter Phys. 2017, 8, 165.
[17] J. Cai, X. Deng, J. Feng, Y. Zhou, Smart Mater. Struct. 2015, 24, 125031.
[170] N. Lazarus, A. A. Wilson, G. L. Smith, *Extreme Mech Lett.* 2018, 20, 46.
[171] B. McCarthy, V. M. Bright, J. A. Neff, *Sens. Actuators, A* 2003, 103, 187.
[172] F. Zhang, Z. Fan, Y. Zhang, *Int. J. Solids Struct.* 2018, 147, 254.
[173] W. Wang, C. Li, H. Rodrigue, F. Yuan, M. Han, M. Cho, S. Ahn, *Adv. Funct. Mater.* 2017, 27, 1604214.
[174] Y. W. Yi, C. Liu, *Sens. Actuators* 1999, 78, 205.
[175] J. Chen, C. Liu, *Microelectromech. Syst.* 2003, 12, 979.
[176] J. H. Cho, S. Hu, D. H. Gracias, *Appl. Phys. Lett.* 2008, 93, 043505.